IJCSE International Journal of Computer Sciences and Engineering Open Access
Research Paper Volume-4, Issue-9 E-ISSN: 2347-2693

Enhancement of An Algorithm To Extract Text-Lines From Images For Blind And Visually Impaired Persons Through Parallel Approach.

Niketan Jivane[1*], Supriya Jivane[2], S.Rajkumar[3] and K.Marimuthu[4]

[1*]Department of Computer Science and Engineering
[1,3,4] Vellore Institute of Technology, Vellore, Tamilnadu, India
[2] M.B.E. Society's College of Engineering, Ambajogai, Maharashtra, India

Available online at: www.ijcseonline.org

Received: 22/Aug/2016 Revised: 02/Sept/2016 Accepted: 20/Sept/2016 Published: 30/Sep/2016

***Abstract*— Applications with Text Extraction form image or videos are boon for people with blindness to assist them in their day to day life. The emergence of high resolution cameras on mobile devices can be used to extract text-line. The literature presents that there are a number of algorithms that have been proposed for the extraction of text-lines from images. The Computation cost of these methods is very high and failed to address the challenging problem in text extraction due to the scale and orientation of the characters. In this paper, we propose a method to enhance an algorithm for quick retrieval of text from image documents using Parallel Matlab. Finally, it demonstrates the improvement in extracting text lines from the existing algorithm through results.**

Keywords- component; Connected component based algorithm; Contrast segmentation algorithm; Text Recognition; Edge Detection; Texture based methods; Text Localization; Precision and Recall; Stroke width transforms (SWT); Color Polarity computation; Adaptive thresholding; Modified Dam point labeling; Inward Filling.

I. INTRODUCTION

Extraction of Text-Lines form image is very important steps for various image processing tasks such as analysis of layout, image segmentation, optical character recognition (OCR). This embedded text can provide valuable information. Wrong Segmentation and extraction of characters may produce bad accuracy. The text extraction from images is very challenging because of the availability of degradations, various types of noises, and differences in script characteristics and writing style. The orientation and scale of text are also varied, the distance between two characters is irregular, and chances of character touching to each other's are very high. In order to give solution to these problems, many algorithms have been proposed in past few decades. However, existing algorithms fails in both accuracy and speed up.

Most of the conventional work focused on specific sets of characters. That is, the recent past algorithm takes the consideration of variations which are caused by individual writers by using language-related features. Such as in some images two characters are usually separated from each other so using connected component (CC) based approach, then the connected components are extracted and portioned into text lines.

The recent technologies are mostly intended to extract text from images with clean backgrounds. So it is necessary to extract text from images with complex natural backgrounds, or images with general backgrounds. There are lots of applications where we can use this extracted text. These applications consist of digital libraries, information retrieval systems, various multimedia systems, computer vision, and in various geographical systems. The main feature of text detection is to find the regions of images which contain text and then extract that text. This extracted text can be directed to the user or can also be given to an optical character reader model to recognize the text.

There are two main methods that are widely used to determine spatial cohesion. These methods are based on connected component [1, 8] and characteristics of text edge [6, 7]. Visual Perception is very important aspect to detect text regions [8]. An image is divided into different categories according to whether or not it contains any text data. This fact can be used to classify the text regions. There are several methods which use classification techniques such as Artificial Neural Network [9], k-means Clustering Method [10], and Support vector machine (SVM) [11, 16].

As stated in [10], each text has some similar typical characteristics in terms of frequency and orientation of the characters and also has spatial cohesion. The meaning of spatial cohesion is referring to the fact that characters of same string appear close to each other and have similar height, spacing and orientation [10].

There are various algorithms based on colour clustering, the general steps consist of pre-processing of the input image to remove any noise from image is present. Then the grouping

of images according to different colour layer and gray component is done. This is done because it is assumed that colour data in text characters are different from that colour data in background of the image. Then the potential text regions are identified using connected component based approach from these colour layers. Then value of each row and column is analyzed using aligning and merging technique. This method works well for English and Chinese characters, but there are problems due to the uneven presence of light and due to reflection in the test images.

The algorithm discussed in [12] also uses colour continuity as important feature for text extraction. It uses the wavelet transforms and it also joins high level image features with low level image features for extraction of text regions. Here bottom up approach is used. Here chip generation process is used. This process uses the spatial cohesion property of each character. The chips are nothing but collections of pixels in the image. The drawback of this algorithm is it does not detect very small characters.

This paper looks into the extraction of text from images using a connected component based method. In our algorithm we can extract text from images with enhanced performance and greater accuracy. The algorithm uses various geometrical properties of images such as aspect ratio, brightness, pixel distance, etc... It also uses factors like dimension of lights, orientation of image and image scale. The accuracy, precision and recall rates are analyzed to give the success and limitations of the approach. The detection and extraction of text from images are done using parallel processing techniques so as to increase the speed of the system and to explore the maximum computation power to reduce wastage of CPU power thus enhance the performance.

The research paper is categorized and structured as follows. Section II presents the literature survey. The proposed method is discussed in Section III; the detailed Implementation and result analysis is presented in Section IV. The conclusion is given in Section V with future work in Section VI.

II. RELATED WORK

There have been a lot of researches in this area, and a number of algorithms have been proposed for the extraction of text-lines in images as well as text documents and videos. Some of them are as follows:

The two state frameworks for text extraction were proposed by A. Mosleh. al. [3]. Here the text is located using unsupervised clustering, which is implemented on connected components which are generated by Stroke Width Transform (SWT).

Yen Lin Chen.al.[6] in his proposed an algorithm for text extraction based on decomposing the image into various distinct planes. These planes include text regions, non-text regions, background regions, etc. Here knowledge based approach is used to extract text from images. Then in painting is done after the text removal.

Novel Method of text extraction is proposed by Won Jun Kim al. Based on frame for overlay. Here in first step transaction map is generated. Then the text regions are extracted based on occurrence of that text in each text region. In next step using localization the text is extracted.

In 2005, M. R. Lyu al. , J. Song al. , and M. Cai al.[7] while doing their Ph.D work in Chinese University of Hong Kong in Shatin, Hong Kong proposed a comprehensive method for multilingual text detection, localization and extraction. Basic three methods collectively perform text detection: edge detection, local thresholding and hysteresis edge recovery. Then coarse-to-fine localization scheme is applied. And at the last, adaptive thresholding, dam point labeling and inward filling are applied for the text extraction.

In his research work, J. Malobabic al. explored method of detection of text using measures such as horizontal difference and morphological processing. He used Wolf-jolion algorithm [8, 9, and 10] for enhancing the results of text extraction.

In 2012, Mohammad Khodadadi and Alireza Behrad al.[11] of Shahed University, Tehran has proposed an algorithm for subtitle detection, extraction and in painting in a color image. Using stroke filters, new segmentation and verification algorithm based on image profile a text is detected. Background and text color in candidate block are estimated using color histogram. At the last in-painting algorithm based on matching algorithm is used to reconstruct the initial image contents in text areas.

In 2004, Ralph Ewerth working in Jena University of Applied Sciences, Germany with his Ph.D. work discovered that text localization and recognition in images is important for searching information in digital photo archives, databases, video and web sites. However, since the text is often printed against a complex background, it is often difficult to detect. So he found a robust text localization approach, which can automatically detect horizontally aligned text with different sizes, fonts, colors and languages.

Sneha Sharma.al., proposed methods the algorithms which use a set of images of natural scenes that vary along the dimensions of lighting, scale and orientation and compare the two basic approaches for text extraction in natural images edged based and connected component based method. And further accuracy, precision and recall rates for each approach

are analyzed to determine the success and limitations of each approach. And gave recommendations for improvements are given based on the results [12].

Umapada Pal.al.[13] who is currently working with Computer Vision and Pattern Recognition Unit, Indian Statistical Institute, India, has given a model for automatic identification of English, Chinese, Arabic, and Devnagari and Bangla script line.

Dr.M.Hanmandlu, working as Professor at IIT Delhi, India has proposed fuzzy model based recognition of handwritten Hindi characters in images and videos. He along with M. Ramana Murthy and Madasu carried out this work at IIT Delhi. This method is useful for Developing OCR for Hindi Language.

U. Bhattacharya working with Computer Vision and Pattern Recognition Unit, Indian Statistical Institute, India proposed a comprehensive method for multilingual text detection, localization and extraction. He has discovered the optical character recognition system for images containing Devanagari and Bangla languages.

Dr.V.Kavitha of Anna University, Tamilnadu, India proposed a novel framework for video text detection and removal by generating a morphological binary map after calculating difference each candidate which are gets generated by connecting candidate regions using a morphological dilation operation. Then, the text localization takes place by projection of text pixels in morphological binary map and last step is for text extraction.

Dr. S. Gopinath who is currently working with ABB GISL, Corporate Research, and Bangalore, India has proposed an automatic text extraction, removal and in painting in complex document images, by decomposing the document image into distinct object planes such as textual regions, non-textual objects, background textures etc. The texts with different characteristics from each object plane are detected using knowledge based text extraction and identification. Then an effective adaptive in painting neighborhood adjustment scheme is applied immediately following text removal.

So, all of these researchers have given various approaches for text extraction as well as text removal. Some of them used simple images; some of them used images with complex background. So all these approaches uses various machine learning approaches and optimization techniques which includes supervised and unsupervised learning, Artificial Neural network , Method of Belief Propagation and Conditional Random Fields(CRF) etc

*Corresponding Author:
C. T. Lin
e-mail: ct.lin@hotmail.com , Tel.: +00-12345-54321

III. DESIGN AND METHOLOGY

A. *FLOWCHART*

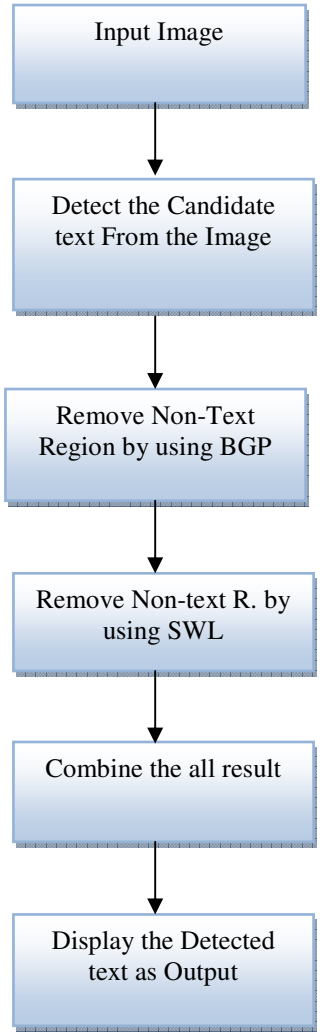


FIGURE 1. Procedure or flow chart of Text Extraction from Image

B. *ALGORITHM*

1. Give the Input images which contain the text Region.
2. Detect the Text region of the images with the help of the MSER function.
3. After the text detection next step are the Detect and Remove the Non-text Region of the images with the help of the Basic Geometric property.
4. Next are the Detect and Remove the Non-text region with the help of the Stroke width Length.

5. Combine the final text detected result.
6. Recognize the text with the help of OCR function.

These are the major part of our project which contains to extract the text from the unstructured scenes. The unstructured scenes of the images that include undecided or unstructured scenario for example, you can detect and recognize text automatically from captured video to alert a driver about a road sign. This is different than structured scenes, which contain known scenarios where the position of text is known beforehand. Segmenting the text from unstructured scene is mostly helpful with the additional task such as optical character recognition (OCR). There are many automatic text detection algorithms that detect the text from the images.

a. Detect the applicant text with the help of MSER.
b. Detect and Remove the Non-text region with the help of Basic Geometric Properties.
c. Detect and Remove the Non-text region with the help of Stroke Width Length.
d. Combine the final text detected result.
e. Recognize Detected Text Using OCR.

A. Detect the applicant text with the help of MSER:

The MSER is the feature detector used to find the text regions. It works well Good for text because of the steady color and high difference of text leads to stable intensity profiles. To use the detect MSER Feature function to find all text region in the images and plot the result. Notices that there are many text detection algorithms are exist to detect the text from natural images.

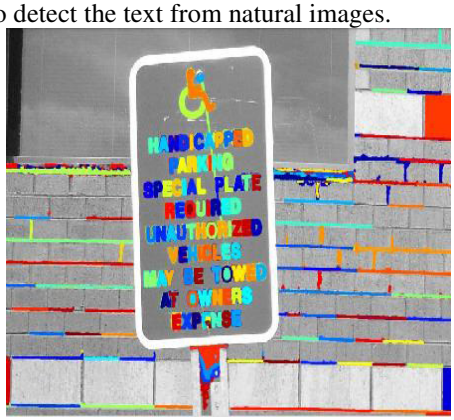


IMAGE 1. Detect the applicant text with the help of MSER.

B. Detect and Remove the Non-text region with the help of Basic Geometric Properties:

Although the MSER algorithm are detect the text from the images and they are also detect the many other stable region of the images that does not contain the text. You can also use the rule-based approach to remove non-text regions. For example, geometric properties of text can be used to filter out non-text regions using simple thresholds. Alternatively, you can use a machine learning approach to train a text vs. non-text classifier. Typically, a combination of the two approaches produces better results [4]. This example uses a simple rule-based approach to filter non-text regions based on geometric properties.

Some of the important properties are:

- ✓ Aspect ratio of the images.
- ✓ Eccentricity of the text region.
- ✓ Euler number
- ✓ Extent
- ✓ Solidity of the images.

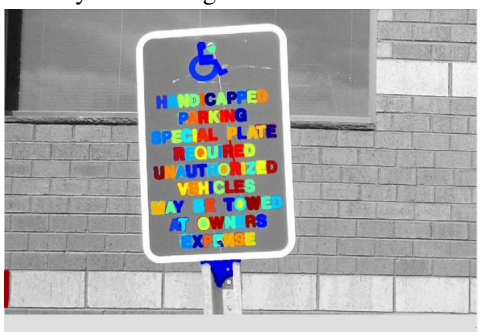


IMAGE 2. After detecting the Non-text region with the help of Basic Geometric Properties.

C. Remove Non-text Regions Based On Stroke Width Length:

Another common metric used to differentiate between the text and non-text is stroke width length. The stroke width is the measure of the width of the curves and lines that make up a character. 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.

Image 3. After Detecting the Non-text region with the help of Stroke Width Length.

D. Combine the final text detected result:

At this part, all the detection results are collected 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, recognize the string like 'THANKS' vs. the set of individual characters {'N','S','K','T','A','H'}, where the meaning of the word is lost without correct ordering.

Image 4. Combine the final text detected result

E. Recognize Detected Text Using OCR:

The final stage contains the detected result, use the OCR function to recognize or identify the text within the bounding box. Note that without finding the text region in the images the output of the OCR is considered as noisier.

Image 5. Recognize Detected Text Using OCR

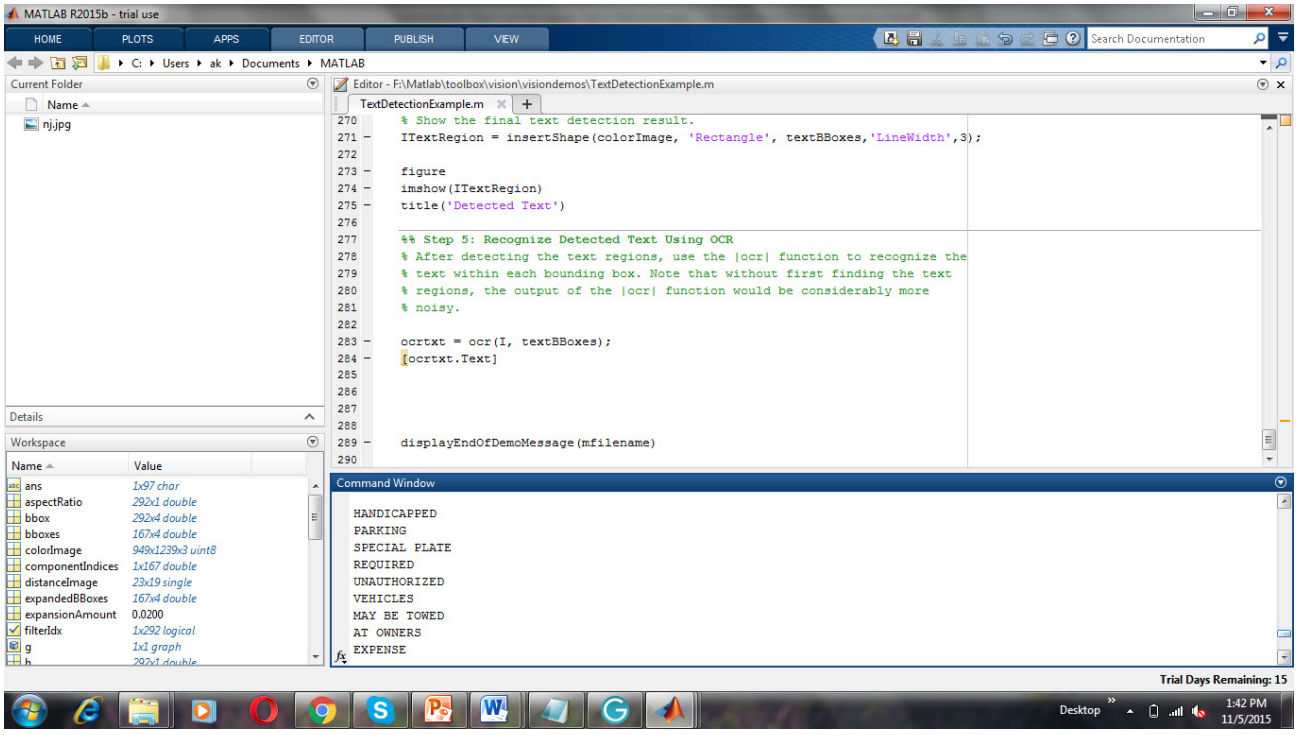


Image 6. Final Result Detected Text on Matlab

IV. IMPLEMENTATION & RESULTS ANALYSIS

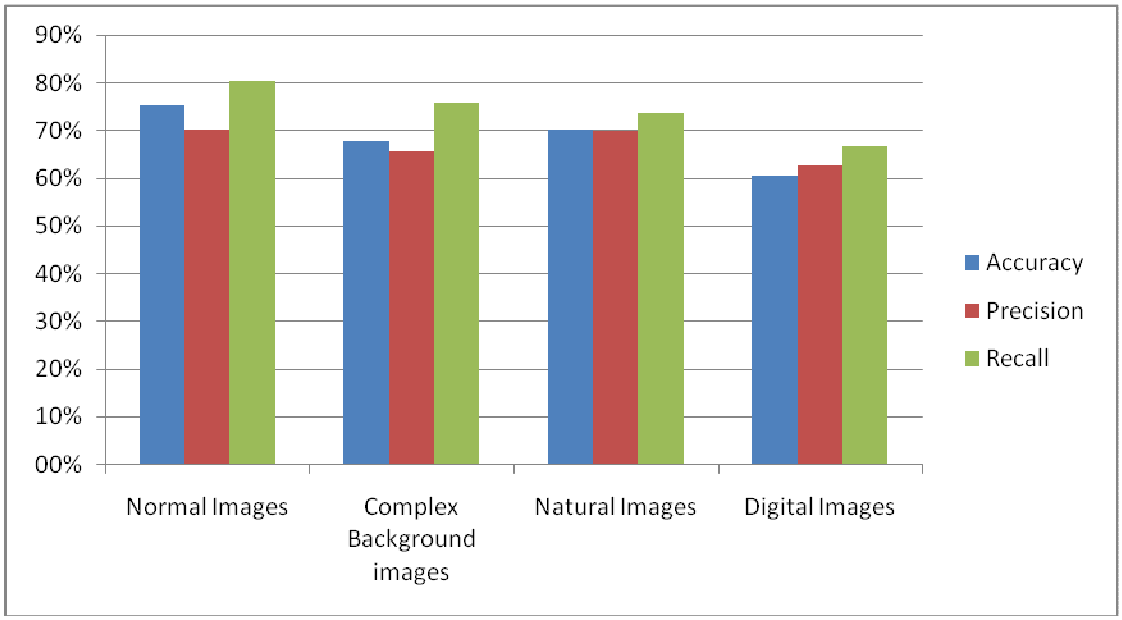
The experiment and analysis is done on system with Sony Vaio, with Intel core i5 1.80 GHZ, 4GB RAM, 64-bit OS. The results are compared with existing methods for text extraction such as edge based detection algorithm, Connected component based algorithm and Wolf- Jolion algorithms. In the proposed algorithm we have modified the connected component based method for parallel implementation. So by using the technique of removing non-text regions with Stroke width variation, we have enhanced the text extraction results. Also by parallel implementation of this modified algorithm, we have increased the speed of operation.

Algorithm	Processing Time (less than10 at time)
Edge Based Text Detection	120 Seconds.
Wolf- Jolion	90 Sec
Connected component Algorithm	60 Sec
Proposed (Serial Execution)	45 Sec
Proposed (Parallel Execution)	20—25 Sec (approx)

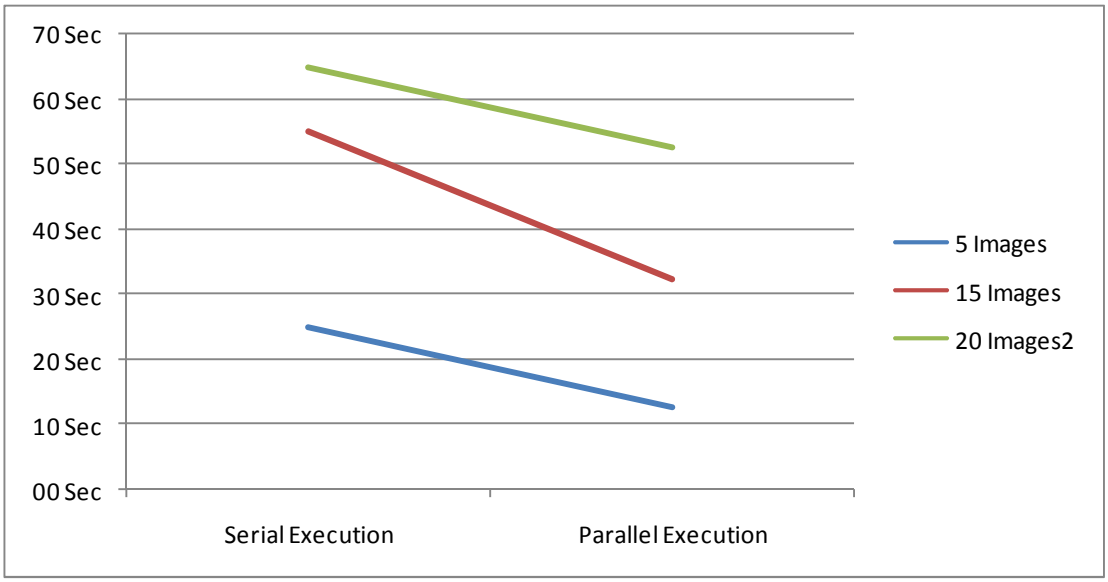
TABLE 1. COMPARISON OF EXECUTION TIME OF ALGORITHM

Table 1. Shows comparison of some existing algorithms with proposed algorithm. The time require for the proposed algorithm and existing methods [1], [22] - [19]. As proposed method is of parallel processing the time requirement was drastically decreased.

The below Graph.1 shows the results of the proposed algorithm for various kinds of images such as Normal Images (i.e images with plane backgrounds), images with complex backgrounds, natural images and digital images (i.e images that are taken from digital camera). In digital images, the distance with which the image is captured is also considered. The result shows that the accuracy of text extraction form normal images is very high as compared to others.



Graph 1. Showing Accuracy of Text Extraction of various images



Graph 2.Showing Processing Time of Serial VS Parallel Execution (in Sec) for 5, 15, 20 images at a time

The above Graph.2 shows the execution time of Serial VS Parallel Implementation. We have shown the execution time of extracting text from 5, 15, 20 images at a time. The result shows that the time is decreased for parallel execution than that of serial execution.

I. CONCLUSION

In the paper the proposed parallel algorithm speeds up the processing as each processor continues execution of the independent modules of code in parallel. The proposed parallel algorithm reduces the time required to extract text from images as compared to the serial text processing. Due to parallel processing of text extraction no other processor is kept idle. Therefore, we get maximum computation power and throughput. Form this we have concluded that there are lots of techniques are available to extract text from images, but using connected component based method we can extract text easily and with greater accuracy. By parallel implementation we have increased the performance of the algorithm. The results show that there is enhancement in the utilization of resources and memory. The speed of entire system is also increased.

Here we can also conclude that we can improve any system by parallel implementation of its operations. So we get higher results with great accuracy. So by implementation of this project we can say that our system can be useful for person with blindness. It helps then to understand the text form images which they could not see through their eyes. This project can also be used in enhanced systems which are making various devices for helping blind people and visually impaired persons. So by detecting text we can convert that text into speech and thus by listening this audio blind or visually impaired person understand the sentiment of the image. So thus this helps in their day to day life and makes their life simpler.

II. FUTURE WORK

In the future, the research is carried out by adding the new algorithm to extract text form very complex and damaged images and videos. The system will be added in the future to convert the text into speech so blind or visually impaired persons can directly understand the text. So with integration of this algorithm with indoor and outdoor text detection we are building the software system for blind persons.

ACKNOWLEDGMENT

The author of this paper would like to thankful for the contribution of VIT University to provide this wonderful opportunity and good facilities to carry out this work.

REFERENCES

[1] Qixiang Ye, Qingming Huang, Wen Gao and Debin Zhao, Fast and Robust text detection in images and video frames, Image and Vision Computing 23, 2005.

[2] AmerDawoud and Mohamed Kame1. : ”Binarization of Document Images Using Image Dependent Model.”

[3] A. Mosleh, N. Bouguila, and A. B. Hamza, “An Automatic Inpainting Scheme for Video Text Detection and Removal,” in IEEE Transactions on Image processing, vol. 22, no. 11, pp.4460–4472, Nov. 2013.

[4] T. Pratheeba, Dr. V. Kavitha, S. Raja Rajeswari, “Morphology Based Text Detection and Extraction from Complex Video Scene,” in Internationl Journal of Engineering and Technology, Vol. 2(3), pp. 200- 206, 2010.

[5] W. Kim and C. Kim, "A new approach for overlay text detection and extraction from complex video scene," IEEE Trans. Image Process., vol. 18, no. 2, pp. 401 –411, feb. 2009.
[6] Yen-Lin Chen, "Automatic Text Extraction, Removal and Inpaiting of Complex Document Images," in International Journal of Innovative Computing, Information and Control, Vol. 8, No.1(A), pp. 303- 327, January 2012.
[7] M. R. Lyu, J. Song, and M. Cai, "A comprehensive method for multilingual video text detection, localization, and extraction," IEEE Trans. Circuits Syst. Video Technol., vol. 15, no. 2, pp. 243 – 255, feb. 2005.
[8] J. Malobabic, N. O"Connor, N. Murphy, and S. Marlow, "Automatic detection and extraction of artificial text in video," in WIAMIS 2004 - 5th International Workshop on Image Analysis for Multimedia Interactive Services, April 2004.
[9] Umesh Kumar Singh, Shivlal Mewada, Lokesh Laddhani and Kamal Bunkar, "An Overview & Study of Security Issues in Mobile Ado Networks", International Journal of Computer Science and Information Security (IJCSIS) USA, Vol-9, No.4, pp (106-111), April 2011
[10] C. Wolf and J.M. Jolion, "Extraction and Recognition of Artificial Text in Multimedia Documents", Technical Report RVF-RR-2002.01, Available: http://rvf.insa-lyon.fr/~wolf/papers/tr-rfv-2002, February 2002.
[11] Mohammad Khodadadi and AlirezaBehrad, "Text Localization, Extraction and Inpainting in color images", ICEE2012, Vol.12, IEEE, May 2012, pp. 1035-1040
[12] Sneha Sharma.:"Extraction of Text Regions in Natural Images."
[13] Pal, U. and Chaudhuri, B.B. 2004. Indian script character recognition: a survey. Pattern Recognition 37, 1887 – 1899.

Authors Profile

Niketan Janakraj Jivane received the Bachelor of Engineering degree in computer science and engineering in 2014 from Smt. Kashibai navale college of Engineering, Vadgaon (BK), Pune, Pune University, Maharashtra, India. He is currently working towards the M.Tech degree at School of Computer Science and Engineering, Vellore Institute of Technology, Vellore, Tamil Nadu, India. His research interest includes Image processing, Machine Learning, Data Mining, Computer Networks.

Supriya Janakraj Jivane currently pursuing the Bachelor of Engineering degree in computer science and engineering from M.B.E. Society's College of Engineering, Ambajogai, Maharashtra, India. Her research interest includes Image processing, Cloud Computing, Computer Networks.

Rajkumar Soundrapandiyan is currently working as an Assistant professor (Senior) in Department of Computer Science and Engineering, VIT University, Vellore, India. He received his B.E degree in Computer Science and Engineering from Anna University, Chennai, India in 2008, M.E degree in Computer Science and Engineering from Anna University, Chennai, India in 2010. He is pursuing his PhD at VIT University, Vellore, India. He has submitted his Ph.D. Thesis on August 2016. His research interest includes Digital image processing, Computer Vision, Visual Perception, Object detection, Medical image processing and Infrared image procesing. He has published several refereed research papers in various reputed international conferences and journals. He is a reviewer for many reputed journals like Computer Biology and Medicine, Journal of Medical Imaging and Health Informatics etc. He is a life member of CSI.

Dr. Marimuthu Karuppiah is currently working as an Associate Professor in the Department of Computer Science and Engineering (SCOPE) in VIT University, Vellore. Prior to joining VIT University, he held academic position of Senior lecturer in Department of Computer Science and Engineering, Syed Ammal Engineering College, Ramanathapuram, India from March 2005 to November 2008. He received his Ph.D. degree in Computer Science and Engineering from VIT University, Vellore, on August 2015 [July 2009 - August 2015]. He completed ME in Computer Science and Engineering from Anna University, Chennai, in the year 2005. He also received BE degree in Computer Science and Engineering from the Madurai Kamaraj University, Madurai, India, in 2003. His current research interests includes cryptography, wireless sensor network security, mobile ad hoc network security, vehicular ad hoc network security, remote user authentication using smart cards. He has authored over 20 papers in reputed international and conferences. He is a life member of Cryptology Research Society of India (CRSI) and Computer Society of India (CSI).

