© 2016 IJSRSET | Volume 2 | Issue 2 | Print ISSN : 2395-1990 | Online ISSN : 2394-4099
Themed Section: Engineering and Technology

Automatically Identifying the Text Name Detection with Speech Output for Blind Persons

D. Ramya, K. Sumathi

ECE Department, SNS College of Technology, Coimabtore, Tamil Nadu, India

ABSTRACT

The camera-based assistive text reading framework to share the information for blind persons to read text name and product packaging from hand-held objects in their day-to-day life is proposed. The work consists of three stages. First is image capturing –by using a mini camera, the text which the user want to read will get captured as an image and have to transfer to the image processing platform. Second is text recognition –by single text recognition algorithm, the text will get filtered from the image in the screen. Third is speech output - the filtered text will be shared into system to get an audio speech output. The entire process is done with the help of MATLAB software.

Keywords: Image capturing, Text, Region detection, Character recognition, Speech output, Mat lab software.

I. INTRODUCTION

Image processing is processing of images using mathematical operations in any form of signal processing for which the input is an image, such as a photograph or video frame. The output of image processing may be either an image or a set of characteristics or the parameters which is related to the image. Most image-processing techniques which involve treating the image as a two-dimensional signal and applying the standard signal-processing techniques to the input.

Image processing is usually refers to digital image processing, optical and analog image processing also are possible.The acquisition of images is referred to as imaging. In computer graphics, images are manually made from physical models of objects, surrounding and lighting, instead of being acquired from natural scenes, and also in most animated movies. Computer vision, on the other hand is often considered a high-level image processing. (e.g., videos or 3D full-body magnetic resonance scans). The millions of visually impaired people in worldwide are still blind. Even in developed countries like united states the people lack the ability to see. Recently, in computer vision and portable computers are helped to to develop the vision technology with the help of optical character recognition (OCR). Reading is one of the basic necessity today, everything around us are in the form of receipts, statements, product packages, menus etc. It Contain printed text and video magnifiers and screen readers will help the blind users and with lower vision help to facilitate text label reading. This can help the blind people identify the hand-held objects much as product packages and objects printed with text label. Processing these devices which are more portable and sophisticated can promote independent living and foster economic and social self-dependency.

This system combined optical character recognition for scanning the text in image and give as voice audio output. The novel text localization algorithm that combines rule based analysis and learning based text are used to read the text in the image. The proposed algorithm are efficient and it handle all the drawbacks of existing models and it extract the text from hand-held product and it captured through the camera or mobile phone. The most important part in reading system for blind person is positioning of object of interest within camera view. The text from the background is also included. To extract the text name from the hand-held

object from the image the motion based method algorithm is used to isolate the region of interest and also text recognition is done only the area of interest.

II. METHODS AND MATERIAL

In proposed work, a prototype system is used to read printed text on hand-held objects for blind persons. In order to solve this problem for blind users, a motion-based method to detect the object of interest. This method can effectively distinguish the object of interest from background or other objects. To extract he text regions from complex backgrounds, a novel text localization algorithm based on stroke orientation and edge distributions in the process. The three processes which include is screen capturing, data processing and audio output.

The screen capture components which capture the scene by web camera in real time with motion based and it will calculate the foreground object at the each frame in the image. The object of interest is localized by the means of foreground. The proposed algorithms that include object of interest and text localization the image region containing text and that is converted as text into readable codes.

Detecting the MSER region and matching is used for the proposed system, with identifying text and non-text region. Next removing the non-text region from the text image, by using SURF (Speeded-Up Robust Features) method image is matching at the end of execution process. Off-the-shelf OCR is used to perform word recognition on the localized text regions and transform into audio output for blind users.

Binary thinning operation is executed to give the exact shape of the each character from text to identify the text perfectly in processing time.

Here laptops are used as the processing device in the current prototype. The audio output components read the correct text code. The Bluetooth ear piece or any mini speaker is used for speech output. The main drawback in this system is complex and heavy device and the blind people found it's difficult to carry it along with them. All computed using matlab software.

III. RESULTS AND DISCUSSION

A. **Input Image**

Figure 1. Input Image

The input image is captured in real time by using web camera is captured as image or an video sequence ,it take few seconds to capture the image in real time process.

B. **Red Channel Image**

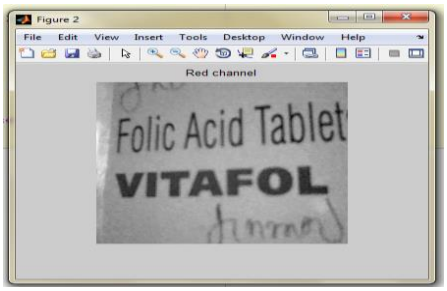

Figure 2. Red Channel Image

The captured input image is converted as colour into red channel image. Background in an image is removed to reduce the noise when implementation process takes place. It provides fast recovery.

C. **Green Channel Image**

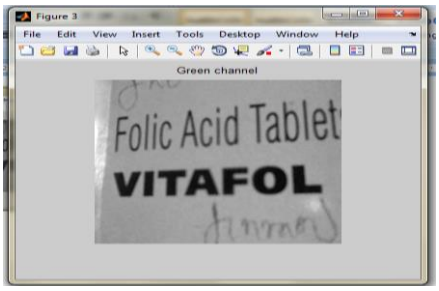

Figure 3. Green Channel Image

The captured input image is converted as colour into green channel image. Background in an image is

removed to reduce the noise when implementation process takes place. It provides fast recovery.

D. **Blue Channel Image**

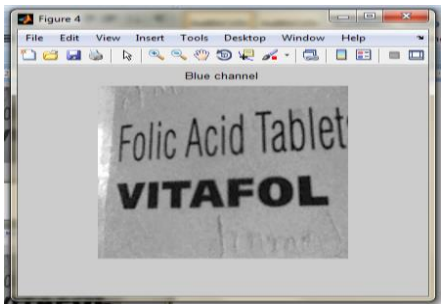

Figure 4. Blue Channel Image

The captured input image is converted as colour into blue channel image. Background in an image is removed to reduce the noise when implementation process takes place. It provides fast recovery.

E. **Detection of MSER Region in Image**

Figure 5. MSER Region image

The MSER feature detector works well for finding text regions. It works well for text because the consistent colour and high contrast of text leads to stable intensity profiles. Using this function all the regions within the image is identified and there are many non-text regions detected alongside the text.

F. After Removing Non –Text Region Based on Geometric Properties

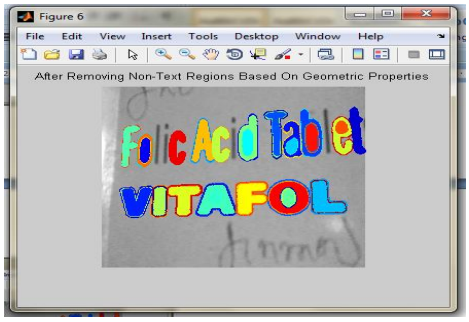

Figure 6. After Removing Non –Text Region

Removed non–text region from an input image is shown in above figure. Removing the non-text region which is presented in an image by using basic geometric properties it identifies the text clearly in an image without any confusion.

G. **Region Image and Stroke Width Image**

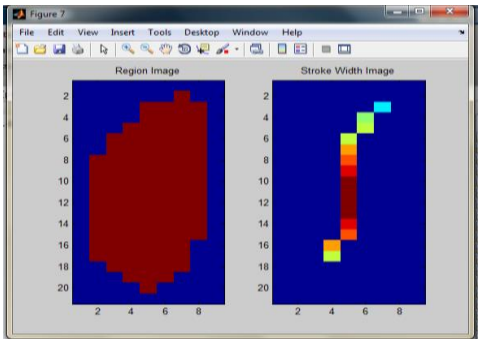

Figure 7. Region Image and Stroke Width Image

The region image and stroke width of an image is shown in the above figure. Feature extraction is done by using stroke width to identify the curves and lines of the text in the images. After feature extraction, the text input data is saved to for further process.

H. Thinning Image

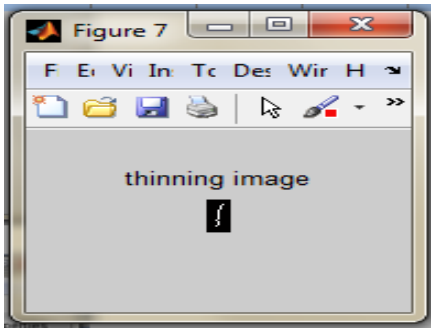

Figure 8. Thinning Image

Thinning is a morphological operation that is used to remove selected foreground pixels from binary images. It preserves the topology (extent and connectivity) of the original region while throwing away most of the original foreground pixels. When matching process takes, the thinning image will be executed to identify the correct shape of the text from an image.

I. **After Removing Non Text Regions based on Stroke width variation**

Figure 9. After Removing Non –Text Regions based on Stroke.

The output of the stroke from an image is detected and is shown in above image. Stroke detection is used to identify the corner or curves or boundaries or lines of the given text in the input image. It can be easily identify the text in image.

J. Expanding The Bounding Boxes Text

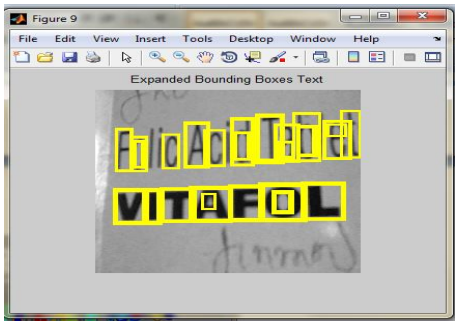

Figure 10. Expanding the bounding boxes text

The bounding box output image of the text is shown in the above image. The smallest bounding is used to measure the area, volume, or hyper volume in higher dimensions within which all the points lie. It is used to speed up computation.

K. **Final Detection of each Character in Image**

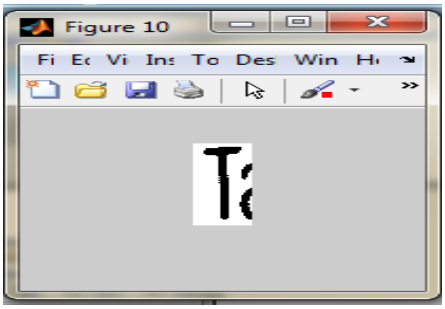

Figure 11. Final Detection of each Character in Image.

The detected each character in a text from an image is shown in above figure. Each character in text is identifies by using optical character recognition method. The character from each text is easily identified.

L .Identified Text Image

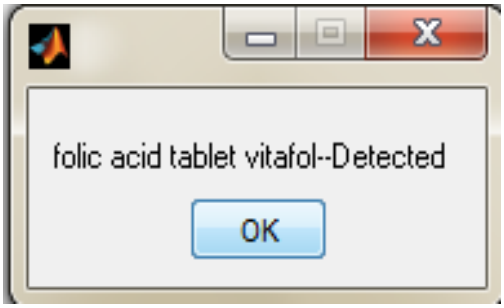

Figure 12. Identified text is detected.

Finally the output of the text is shown in the above image. The image which is selected from that the particular text in real time will be given as speech output to the users. The output will be given as soon as possible to the blind users.

IV. CONCLUSION

The text to speech synthesis framework for the recognition of natural scene text is described. A scheme for character detection and segmentation from real time images is identified. The detection stage, stroke edge is employed to detect possible text regions, and some geometrical features are used to filter out obvious non-text regions. By using K-Means cluster algorithm the image is segmented. The features are obtained from MSER (Maximally Stable Extremal Region) identifying the text region from an image, stroke distribution, and stroke width. The proposed framework of text localization is evaluated on scene images and images of handheld objects captured by blind persons are successfully completed. The proposed algorithm achieves a better result when compared to existing algorithm. By using MSER and matching method the text can be detected easily. It takes faster to implement the text from the input image. The performance is clearly observed in text to speech analysis process.

V. REFERENCES

[1] X.chen and A.yuille." Detecting and reading text in natural scenes."In proceedings of IEEE Conferenceon (CVPR) volume 2, pages II-366. IEEE, 2004

[2] L. Neumann and J. Matas. "Real-time scene text localization and recognition. " In proceedings, Eighth International conference on Document Analyisis and Recognition (CVPR), 2012.

[3] T. de campos, B. Babu and M. varma "Character recognition in natural images. " In VISAP, 2009.

[4] K. Wang, B.Babenko and S.Belongie. "End-to-end scene text recognition." In international conference on computer vision(ICVV), 2011.

[5] Per-Erik Forssen and David G.Lowe "Shape Descriptors For Maximally Stable Extermal Regions."

[6] P. E Forssen." Maximally Stable colour regions for recognition and matching". In IEEE coference on computer vision and pattren Recognition,Minneapolis, USA, June 2007. IEEE Computer Society.

[7] J. Matas, O. Chum, M.Urban and T. Pajdla."Robust wide baseline stero from maximally stable extremal regions." In 13th BMVC, pages 384-393 September 2002.

[8] M.Yokobayashi and T.wakahara. "Segmentation and recognition of characters in scene images using selective binarization in color space and get correlation." In proceedings, Eighth (IDAR), pages 167-171, IEEE, 2005

[9] Saleh Ali K. Al-Omari, Putra sumari, Sadik ". Digital Recognition using Neural Networks." Journal of computer sciences,2009.

[10] Chucai Yi and Aries Arditi "Portable Camera-Based Assistive Text and Product Label Reading From Hand-Held Objects for Blind Persons", vol. 19, no. 3, june 2014.

[11] Amarnath singh, S.Ram kishore, G. swathi, "Portable read pad for object identification for blind through mat lab and android", January 2015.

[12] Prof.R.R.Bhambare, Akshay Koul, Siddique Mohd Bilal, Siddharth Pandey, "Smart vision system for blind", Vol 3, May 2014.

[13] Dimitrios Dakopoulos and Nikolaos G.Bourbakis"WearableObstacle Avoidance Electronic Travel Aids for Blind " vol. 40, no. 1, January 2010.

[14] Boris Epshtein, Eyal Ofek and Yonatan Wexler," Detecting Text in Natural Scenes with Stroke Width transform". in Proc.comput. vision pattern recognit.2010.

[15] S. pushpalata, R. infantabinaya, E. esakkiammal, "Handheld Object Identification and Text reading for Visually challenged" International Conference on Science, Technology, Engineering & ManagementICON-STEM'15 Journal of Chemical and Pharmaceutical Sciences.Issue 10: July 2015 .