Integrated Assistance System for Blind

Boddula Balaji, Pallikonda Sarah Suhasini

Abstract: ***This paper proposes an integrated assistance system for blind and visually impaired people. The system incorporates three basic features viz, color detection, text detection and location finding. The process of color detection includes capturing the objects through a webcam and applying the Color Image Processing techniques. To assist the person in reading, text to speech conversion method is used. Initially, the printed document is captured through webcam, then adaptive Gaussian thresholding, dilation and erosion techniques are applied to detect the text. Finally the detected text is mapped to the corresponding predefined words, which are then converted to speech. For extracting text Tesseract OCR Engine is used. In addition a GPS module is added for tracking the location of the blind person and to send the location details in the form of an SMS to the registered phone numbers. The entire system is implemented using Raspberry Pi.***

Index Terms: Color Detection, Text to Speech Conversion, Tesseract OCR Engine, Location finding, Adaptive Gaussian thresholding, Dilation, Erosion.

I. INTRODUCTION

According to the 2017 analysis of world population, there are about 2% blind people and 5% visually impaired worldwide. Although there have been several devices designed for helping visually disabled to see objects using an alternating sense such as touch, the development of text reading devices is still at an early stage. Early techniques for text detection include Braille Script [5], Nipkow disk [6], Optophone [7], One-Eyed Machine Stenographer, Reading machine which is a first Optical Character Recognition (OCR) machine, Optacon [8], Tesseract [9], etc. S. Impedovo, et al. in their survey paper on OCR explained and compared pre-processing, feature extraction and post-processing methods for commercial reading machines [1]. Text segmentation is done by learning best binarization values for OCR systems. To differentiate object from background using gradient-analysis of image histogram, fast computation of threshold value is proposed by A.H. Ismail and M.H. Marhaban [2]. For Filtering Texture and Flat areas of Image, Jain A, Gupta R [3] proposed Threshold Modulation of Gaussian Filter. Al-Doweesh S.A, et al. [4] presented an iphone application for color detection which is helpful for blind and visually impaired people. In their application, HSL color space is used to produce color for the corresponding input image pixel values. In this work, three features are incorporated viz. i) recognizing the text present in the printed document and converting it into the speech, ii) recognizing the color of the objects and iii) tracking the location of the person and sending his details to the registered numbers in the form of a SMS. Thus this system would be helpful to the visually impaired for reading, color identification and location finding.

Revised Manuscript Received on June 14, 2019.

Boddula Balaji, Student, Department of ECE, Velagapudi Ramakrishna Siddhartha Engineering College, Vijayawada, India.

Pallikonda Sarah Suhasini, Associate Professor, Department of ECE, Velagapudi Ramakrishna Siddhartha Engineering College, Vijayawada, India.

II. METHODOLOGY

The integrated system designed has three subsystems viz. Color Identification, Text Reading and location finding.

A. Color Detection

As shown in Fig.1, in this subsystem color of the object is identified using RGB color space. Through the webcam the image is captured, the pixel value of the corresponding spot selected for identifying the color of the object, is converted to RGB. The resultant color is displayed in text and audio form.

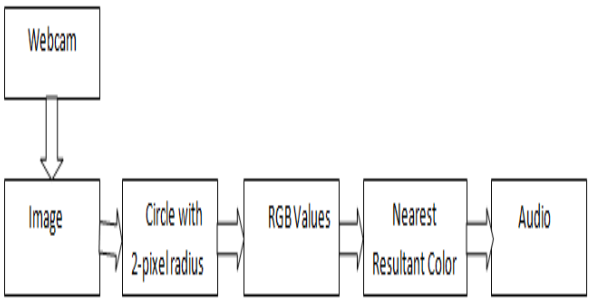

Fig. 1. Block Diagram of Color Detection

Processing Steps

Step 1: The object whose color is to be detected is captured with the help of web cam.
Step 2: With the help of the width and height co-ordinates a point on the image is selected.
Step 3: At the selected point a circle of radius 2 pixels is drawn.
Step 4: The BGR color triplet of the circle is extracted instead of RGB since the Open-CV software which is used to extract the color uses the BGR conversion technique.
Step 5: The detected BGR Colour triplet is converted into the RGB format.
Step 6: The nearest color to the extracted RGB triplet is calculated and resultant color is displayed beside the circle.
Step 7: Finally the recognized color is converted into the speech using the speak library of python.

B. Text to Speech Conversion

The process of text to speech conversion includes capturing the printed document through webcam and applying different Image Processing techniques like adaptive Gaussian thresholding, dilation, erosion present and text is extracted from the processed document using the pytesseract library. Finally, using the detected text is converted into speech.

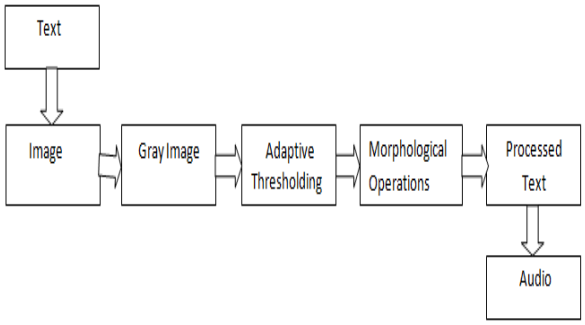

Fig.2 Block Diagram of Text to Speech Conversion

Processing steps:

Step1: Capture the image consisting of required text using the web cam.

Step 2: The color image is converted into the gray scale.

Step 3: Adaptive thresholding is applied on the gray scale image and this technique in the simplest implementation, outputs a binary image representing the segmentation. For each pixel in the image, a threshold is calculated. If the pixel value is below the threshold it is set to the background value, otherwise it assumes the foreground value.

Step 4: Eroding is applied on the gray scale image which will further help in decreasing the noise present in the image and to detach the two connected characters.

Step 5: Erosion avoids the white noise in the object but it shrinks the characters. So as to avoid shrinking, dilation is applied which increases the characters size.

Step 6: The pytesseract module will extract the text from the processed image and it sends the text to the SAPI library which will produce the speech output.

C. Location Finding

The coordinates which are extracted using the GPS antenna are converted into the readable location names using reverse geo decoding and a text SMS is sent to the registered names along with audio to guide the person, as shown in Fig.3.

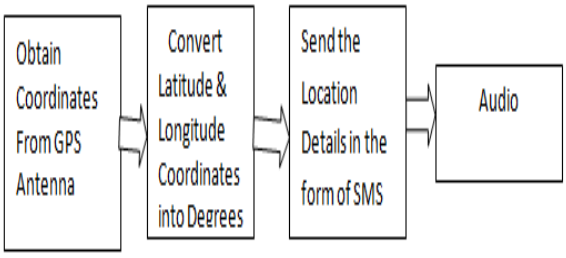

Fig.3 Block Diagram for location finding

III. RESULTS AND DISCUSSION

In this system the color detection, text to speech conversion and the location detection modules are integrated using a Raspberry Pi. Initially the user has to select one of the options either text to speech conversion, color detection or location detection using the three buttons that are interfaced to the Raspberry Pi GPIO pins. The entire process is as shown in Fig.4.

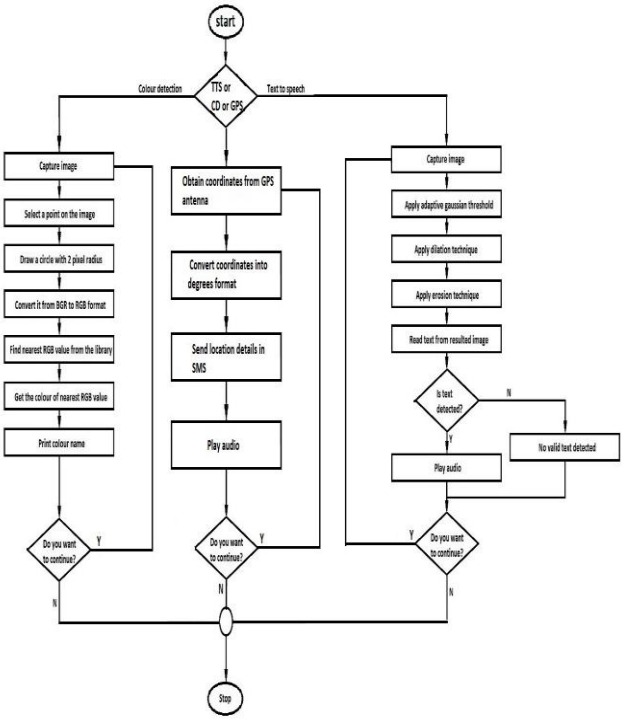

Fig.4 Flow Chart of Assistance System

i) For the mode of operation - OCR

The text to speech conversion is verified for different font sizes and different distances.

Test Case 1: Placed the Document with font size 14 at a distance of 16 cm from the webcam and the outputs are observed.

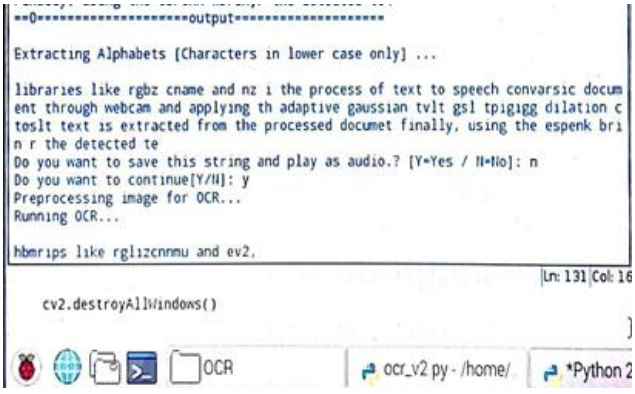

Fig.5 Text recognition o/p for font size 14 and distance 16 cm.

Test Case 2: Placed the Document with font size 16 at a distance of 25 cm from the webcam and the outputs are observed.

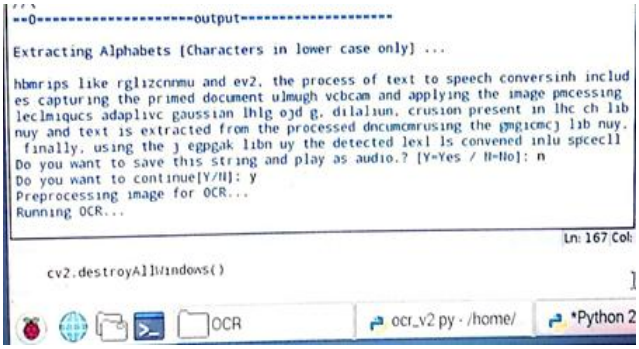

Fig.6 Text recognition o/p for font size 16 and distance 25 cm.

The results of the 4 test cases are shown in Table 1.

Table 1 Accuracy of Text Detection

S. No	Distance in cm	Font size	Total words	Detected words	Accuracy
1	16	14	52	34	65.38
2	25	16	54	32	59.25
3	31	18	54	35	64.38
4	20	20	44	40	90.90

ii) For the mode of operation – Color Detection

Input given is a blue colored paper and the output is observed is the same.

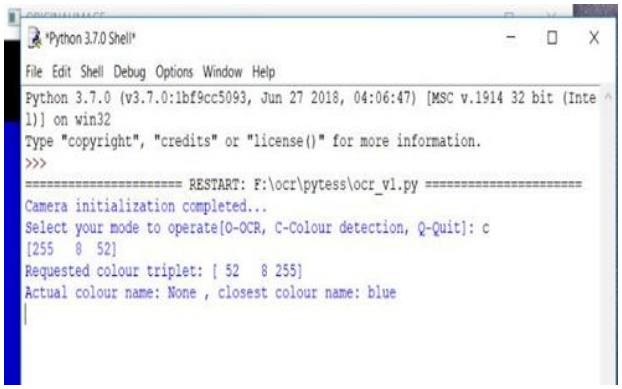

Fig 7 Output for color detection

iii) For the mode of operation – Location Finding

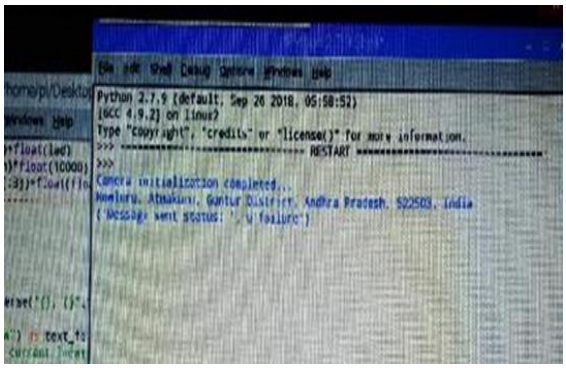

Fig 8 Output for location detection

IV. CONCLUSIONS

With the proposed assistance system for single colored object, the recognition of color is found to be accurate; however for multi colored objects all the colors are not detected. This may be improved by using cameras of good resolution. The system accuracy for the text to speech conversion is found to be high when the font size is large and the distance between the text document and the aperture of the webcam is small. If the person is in any panic situation his location details can be sent to the friend or a family member in the form of a text SMS.

The proposed algorithm for the text detection is limited to the printed documents. This can be further extended to detect the hand written text and multi colored objects with the help of machine learning techniques.

REFERENCES

1. S. Impedovo, L. Ottaviano, S. Occhinegro, "Optical Character Recognition – A Survey", International Journal of Pattern Recognition and Artificial Intelligence, vol. 5, pp. 1-24, 1991
2. A.H. Ismail, M.H. Marhaban, "A Simple Approach To Determine the best threshold value for automatic image thresholding", IEEE International Conference on signal and image processing, 2009.
3. Jain A, Gupta R, "Gaussian filter threshold modulation for filtering flat and texture area of an image", International Conference on Advances in Computer Engineering and Applications, ICACEA, 22,760-763, 2015.
4. Al-Doweesh S.A., Al-Hamed F.A., Al-Khalifa H.S., "What Color? A Real-time Color Identification Mobile Application for Visually Impaired People" Stephanidis C. (eds) HCI International, 2014.
5. Jiménez J, Olea J, Torres J, Alonso I, Harder D, Fischer K., "Biography of louis braille and invention of the braille alphabet", Surv Ophthalmol, 2009,54(1):142-149
6. Timothy R. Corle, Gordon S. Kino, "Confocal Scanning Optical Microscopy and Related Imaging Systems", 1996
7. E. E. Fournier D'albe, "The Optophone: An Instrument for Reading by Ear", Nature, 105, 295–296 ,1920
8. H. Goldish, Louis & E. Taylor, Harry, " The Optacon: A Valuable Device for Blind Persons", New Outlook for the Blind. 68, 1974
9. R. Smith, "An Overview of the Tesseract OCR Engine", Proc. International Conference on Document Analysis and Recognition, 2007