ISSN 2348-1196 (print)
International Journal of Computer Science and Information Technology Research ISSN 2348-120X (online)
Vol. 3, Issue 1, pp: (87-92), Month: January - March 2015, Available at: www.researchpublish.com

# Compact Camera Based Assistive Text Product Label Reading and Image Identification for Hand-Held Objects for Visually Challenged People

R. Infant Abinaya[1], E.Esakkiammal[2], Mrs.Pushpalatha[3]

Department of Information Technology, Jeppiaar Engineering College, Chennai, India

***Abstract:*** **Assistive technologies are being developed for visually impaired people in order to live confidently. This project work proposes a camera-based assistive text reading framework and to process the captured object and obtain its details to help blind persons read text labels and product packaging from hand-held objects in their daily lives. The project work is framed into three stages. First, Image capturing – Using a mini camera ,the text which the user need to read get captured as an image and have to send to the image processing Platform. Secondly, Text recognition – Using text recognition algorithm, the text will get filtered from the image. Finally, Speech output - A filtered text will be passed into this system to get an audio output. This project work can be able to insist the blind people in their daily life. The entire application is based on Raspberry Pi.**

## 1. INTRODUCTION

There are over 314 million visually challenged people worldwide , a major part of this population are still blind even in developed countries like united states, the national health interview survey conducted in 2008 reported that over 85 of the adult Americans lack the ability to see. In recent time's development in computer vision, digital cameras and portable computers help to aid these individuals by developing camera based products that integrate computer vision technology with already existing products such as optical character recognition (OCR). Reading is one of the basic necessity today .everything around us are in the form of reports, receipts, bank statements, product packages, restaurant menus etc... Contain printed text on it although optical aids and video magnifiers and screen readers help blind users and those with lower vision help to facilitate text reading. There are few devices that can render better access to common hand-held objects much as product packages and objects printed with text. Formulating devices which are even more portable and sophisticated can promote independent living and foster economic and social self-dependency.

There are already some portable systems in use that cannot handle product abiling for example, bar code readers help indentify various products in the extensive product database to enable blind users to access information about products through Braille and speech . But there was difficulty in finding him possession of the barcode so pen scanners might be employed in the cases. These system combine OCR software for the propose of scanning and text recognition and integrated voice output But most OCR software proved insufficient to handle images with complicated background.

Reader mobile runs on cell phone and reads document which are really flat with the dark surface and mostly contain text. Accurately read black print on a white background but cannot read text with complex background and text on uneven surface.

There are no existing reading assistance that can read from all kinds of challenging background and patterns .Our proposed algorithm that effectively handle likely all the drawbacks of existing models and extract text from hand-held objects and nearby sources which ever captured through the camera. The most challenging part in assistive reading system for blind people in positioning of object of interest within the camera view. In order to focus the object within the camera

view a camera with whie angle is used as an approximate solution. Often text from the surrounding areas also included. Thus to extract the hand-held objects from the image we proposed motion based method to isolate the region of interest and text recognition is done only on the area of interest. For text orientation the paper includes text strings in screen images keep approximately horizontal alignment. The algorithms developed for localization of text regions in the screen images are divided into two categories: rule based, learning based

Rule Based algorithm include pixel-level image processing to extract text information from predefined text layout such as character size, aspect ratio, edge density ,character stricter , color uniformity of text string etc.. Edge pixel density is analyzed with laplacian operator and maximum gradient indentify text regions and these different maps and perform global binorization to obtain text regions. Stroke with transforms to localize text characters color based text segmentation is performed through a Gaussian mixture model for calculating confidence value for text region. Leering based algorithm model the structure of the text an extract text features for text classifiers.

To solve the task at hand, extract text information from complex background with multiple and variable text pattern a text localization algorithm combine rule based and learning based algorithm which define novel features based on stroke orientation and edge distribution. These, intent, generate representative and discriminative text features which distinguish text character from backgrounds.

## 2. LITERATURE SURVEY

**A System-Prototype Representing 3D Space via Alternative-Sensing for Visually Impaired Navigation:**

This paper proposes a mode of interaction with the surrounding 3-D space to the visually impaired for collision free navigation is a goal of great significance that includes several key challenges [1].

**A System for Identification of Dangerous Spots and POIs for the Blind and Visually Impaired:**

The paper presents the course of work on construction of a prototype system for demarcating dangerous areas and points of interest for blind and visually impaired people in large urbanized areas [2].

**A survey on wearable devices used to assist the visual impaired user navigation in outdoor environments:**

This paper introduces a comprehensive survey of wearable systems designed to assist the visual impaired user's navigation in everyday life outdoor scenarios [3].

## 3. EXISTING MODEL

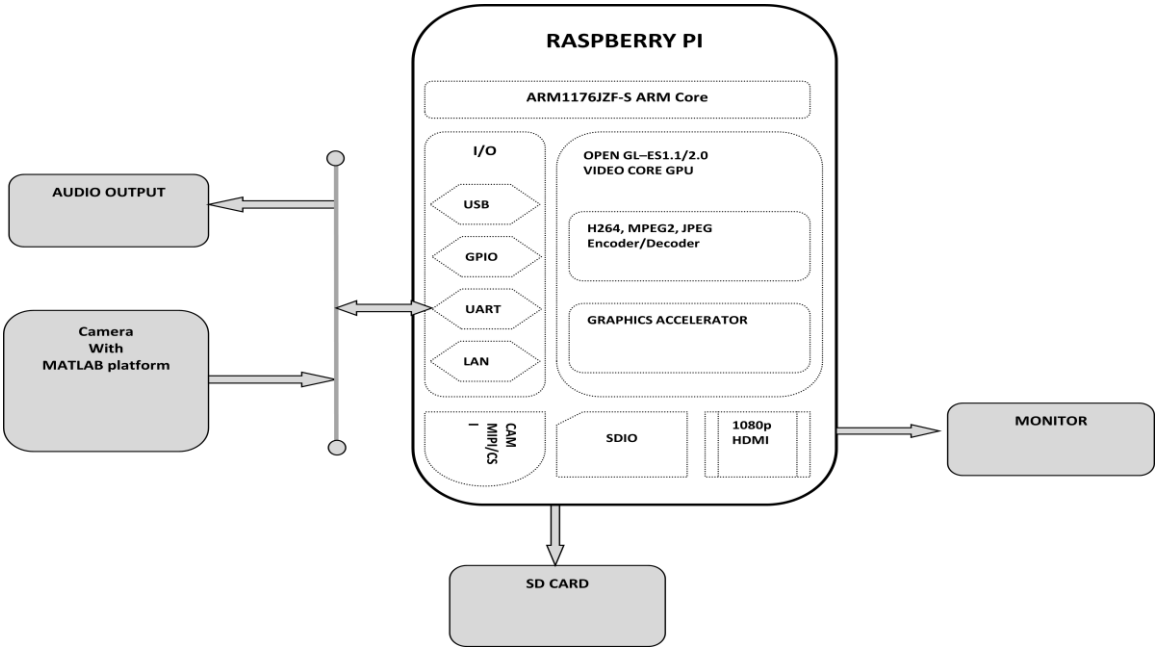
This paper puts forth a prototype system of assistive text reading and image processing .The three fictional components include screen capturing, data processing and audio output. The screen capture components captures the scene by motion-based object-deduction using a camera attached to a pair of sun glasses and 'R' denote the calculated foreground object at the each frame. The object of interest is localized by the mean of foreground .Data processing component deploy our proposed algorithms that include object of interest , text localization to obtain image region containing text and to convert the intensified text into readable codes.Mini laptops are used as the processing device in the current prototype. The audio output components read the recognized text code. The Bluetooth ear piece facilitates speech output.

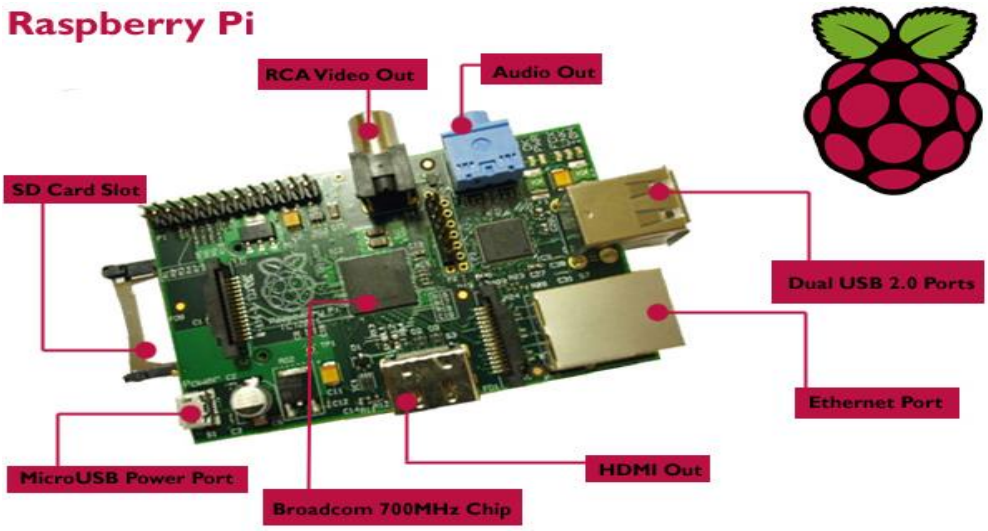
The drawback death with these systems is that it is a complex and heavy device and the blind people found it difficult to carry it along with it. The internal architecture consisted of 8051 microcontroller and separate chips for CPU, GPU, and USB controller, RAM which made this model heavy and complicated. Thus our proposed work deals with designing a device which would be handy and easily portable for visually impaired people to carry it along.

## 4. FUTURE WORK

The Raspberry Pi is a credit card sized single computer or Sock uses ARM1176JZF-S core. SoC, or System on a Chip, is a method of placing all necessary electronics for running a computer on a single chip. Instead of having an individual chip for the CPU, GPU, USB controller, RAM everything is compressed down into one tidy package. Raspberry Pi needs an Operating system to start up. In the aim of cost reduction, the Raspberry Pi omits any on-board non-volatile memory used to store the boot loaders, Linux Kernels and file systems as seen in more traditional embedded systems. Rather, a SD/MMC card slot is provided for this purpose. After boot load, as per the application program Raspberry Pi will get execute.

ISSN 2348-1196 (print)
International Journal of Computer Science and Information Technology Research ISSN 2348-120X (online)
Vol. 3, Issue 1, pp: (87-92), Month: January - March 2015, Available at: www.researchpublish.com




## 5. FUNCTIONS OF INTERNAL COMPONENTS

**SOC:**

Broadcom BCM2835 → High Definition Embedded Multimedia Application Processor.

**Power Supply:**

Micro USB power supply

5v, 700mA=3.5w which is less than a bulb. Standard charging port. Universal charging solution by GSMA.

**Monitor:**

Supports only HDMI.Old analogous TV, modern Digital TV desktop monitor or even your Smartphone as a display.

**Sd Memory Card:**

Min 2GB, expandable up to 32GB. Supports only Linux as Fedora, Archlinux & Desbian.

Desbian is the most recommended one; the reason is it supports python programming language.The Raspberry Pi versions of kernels are as follows: Fedora – Pidora , Desbian – Raspbian.

ISSN 2348-1196 (print)
International Journal of Computer Science and Information Technology Research ISSN 2348-120X (online)
Vol. 3, Issue 1, pp: (87-92), Month: January - March 2015, Available at: www.researchpublish.com

## 6. FUNCTIONAL MODULES

**1. Image Capturing:**

The object from which the text label has to be extracted is captured using a mini-camera which is embedded within our device.

**2. Text Extraction:**

Text from the captured image is extracted using a MATLAB PLATFORM.

2.1: Read Image:

This cell of codes read the image to MATLAB workspace.

2.2: Convert to grayscale image:

This cell of codes converts the RGB to gray.

2.3: Convert to binary image:

This cell of codes converts the gray to binary image.

2.4: Edge detection:

This cell of codes detects the edge of the image.

2.5: Morphology:

This cell of codes performs the image dilation and image filling on the image.

2.6: Blobs analysis:

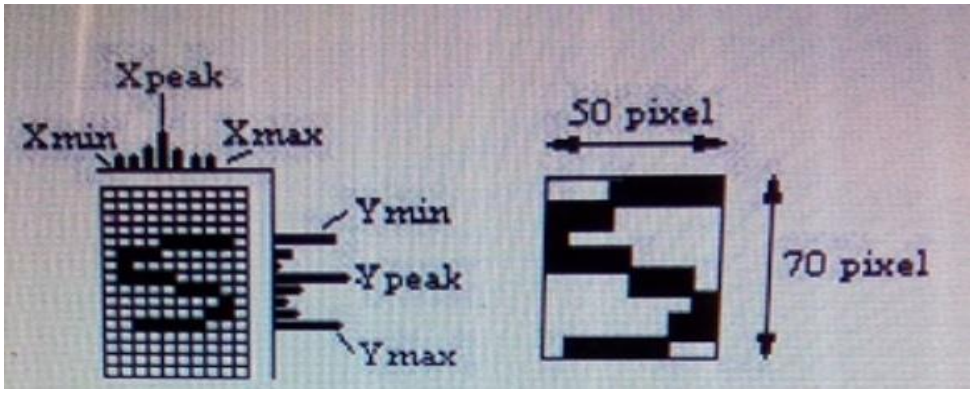
This cell of codes finds all the objects on the image, and finds the properties of each object.

2.7: Plot the Object Location:

This cell of codes plots the object locations.

***Feature Extraction:***

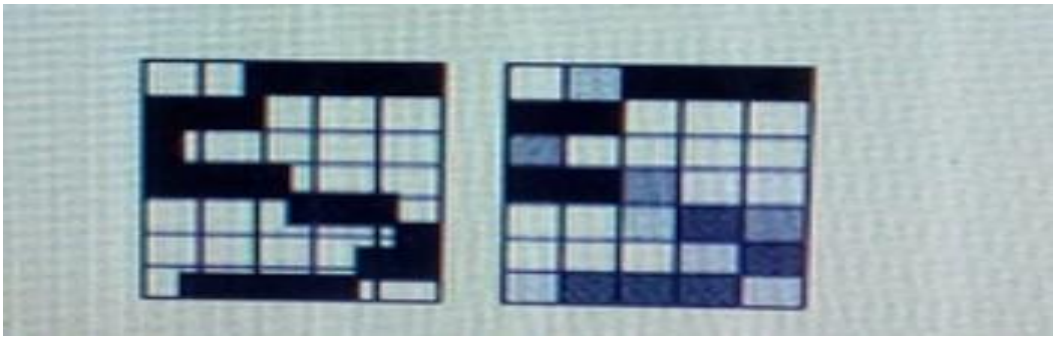
The sub-images have to be cropped sharp to the border of the character in order to standardize the sub-images. The image standardization is done by finding the maximum row and column with 1s and with the peak point, increase and decrease the counter until meeting the white space, or the line with all 0s. This technique is shown in figure below where a character "S" is being cropped and resized.




Cropped and resized picture the image pre-processing is then followed by the image resize again to meet the network input requirement, 5 by 7 matrices, where the value of 1 will be assign to all pixel where all 10 by 10 box are filled with 1s. Image resizes again to meet the network input requirement.

Finally, the 5 by 7 matrices are concatenated into a stream so that it can be feed into network 35 input neurons. The input of the network is actually the negative image of the figure, where the input range is 0 to 1, with 0 equal to black and 1

indicate white, while the value in between show the intensity of the relevant pixel. By this, we are able to extract the character and pass to another object for future classification or training purpose on the neuron network.



Creating Vectors data for the Neural Network (objects)

Creating and training of the Neural Network

Testing the Neural Network

***Application and Graphical User Interface:***

The character recognition application can be used to type every command inside the MATLAB console and workspace and also to use already pre-prepared Graphical User Interface. The GUI consists of files that include all necessary programming code, and also files that include visible interface shapes and forms. The interface employs the workflow of recognition process. After loading the image, we select the character and after that crop it, pre-process, extract feature and finally identify. On every stage, GUI shows us a new image, which is unique for the each step. The images can be viewed in the Main window, RGB, Binary, Crop to Edges and features window.

***Comparison:***

The image processed under mat lab platform in now compared with the images which have been already fed into the database. When the captured image matches with any of the images in the database, then that text is outputted as audio.

**3. Voice Module:**

This is an enhanced 8 channel recordable voice module. Each channel can hold up to 1 minute of recorded voice and/or music .The built-in microphone and push to record button makes recording easy and instant. The connection to external amplifiers, audio equipment, and paging systems are done using a line-level output jack.

## REFERENCES

[1] World Health Organization. (2009). 10 facts about blindness and visual impairment [Online]. Available: www.who.int/ features/factfiles/ blindness/blindness_facts/en/index.html

[2] T.Phan,P.Shivakumara,and C.L.Tan,“ALaplacianmethodforvideotext detection,” in Proc. Int. Conf. Document Anal. Recognit., 2009,pp. 66–70.

[3] International Workshop on Camera-Based Document Analysis and Recognition (CBDAR 2005, 2007, 2009, 2011). [Online]. Available: http://www.m.cs.osakafu-u.ac.jp/cbdar2011/

[4] N. Nikolaou and N. Papamarkos, “Color reduction for complex document images,” Int. J. Imaging Syst. Technol., vol. 19, pp. 14–26, 2009.

[5] KReader Mobile User Guide, knfb Reading Technology Inc. (2008). [On- line]. Available: http://www.knfbReading.com

[6] D. Dakopoulos and N. G. Bourbakis, “Wearable obstacle avoidance elec- tronic travel aids for blind: A survey,” IEEE Trans. Syst., Man, Cybern., vol. 40, no. 1, pp. 25–35, Jan. 2010.

[7] B. Epshtein, E. Ofek, and Y. Wexler, “Detecting text in natural scenes with stroke width transform,” in Proc. Compute. Vision Pattern Recognit., 2010, pp. 2963–2970.

[8] .A. Shahab, F. Shafait, and A. Dengel, "ICDAR 2011 robust reading com- petition: ICDAR Robust Reading Competition Challenge 2: Reading text in scene images," in Proc. Int. Conf. Document Anal. Recognit. 2011, pp. 1491–1496.

[9] R. Manduchi and J. Coughlan, "(Computer) vision without sight," Com- mun. ACM, vol. 55, no. 1, pp. 96–104, 2012.

[10] The Portset Reader, TVI Technologies for the Visually Impaired Inc., Hauppauge, NY, USA. (2012). [Online]. Available: http://www.tvi-web. com/products/porsetreader.html

**Author's profile:**

**R.Infant Abinaya** is currently doing her Fourth year B. TECH Information Technology at Jeppiaar Engineering College in Chennai, Tamil Nadu She has presented a paper based on Green computing technical symposium and has attended many workshops. Email – abi.santhyagu@gmail.com.

**E.Esakkiammal** is currently doing her Fourth year B. TECH Information Technology at Jeppiaar Engineering College in Chennai, Tamil Nadu .She has presented a paper based on Green computing in technical symposium and has attended many workshops. Email – shanthi140893@gmail.com.

**Mrs.L.Pushpalatha** is working as an Associate Professor in the Department of Information technology at Jeppiaar Engineering College in Chennai, Tamilnadu. Her area of interest is networking technology.