

# Raspberry Pi Based Reader for Blind Peoples

C.PON NIRMAL<sup>\*\*</sup>, M.BHARATH<sup>\*</sup>, P.SINDHA MATHAR<sup>\*</sup>, M.THAMARAİKANI<sup>\*</sup>, K.SRIDHAR<sup>\*</sup>

<sup>\*\*</sup> Assistant Professor, SCAD College of Engineering and Technology, Tirunelveli, India.

<sup>\*</sup>UG Scholar, SCAD College of Engineering and Technology, Tirunelveli, India.

**ABSTRACT:** Optical character recognition (OCR) is the identification of printed characters using photoelectric devices and computer software. It converts images of typed, handwritten or printed text into machine encoded text from scanned document or from subtitle text superimposed on an image. In this research these images are converted into audio output. OCR is used in machine process such as cognitive computing, machine translation, text to speech, key data and text mining. It is mainly used in the field of research in Character recognition, Artificial intelligence and computer vision. In this research, as the recognition process is done using OCR the character code in text files are processed using Raspberry Pi device on which it recognizes character using tesseract algorithm and python programming and audio output is listened. To use OCR for pattern recognition to perform Document image analysis

(DIA) we use information in grid format in virtual digital library's design and construction. This research mainly focuses on the OCR based automatic book reader for the visually impaired using Raspberry Pi. Raspberry Pi features a Broadcom system on a chip (SOC) which includes ARM compatible CPU and an on chip graphics processing unit GPU. It promotes Python programming as main programming language with support for BBC BASIC

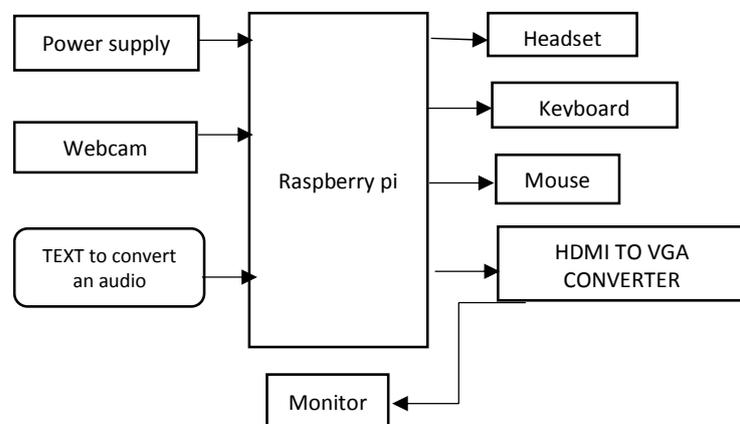
**KEYWORDS:** Energy efficient Character recognition, document image analysis(DIA), Raspberry Pi, audio output, OCR based book reader, python programming.

## INTRODUCTION

Gives an algorithm for detecting and reading text in natural images for the use of blind and visually impaired subjects walking through city scenes. The overall algorithm has a success rate of over 90% on the test set and the unread text is typically small and distant from the viewer. [2] have proposed a novel scheme for the extraction of textual areas of an image using globally matched wavelet filters. A clustering based technique has been devised for estimating globally matched wavelet filters using a collection of ground truth images. [3] proposes a support vector machine (SVM) is used to analyse the textual

properties of texts. The combination of CAMSHIFT and SVM's produces both robust and efficient text detection. [4] tells about the navigational technologies available to blind individuals to support independent travel, our focus is on blind navigation on large scale. [5] presents an approach to automatic detection and recognition of signs from natural scenes and its application to sign translation task that further propose a local intensity normalization method to effectively handle lighting variations followed by a Gabor transform to obtain local features. [6] presents a comparative survey among portable/wearable obstacle detection/avoidance systems to inform about the progress in assistive technology for visually impaired people.

## Block Diagram



Block diagram of book reader with Raspberry pi.

The above figure shows the block diagram of a book reader, a proto typed system, which reads the printed text on hand held objects on assisting the blind persons. When the application is started first it checks whether all the devices are available and also it checks for the connection. The graphical user interface (GUI) has than optional label for displaying the image from the camera, a status box for representing the image. The Raspberry board comes with integrated peripherals like USB, ADC and

Serial. On this board we are install the LINUX operating system with necessary drivers for all peripheral devices.

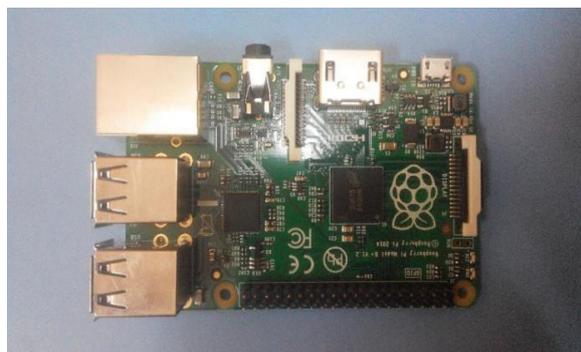
### 1.1. Working principle

When capture button is clicked this system captures the product image placed in front of the web camera which is connected to ARM microcontroller through USB .After selecting the process button the captured label image undergoes Optical Character Recognition(OCR) Technology. OCR technology allows the conversion of scanned images of printed text or symbols into text or information that can be understood or edited using a computer program. In our system for OCR technology we are using TESSERACT library. Using Flite library the data will be converted to audio. Camera acts as main vision in detecting the label image of the product or board then image is processed internally and separates label from image by using open CV library and finally identifies the product and identified product name is pronounced through voice. Now it identifies received label image is converted to text by using tesseract library. Once the identified label name is converted to text and converted text is displayed on display unit connected to controller. Now converted text should be converted to voice to hear label name as voice through ear phones connected to audio jack port using flite library.

### 1.2. Raspberry Pi – description

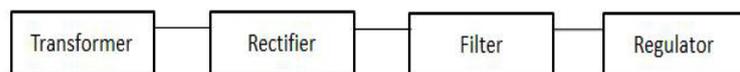
Raspberry Pi is a low cost, credit card sized computer that plugs computer monitor and TV and uses standard keyboard and mouse that uses python programming.

There are two models of Raspberry Pi, model A and model B. These two are bit similar with few advance features on model B compared to model A. Model B has 512 MB RAM, two USB port whereas Model A has 256 MB RAM and just a USB port. Besides, Model B has Ethernet port while Model A does not. Some of the components of the raspberry Pi include the SD card, Micro USB power, HDMI out, Ethernet and USB port, RCA video out and audio out, GPIO headers and chips.



### Hardware required for the Raspberry Pi

Transmitting, the hardware components of the Raspberry Pi include power supply, storage, input, monitor and network. POWER SUPPLY UNIT- It is the device that supplies electrical energy to the output loads



It gives a well regulated power supply of +5v with a output current compatibility of 100 mA

### 1.4. WEB CAM

A webcam is a video camera which feeds its images in real time to a computer or computer network, often via USB, Ethernet or Wi-Fi.

### 1.5. LENS

Webcams typically include a lens, an image sensor, support electronics, and may also include a microphone for sound.

## II. FEATURES OF LOGITECH WEBCAM

- Plug-and-play setup (UVC)
- Video capture: Up to 640 x 480 pixels
- Photos: Up to 1.3 megapixels (software enhanced)
- Frame rate: Up to 30 frames per second (with recommended system)
- Hi-Speed USB 2.0 certified
- Fixed focus
- Universal clip fits notebooks, LCD or CRT monitor

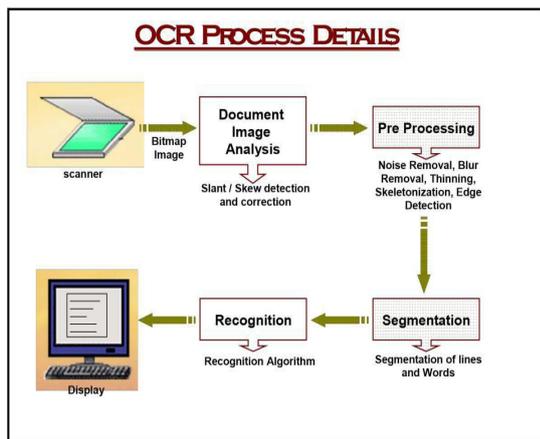
Logitech webcam.



## OPTICAL CHARACTER RECOGNITION (OCR)

Optical Character Recognition or OCR is the text recognition system that allows hard copies of written or printed text to be rendered into editable, soft copy versions. It is the translation of optically scanned bitmaps of printed or written text into digitally editable data files. An OCR facilitates the conversion of geometric source object into a digitally representable character in ASCII or Unicode scheme of digital character representation.

OCRs are of two types: for recognizing printed characters and for hand written text OCR PROCESS



OCR process.

The process include scanning, document image analysis (DIA), pre-processing ,segmentation and recognition.

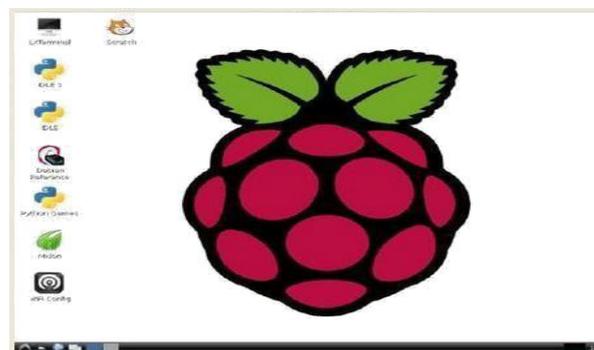
### SOFTWARE DESCRIPTION

#### Installation Of Operating System On Raspberry Pi

Raspberry Pi is a small computer; hence operating system (OS) should be installed. As the Raspberry doesn't have hard drive, OS is installed in the external memory. For that, memory card (SD card) is used for the installation of operating system and all the required software and supporting files are stored in the same SD card.

#### Use Of Desktop Screen, Keyboard And Mouse For Raspberry Pi

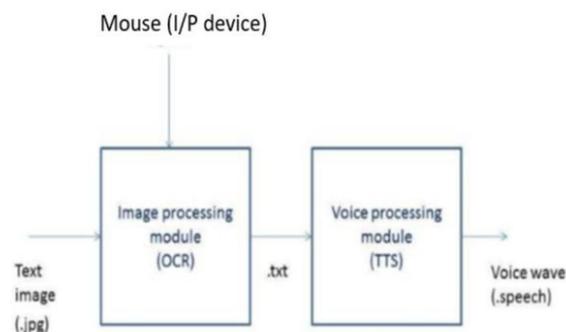
This is the first software need to be installed which can be downloaded from the link, Download Xming and install it in the laptop. After completion of installation, run the application called 'XLaunch' This is the primary software need to be installed. It can be downloaded in the provided link as follow, Download Putty. As, it is downloaded, it needs to be installed following some few normal steps of installation. For Configuration, double click the icon of Putty after the completion of installation and enter the IP address of Raspberry Pi as shown in the below figure.



Raspberry Pi.

#### METHODS

Text-to-speech device consists of two main modules, the image processing module and voice processing modules (Fig.2). Image processing module captures image using camera, converting the image into text. Voice processing module changes the text into sound and processes it with specific physical characteristics so that the sound can be understood.



#### Block Diagram of Text-To-Speech Device

Fig 2 shows the block diagram of Text-To-Speech device, 1st block is image processing module, where OCR converts .jpg to .txt form. 2nd is voice processing module which converts .txt to speech.

#### A. Image Processing Module Using Optical Character Recognition

OCR is important element in this module. OCR or Optical Character Recognition is a technology that automatically recognize the character through the optical mechanism, this technology imitates the ability of the human senses of sight, where the camera becomes a replacement for eye and image processing is done in the computer engine as a substitute for the human brain [2]. Tesseract OCR is

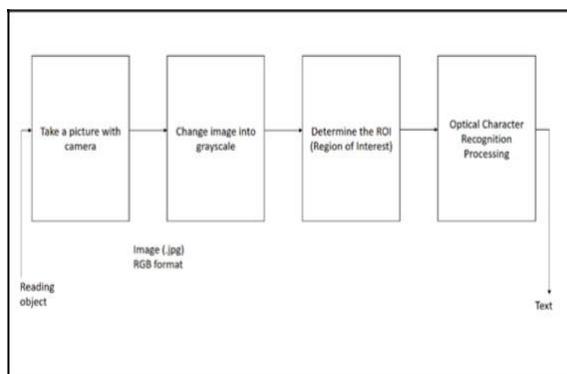
a type of OCR engine with matrix matching [3]. The selection of Tesseract engine is because of its flexibility and extensibility of machines and the fact that many communities are active researchers to develop this OCR engine and also because Tesseract OCR can support 149 languages. In this project we are identifying English alphabets. Before feeding the image to the OCR, it is converted to a binary image to increase the recognition accuracy. Image binary conversion is done by using Imagemagick software, which is another open source tool for image manipulation. The output of OCR is the text, which is stored in a file (speech.txt). Machines still have defects such as distortion at the edges and dim light effect, so it is still difficult for most OCR engines to get high accuracy text [4]. It needs some supporting and condition in order to get the minimal defect [1].

*Tesseract OCR Implementation-*

The input image captured by the Raspberry Pi camera module having a 5 MP sensor. Based on the specifications of the Tesseract OCR engine, the minimum character size that can be read is 20 pixels uppercase letters. Tesseract OCR accuracy will decrease with the font size of 14pt.

*Software Design-*

Software processes the input image and converted into text format. The software implementation is showed in



**Image Processing Module flow layout[2]**

The image is taken by the user via GPIO pin (23) that is connected to the button, using interrupt function. Furthermore, the picture is taken by using rasp still program with sharpness mode to sharpen the image. The resulting image has a .jpg format with a resolution of 2592 x 1944 pixel[1].

**B. The Voice Processing Module**

In this module text is converted to speech. The output of OCR is the text, which is stored in a file (speech.txt). Here, Festival software is used to convert the text to speech. Festival is an open source texttospeech (TTS) system, which is

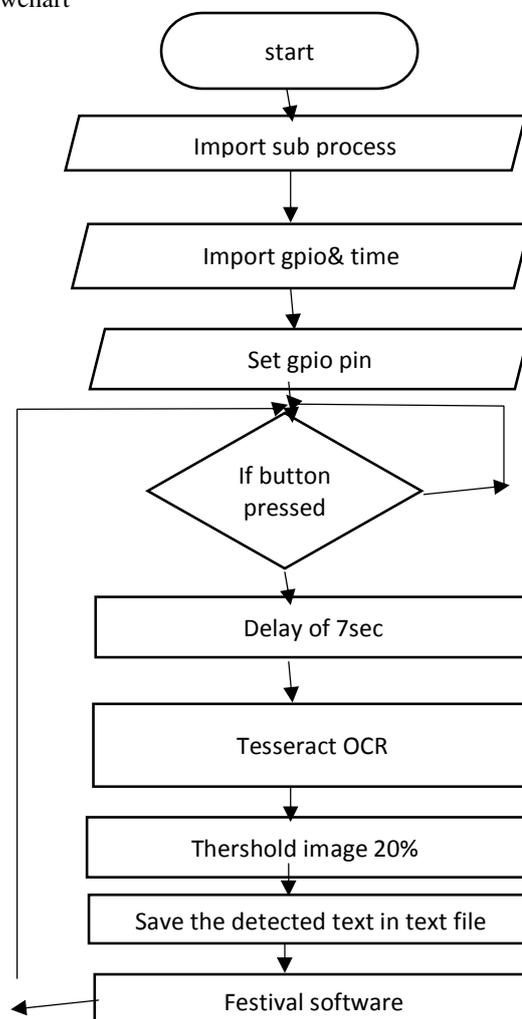
available in many languages. In this project, English TTS system is used for reading the text.

**DESIGN IMPLEMENTATION**

The testing was done using Raspberry Pi platform with the following specifications:

- Raspberry Pi 3 (Model B)
- Raspberry Pi Camera Module 5 MP
- Speakers
- Micro-USB charger (Power Supply)
- Peripherals (Mouse, Keyboard, Monitor)

The main program provides functions to retrieve and process the input image, convert it into a sound signal. Picture will be taken as soon as push button switch is pressed then this Captured image is thresholder before feeding it to OCR to increase the accuracy. Overall flow of program is done as in flowchart

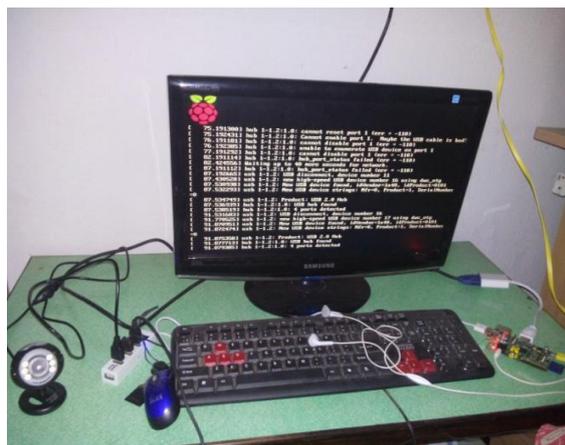


**Flow chart of text-to-speech device program[4]**

**IV. RESULT**

- Observed outcome of project:
- Text is extracted from the image and converted to audio.
- It recognizes both capital as well as small letters.
- It recognizes numbers as well.
- Range of reading distance was 35-45cm.
- Character font size should be minimum 18pt

- Signals, Systems & Automation, pp. 49-53,2009.
- "Raspberry Pi Cook Book" by Simon Monk.
- Text Localization and Extraction in Images Using Mathematical Morphology and OCR Techniques,2013.



**Figure 13:** Whole set up with Raspberry pi, keyboard, mouse, webcam and monitor.

## V. CONCLUSION

Text-to-Speech device can change the text image input into sound with a performance that is high enough and a readability tolerance of less than 2%, with the average time processing less than three minutes for A4 paper size. This portable device, does not require internet connection, and can be used independently by people. Through this method, we can make editing process of books or web pages easier.

## VI. REFERENCES

- ARCHANA A. SHINDE, D. "Text Pre-processing and Text Segmentation for OCR. International Journal of Computer Science Engineering and Technology", pp. 810-812, 2012.
- R. Mithe, S. Indalkar and N. Divekar." Optical Character Recognition" International Journal of Recent Technology and Engineering (IJRTE)", ISSN: 2277-3878, Volume-2, Issue-1, March 2013.
- R. Smith. "An Overview of the Tesseract OCR Engine", USA: Google Inc.
- H. Shah and A. Shah. "Optical Character Recognition of Gujarati Numerical", International Conference on

