Smart Reader for Visually Impaired Using Raspberry Pi

S Sarkar, G Pansare, B Patel, A Gupta, A Chauhan, R Yadav and N Battula*

School of Mechanical Engineering, Vellore Institute of Technology, Vellore, Tamil Nadu - 632014, India

*e-mail: nithin.battula@vit.ac.in

Abstract. With the humongous amount of texts and written media available today it has been increasingly necessary to make these available to people from all walks of life specially catering to the visually impaired people, thus devising a system which can assist in this task is of prime importance. The proposed device aims to solve this predicament by using Raspberry Pi module B+ to convert printed and hand written into easily accessible playable speech. It also stores speech which can be replayed at a time deemed suitable. The main focus of this paper is to develop a smart reader system which coverts hand written and printed text to speech. It has been observed that the scanned texts were converted into easily audible speech heard via the speaker with high efficiency.

1. Introduction

Today there is a huge amount of written material available everywhere, and converting that into Braille is not a suitable option due to the vastness of the content that is available and the lack of return on such a huge investment. But that would mean denying the blind an access to these huge quantities of scholarly material. A suitable and viable alternative for this predicament is designing a smart device which can convert this print media into speech format and it can in turn play it out for the visually impaired. Such a device is radical in its kind and is a huge benefit, although the initial investment may be high but in the long run it is pretty cost effective. The main objective of this paper is to make use of such a system specifically the Raspberry Pi system along with its ancillaries in unison and use it to convert text into speech and thus assist the visually impaired.

The proposed algorithm uses a camera module with which it can capture the desired text and then convert it to binary representation [1], i.e. converting the image into a gray-scale image. From this gray-scale image the individual characters are extracted and recognized all of which is carried out by the Optical Character Recognition Algorithm. Upon undergoing the processes of Scanning, Preprocessing, Segmentation and Feature Extraction, the scanned text is finally ready to give its output by means of the speaker connected to the Pi module. Even though such systems are present, most of them are in the crude forms and developing a commercially viable setup will be a huge aid for the visually impaired thus giving them access to unprecedented amounts of text and written media. Such as system which involves only a one-time investment is thus a vital assist tool. The main objective of this project is converting print and written media into playable audio with high efficiency. A unique addition in this device is to record speech in the memory and replay these audio files at a convenient time.
2. Methodology

The block diagram of the setup consists mainly of two input parts. The first part is the text to speech conversion and the next is where the camera module converts the scanned image file into speech. The entire project is implemented on the Raspberry Pi 3 board which consists of USB ports, GPIO pins for input and output, a dedicated Pi camera for scans, a micro SD card slot and finally the speaker which will give the output.

![Block diagram of the hardware setup](image)

**Figure 1.** Block diagram of the hardware setup

The entire module is powered by a USB cable connected to a computer. The press buttons are used to activate the program and the speaker for output [2]. The printed text to be scanned is then placed under the Pi camera to ensure a good quality image and few numbers of distortions. The acquired image is then transferred to an OCR algorithm whose task is to convert the captured image file into a playable text file. This OCR algorithm checks the scanned image for alphabets, numbers and characters and then gives a corresponding text output after comparing it with the pre-loaded database.

2.1. Hardware

The list of main hardware components used in this simulation are Pi Camera, Raspberry Pi3 Model B+, speaker/ headphone, push buttons, GPIO pins, bench Support, book, and monitor.

![Camera module](image)

**Figure 2.** Camera module

The Raspberry Pi V2.1 camera is based on the Sony IMX219 silicon CMOS back lit sensor which produces 8 megapixels images that are 3280 x 2464 pixels in size. The IMX219 sensor operates in the visible spectral range [3] i.e. 400-700 nm and uses a Bayer array with a BGGR pattern. There are two versions of the Pi camera, the regular version and the No-Infrared (NoIR) version. The NoIR version has a filter on the lens which allows to capture images beyond the visible range. The scope of this study though does not involve the NoIR version.

| Camera Parameter       | Specification       |
|------------------------|---------------------|
| Lens Focal Length      | 3.04 mm             |
| f-number               | 2                   |
| Instantaneous field of view | 0.368 mrad |
| Full-frame of view     | 59.17° (H) X 58.3° (V) |
The Raspberry Pi V2.1 camera module has a fixed focal length of 3.04 mm [3] and a single f-number of F2.0 which is typically focused from near-field to infinity. The images are captured at ISO settings between 100-800 which can be changed manually by increments of 100, which has though not been verified above 600. Some of the most significant feature of the Raspberry Pi camera version used has been shown is table 1.

The dimensions of the entire camera module are about 25mm x 25mm x 9mm and it weighs just about 3 grams. The camera connects to the Raspberry Pi board via a 15-pin mobile industry processor interface (MIPI). This is chosen because it is accessible easily and widely. The module has small physical size and the quality of storage can be modified as per requirements, it’s power-efficient and cost effective and it is possible to connect multiple cameras and process them parallelly [4].

The Raspberry Pi is a single credit card sized single board computer which has been developed by the Raspberry Pi Foundation, United Kingdom. This board has incredible computing power [5] and has the capability to do amazing projects. The device is perfect for performing all types of computing tasks and interfacing various types of devices via GPIO. The board consists of a Broadcom based ARM processor, a graphics chip, a RAM, GPIO and other connectors for external devices. It has Broadcom BCM2835 SoC processor with 700MHz ARM1176JZF-S core, 512 MB RAM, Videocore 4 GPU supports up to 1920x1200 resolution, MicroSD card slot, 10/100 Mbps Ethernet port, 4 x USB 2.0 ports, HDMI, audio/video jack, GPIO header containing 40 pins, MicroUSB power port providing 2A current supply, DSI and CSI ports.

Raspberry Pi operating system is installed in this device and Python is used as the main programming language but there is also support for BASIC, C, C++, JAVA, Perl and Ruby.

The main function of the speaker in the entire module is to give the final output of the scanned text and help in the completion of the motive behind the study. There are a few necessary conditions which have been ensured like recognition of a crowded place and modify the volume of sound accordingly and recognition of silent places and adjust the sound similarly [7].

The main function of speaker is to assist in text-to-speech conversion. gTTS module for python has is a Text-to-Speech (TTS) engine whose main function is to convert and thus create a spoken version of the text which has been scanned previously by the Pi camera. This TTS enables the reading of the text [8].

The main function of the push button is to switch something on or off depending on the necessity of the action. In this project a tangible push button is connected to a pin on the Raspberry Pi3, while the opposite side has a ground connection. When the button is pulled up the circuit is connected while when it is pushed down it is connected to the ground and the circuit is in off condition. These tactile push button switches can be easily programmed to provide a variety of output messages and switching messages.

The attaching of the tangible push button to the Raspberry Pi is an easy job, but the most important thing to remember that the GPIO pins are +3.3V DC compliant. If a voltage greater than +3.3 V DC is applied it will end up damaging the Raspberry Pi module.

An important and a very powerful feature of the Raspberry Pi is the row of GPIO (general purpose input/output) pins along the top edge of the board. The 40-pin GPIO is now found on all Raspberry Pi boards.

Any of the 40 GPIO pins can be designated as an input or output and they can be used for a wide range of applications. The board consists of two 5 V pins, three 3V pins and number of unconfigurable ground pins which have 0V.

The Optical Character Recognition (OCR) is the text recognition system which allows the hard copies and all other written text to be converted into editable soft copies [9]. An OCR facilitates the conversion of the geometric source object into digitally representable ASCII or Unicode scheme. OCR are of two types: for recognizing printed characters and for recognition of hand written text.

2.2 Software

The main open source programming language used in this Raspberry Pi project is Python [10]. Python is the very versatile, easily acquirable and has a simple syntax and thus the language becomes a very powerful tool in developing programs. Raspberry Pi also faces no compatibility issues while programming it via Python. Due to the wide range of libraries that are included by Python, it can be used for a variety of applications.
Python therefore comes pre-installed in Raspbian so that it can be used upon receiving. The two main choices for writing Python on the Raspberry Pi are using MU editor and editing remotely over SSH.

The MU editor is one of the pre-installed Python IDEs that can be used to write to the programs.

It often happens that a computer, keyboard and monitor are not always available to write the Python codes on Raspberry Pi. Then Raspberry Pi has a special feature which allows to connect to the module remotely over SSH.

The VLC media player is an open source cross platform media player and streaming software which has been implemented in our Raspberry Pi module; its objective being providing the platform via which the sound output is emitted.

3. Results and discussion

Figure 3 is the image of the original simulation setup which has been done during the course of this study. The assembly of the components worked in unison and gave the desired output.

![Simulation setup](image)

**Figure 3. Simulation setup**

3.1 Python code

3.1.1 Text to speech conversion procedure. The text to speech conversion is mainly divided into two main modules, the image processing and the voice processing. The “picamera” module is used to capture images via the camera and convert images into text using “pytesseract” OCR image processing module of python. The voice processing module (gTTS) converts this scanned text into sound which has specific characteristics so that it can be understood by the intended audience.

![Block Diagram of TTS](image)

**Figure 4. Block Diagram of TTS**

Figure 4 shows the block diagram of a typical block diagram of the text to speech conversion module, where the first block is the image processing module. Here the image processing module converts image into texts and the next is the voice processing module which converts this text into voice.

The design of the entire module involves a fully integrated system where all devices do their tasks in harmony, the camera feeds the input text and it forms a digitized document. The image processing module processes the scanned document. The image processing module can recognise the sequence of characters present and can also indicate whether the current reading line is enabled [11]. The scanned
images may vary in size some to a few idioms whereas some may be full sentences, therefore storage space management is very necessary and thus space digitization [12] of texts is very helpful. The image processing module methodology also makes use of methods like machine-learning and text-to-speech conversion to capture the image and finally this recognized speech is translated to the speaker for output.

3.1.2 Text displayed and speaker output. The hand-written text shown in Figure 5 is the actual sample of the written media which had been scanned by the Raspberry Pi camera module. The speaker readout the following text without any mistakes. It was tested with various other type of text for different hand writings and it recognised all of them almost all the time and gave the correct output.

![Figure 5. Hand written text shown to Pi camera module.](image)

The Raspberry Pi Module does not have a language barrier, it can be successfully programmed to process any language which the user finds comfortable working with. We tested only for English language as modules are available for it. It has to be tested further for other languages. We observed that it works correctly 90 % of the time because of the OCR algorithm. Moreover, a special bookmark feature can be enabled such that any important information can be highlighted for future reference.

4. Conclusion
We developed a smart reader to assist the visually impaired. The device has demonstrated all the desired working criteria and thus the success of the device can be considered. Such a reading assistant has its scope limited not only to the visually impaired, but it can be widely used by people from all walks who want to have the luxury of a book reader available to them. This device thus acts both as a necessity and as a comfort supplement. The one-time cost of the entire device is around $70 (₹ 5000) and once installed successfully it has the potential to be the ideal assistant for the people it has been chiefly intended to. This device has known to find its importance in schools & libraries and also in many other fields where a voice assistant is needed.

References
[1] Aditi S, Annapoorani S P, Kanchana A 2018 Book Reader Using Raspberry Pi for Visually Impaired Int. Res. J. Eng. Tech. 5(03).
[2] George L, Rinsila S, Baby R, Thomas A, 2020 Raspberry Pi based Reader for Blind, J. Opt. Commun. Electron. 5(03) 11–16.
[3] Pagnutti M A, Ryan R E, Gold M J, Harlan R, Leggett E and Pagnutti J F 2017 Laying the foundation to use Raspberry Pi 3 V2 camera module imagery for scientific and engineering purposes J. Electron. Imaging, 26(1) 013014.
[4] Anandhalli M, Baligar V P 2018 A Novel Approach in real time vehicle detection and tracking using Raspberry Pi Alex. Eng. J. 57(3) 1597–1607.
[5] Nayyar A, Puri V 2015 Raspberry Pi-A Small Powerful, Cost effective and efficient form factor computer: A review, Int. J. Adv. Res. Comput. Sci. Software. Eng. 5(12) 720–737.
[6] Sachdeva P, Katchii S 2014 A Review Paper on Raspberry Pi Int. J. Curr. Eng. Tech. 4(6) 3818-19.
[7] Nwazor N O 2019 A Raspberry Pi Based Speaker Recognition System for Access Control Int. Res. J. Eng. Tech. 6(3).
[8] Mallesh S J, Basavaraj H, Gagan L, Prateek R P 2018 Smart Speaker using Raspberry Pi Int. J. Trend Sci. Res. Dev. 2(4) 1682–85.
[9] Aaron J S, Sanjana S, Monisha M 2016 OCR based automatic book reader for visually impaired using Raspberry Pi Int. J. Innovative Res. Comp. and Commun. Eng. 4(7) 1111-18.
[10] Hoyo A, Guzman J L, Moreno J C 2015 Teaching Control Engineering Concepts using Open Source tools on a Raspberry Pi board *IFAC-PapersOnLine* 48(29) 99–104.

[11] Jojie A, George A and Nayana, J D D Book Reader for Blind. *IOSR J. Eng. ISSN (e)*, 2250-3021.

[12] Mallapa D G, Shruti S S, Shruti B H, Vijayalaxmi I B 2017 "B-LIGHT: A reading aid for the blind people using OCR and OpenCV" *Int. J. Sci. Res. Eng. Tech.* 6(5) 546–48.