

Drushti- A Smart Reader for Visually Impaired People

Shobha Sharma ¹, Dhanush.R.Bhat ², Neerendra.R.Hegde ³, Yajnesh A.P ⁴, Mythri. D ⁵

Assistant Professor ¹, UG Student ^{2,3,4,5}

Department of Information Science and Engineering SDM Institute of Technology,
Ujire, Karnataka, India

Abstract:- According to the World Health organization (WHO), 285 million people are estimated to be visually impaired worldwide, among which 90% live in developing countries and forty five million are blind individuals worldwide. Though there are many existing solutions to the problem of assisting individuals who are blind to read .In particular, there is a need for a portable text reader that is affordable and readily available to the blind community. This project proposes a smart reader for visually challenged people using Raspberry Pi. This paper addresses the integration of a complete Text Read-out system designed for the visually challenged. A camera will be used to take input, speaker and LCD to give output. The system consists of a webcam interfaced with Raspberry Pi which accepts a page of printed text. The OCR (Optical Character Recognition) package installed in Raspberry Pi scans it into a digital document. Once it is scanned, the text is read out by a text to speech conversion unit (TTS engine) installed in Raspberry Pi. The output is fed to an audio amplifier before it is read out. The image to text conversion and text to speech conversion is done by the OCR software installed in Raspberry Pi. The system finds its interesting applications in libraries, auditoriums, offices where instructions and notices are to be read and also assists in filling of application forms.

Keywords:- Raspberry Pi, OCR(Optical Character Recognition), TTS(Text to Speech) Engine, Web Camera.

1. INTRODUCTION

An Embedded System is a combination of computer hardware and software, perhaps additional mechanical parts, designed to perform a specific function. An embedded system is a microcontroller-based, software driven, reliable, real-time control system, autonomous, human / network interactive, operating on diverse physical variables in diverse environments sold into a competitive and cost conscious market.

We present a smart device that assists the visually impaired which effectively and efficiently reads paper-printed text. The proposed project uses the methodology of a camera based assistive device that can be used by people to read Text document. The framework is on implementing image capturing technique in an embedded system based on Raspberry Pi board. The design is motivated as it is small-scale and mobile, which enables a more manageable operation with minimal setup. In this project we have proposed a text read out system for the visually challenged. The

proposed integrated system has a camera as an input device to feed the printed text document for its conversion into a gray scale image and the scanned document is processed by a software module known as OCR (optical character recognition engine). As part of the software development, the Open CV (Open source Computer Vision) libraries is utilized to do image capture of text for character recognition. Most of the access technology tools built for people with blindness and limited vision are built on two basic building blocks of OCR software and Text-to-Speech (TTS) engines. Optical character recognition (OCR) is the translation of captured images of printed text into machine encoded text. It is defined as the process of converting scanned images of machine printed into a computer processable format. The final recognized text document is fed to the output device depending on the choice of the user. The output device can be a headset connected to the Raspberry Pi board or a speaker which can spell out the text document aloud.

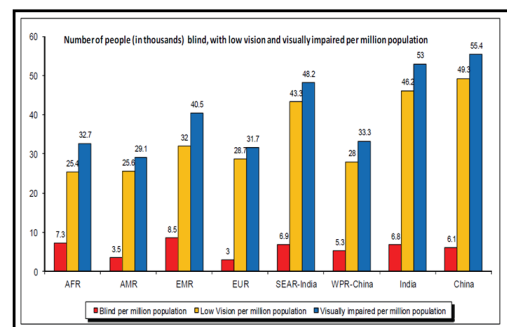


Figure 1. Prevalence of Blindness as per estimate of 2017

According to Figure 1 , In our planet of 7.4 billion humans, 40% are visually impaired out of which 6% people are completely blind, i.e. have no vision at all, and 35% have mild or severe visual impairment (WHO, 2017). It has been predicted that by the year 2020, these numbers will rise to 75 million blind and 200 million people with visual impairment. There have been numerous efforts in this area to help visually impaired to read without difficulties. By this project, we would be able to detect the text effectively and efficiently which would work towards the benefit of these people.

2. LITERATURE SURVEY

According to Bindu Philip and R. D. Sudhaker Samuel

paper on “Human Machine Interface-A Smart OCR for the visually challenged”. The integration of a complete Malayalam Text Read-out system was designed for the visually challenged. The system accepts a page of printed Malayalam text with English numerals, scans it into a digital document which is then subjected to skew correction, segmentation, before feature extraction to perform classification. Once classified, the text in Malayalam is read out by a text to speech conversion unit.

A paper by V. Ajantha Devi, Dr. S Santhosh Baboo on “Embedded Optical Character Recognition on Tamil Text Image using Raspberry Pi”. Optical Character recognition is used to digitize and reproduce texts that have been produced with non-computerized system. Digitizing texts also helps reduce storage space.

A paper by J.N. Balaramkrishna , J.Geetha on ”The Smart Reader from Image using OCR and Open CV with Raspberry Pi 3” This kind of system helps visually impaired people to interact with computers effectively through vocal interface. Text-to-Speech is a device that scans and reads English alphabets and numbers that are in the image using OCR technique and changing it to voices.

A paper by Asha G. Hagargund, Sharsha Vanria Thota, Mitadru Bera, Eram Fatima Shaik on “Image to speech conversion for visually Impaired” The device that proposed aims to help people with visual impairment. A device that converts an image text to speech. The basic framework is an embedded system that captures an image, extracts only the region of interest (i.e. region of the image that contains text) and converts that text to speech. It is implemented using a Raspberry Pi and a Raspberry Pi camera Two tools are used convert the new image (which contains only the text) to speech. They are OCR (Optical Character Recognition) software and TTS (Text-to-Speech) engines. The audio output is heard through the raspberry pi’s audio jack using speakers or earphones.

3. BLOCK DIAGRAM

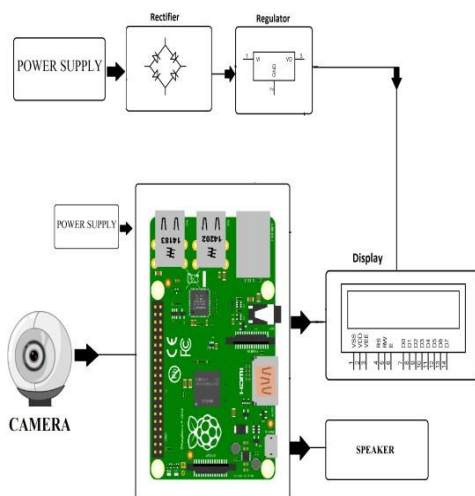


Figure 2. Block Diagram of Proposed System

The Figure 2 illustrates the block diagram of our proposed system. The framework for the proposed system is the Raspberry Pi board. The Raspberry Pi 3 B+ is a single board computer which has 4 USB ports, an Ethernet port for internet connection, 40 GPIO pins for input/output, CSI camera interface, HDMI port, DSI display interface, SOC (system on a chip), LAN controller, SD card slot and an audio jack. The power supply is given to the 5V micro USB connector of Raspberry Pi through the Switched Mode Power Supply (SMPS). The SMPS converts the 230V AC supply to 5V DC. Web Camera is connected to the USB port of Raspberry Pi. Raspberry Pi has an OS named RASPBIAN which process the conversions. The audio output is taken from the audio jack of the Raspberry Pi. The converted speech output is amplified using an audio amplifier. A Power Supply Unit is a device that supplies electrical energy to the output loads.

A Power Supply is also given to the LCD Display for Display Purpose. The Capacitors, Resistors and Voltage Regulators are all embedded on a PCB Board and is then connected to a Power Source for functioning of the LCD Display. The PCB Board is connected to the Raspberry Pi through GPIO Pins in order to bring about collaboration between Raspberry Pi and Printed Circuit Board.

The Document to be read is placed on a base and the camera is focused to capture the image. The captured image is processed by the OCR software installed in Raspberry Pi. The captured image is converted to text by the software. The text is converted into speech by the TTS engine. The final output is given to the audio amplifier from which it is connected to the speaker. Speaker can also be replaced by a headphone .

4. SOFTWARE REQUIREMENTS

- **Raspbian OS**– It is an Operating System to get the Raspberry Pi started.
- **IDLE**- Python Integrated Development and Learning Environment.
- **OpenCV**- Open source Computer Vision libraries is utilized to do image capture of text
- **OCR**- Optical character recognition is the translation of captured images of printed text into machine-encoded text.
- **TTS Engine** - A Text-To-Speech (TTS) synthesizer is a computer-based system that should be able to read any text aloud, whether it was directly introduced in the computer by an operator or scanned image.

5. METHODOLOGY

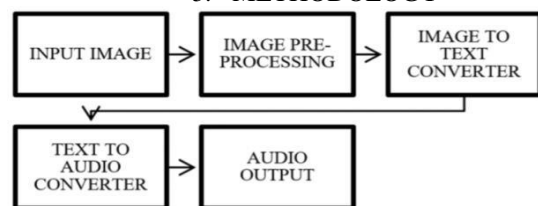


Figure 3. The Methodology Stages

According to Figure 3 our Proposed Project has been divided into the Following Stages:

1. Input Image
2. Image Pre-Processing
3. Image To Text Converter
4. Text to Audio Converter
5. Audio Output



Figure 4. Project Setup

The Figure 4 illustrates the Project Setup of our Device Drushti on the basic external connections done according to the Block Diagram.

5.1 Input Image



Figure 5. LCD Display Indicating to Press a Switch



Figure 6. LCD Display Indicating that a Image is Captured

The Figure 5 illustrates the LCD Display indicating to press a Switch. Once the Switch is pressed, the Web Camera connected to the USB Port of Raspberry Pi activates itself to capture the Image which is placed on the Base. The Figure 6 illustrates the LCD Display indicating that an image is captured. At this stage an image is captured through the Web Camera and the captured image becomes the Input Image. Before the Image is captured, the Web Camera must be properly focused for high image clarity which helps in better recognition of Characters. If the Camera is inefficient to capture the image then we would receive a message displaying “No Image Captured” on the LCD Display.

5.2 Image Pre-Processing

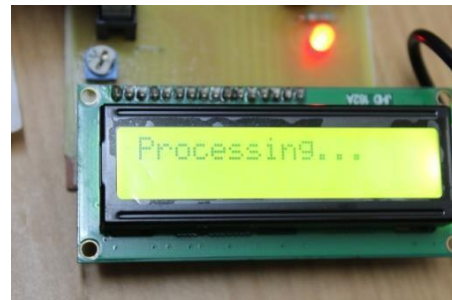


Figure 7. LCD Display Indicating that Image is getting processed

The Figure 7 illustrates the LCD Display indicating that the image is being processed. At this stage the captured image will be converted from Colour Image to a Gray Scale Image. The Conversion into a Gray Scale Image will blur the Background of the Image. This helps in Enhancing the Text from the Gray Scale Background and makes it easy for the OCR Software to recognize the characters.

5.3 Image to Text Converter

The Image to Text Conversion is done using the Optical Character Recognition Software which is translation of captured images of printed text into machine-encoded text. The image is submitted as input to an OCR engine. The OCR mainly consists of two process which is Feature Extraction and Classification. The initial process involves recognition of features of the character in the image, then the classification is carried out. The OCR engine matches portions of the image to shapes instructed to recognize. Given logic parameters that the OCR engine has been instructed to use, the OCR engine will make its best guess as to which letter a shape represents. OCR results are written as text.

5.4 Text to Audio Converter

The Text to Audio Conversion is done using the Text to Speech Synthesizer which is a computer- based system that should be able to read any text aloud, whether it is directly introduced in the computer by an operator or scanned image. A text-to-speech system is composed of two parts: a front-end and a back-end. The front-end has two major tasks. Firstly, it converts raw text containing symbols like numbers and

abbreviations into an equivalent written-out words. This process is often called as text normalization, pre-processing, ortokenization. The front-end then assigns phonetic transcriptions to each word, divides and marks the text into units, like phrases, clauses, and sentences. The process of assigning phonetic transcriptions to words is called text-to-phoneme conversion. The back- end often referred to as the synthesizer—then converts the symbolic linguistic representation into sound.

5.5 Audio Output



Figure 8. LCD Display Indicating that the Text has been converted into Speech

The Figure 8 illustrates the LCD Display indicating that the Text is being read out loud through the Audio Speaker. We can also use Earphones in order to hear the output.

6. ADVANTAGES

- Text is extracted from the image and converted to audio.
- Case Sensitive.
- Numeric Recognition.

7. DISADVANTAGES

- Range of reading distance was 30-40cm.
- Character font size should be minimum 14pt.
- Maximum tilt of the text line is 4-5 degree from the vertical.

8. CONCLUSION

We have implemented an image to speech conversion technique using Raspberry Pi. Our algorithm successfully processes the image and reads it out clearly. This is an economical as well as efficient device for the visually impaired people. We have applied our algorithm on many images which has succeeded. The device is compact and helpful to the society.

9. FUTURE SCOPE

In the future we can use more Robust and Efficient algorithms to read the image and separate the text from the images. The Captured Image was blur and there is a need to de-blur the Image in less time so that we can separate the data efficiently and convert them into speech. By considering all these aspects our proposed project is going to work towards the benefit of the society and would benefit the visually impaired as well as Blind.

10. REFERENCES

- [1] Bindu Philip and r. d. Sudhaker Samuel 2009 "Human machine interface – a smart ocr for the visually challenged" International journal of recent trends in engineering, vol no.3, November .
- [2]. K Nirmala Kumari, Meghana Reddy J [2016]. Image Text to Speech Conversion Using OCR Technique in Raspberry Pi. International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering Vol. 5, Issue 5, May 2016.
- [3] V. Ajantha devi, dr. Santhosh baboo "Embedded optical character recognition on tamil text image using raspberry pi" international journal of computer science trends and technology (ijest) – volume 2 issue 4, jul-aug 2014
- [4] Jaiprakash verma, khushali desai "Image to sound conversion" International journal of advance research.
- [5] 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.
- [6] Character Detection and Recognition System for Visually Impaired People by Akhilesh A. Panchal, Shrugal Varde, M.S. Panse.