ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

Smart Text Reader from Image Using OCR and OpenCV with Raspberry PI 3

J. N. Balaramakrishna
M.Tech (Embedded Systems),
Department of ECE
Jogaiah Institute of Technology and Sciences College of Engineering,
Kalagampudi, Palakol - 534 268,
West Godavari Dt., A.P.

Ms. J. Geetha, M.Tech
Assistant Professor
Department of ECE
Jogaiah Institute of Technology and Sciences College of Engineering,
Kalagampudi, Palakol - 534 268,
West Godavari Dt., A.P.

ABSTRACT
In this project an innovative, efficient and real-time cost beneficial technique that enables user to hear the contents of text images instead of reading through them as been introduced. It combines the concept of Optical Character Recognition (OCR) and Text to Speech Synthesizer (TTS) in Raspberry pi. This kind of system helps visually impaired people to interact with computers effectively through vocal interface. Text Extraction from color images is a challenging task in computer vision. 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. This paper describes the design, implementation and experimental results of the device. This device consists of two modules, image processing module and voice processing module. The device was developed based on Raspberry Pi v2 with 1.2 GHz processor speed.

INTRODUCTION
In our planet of 7.4 billion humans, 285 million are visually impaired out of whom 39 million people are completely blind, i.e. have no vision at all, and 246 million have mild or severe visual impairment (WHO, 2011). It has been predicted that by the year 2020, these numbers will rise to 75 million blind and 200 million people with visual impairment [7].

There have been numerous advances in this area to help visually impaired to read without many difficulties. The existing technologies use a similar approach as mentioned in this project, but they have certain drawbacks. Firstly, the input images taken in previous works have no complex background, i.e. the test inputs are printed on a plain white sheet. It is easy to convert such images to text without pre-processing, but such an approach will not be useful in a real-time system [1][2][3]. Also, in methods that use segmentation of characters for recognition, the characters will be read out as individual letter and not a complete word. This gives an undesirable audio output to the user. For our project, we wanted the device to be able to detect the text from any complex background and read it efficiently. Inspired by the methodology used by Apps such as “CamScanner”, we assumed that in any complex background, the text will most likely be enclosed in a box eg billboards, screens etc. By being able to detect a region enclosing four points, we assume that this is the required region containing the text. This is done using warping and cropping. The new image obtained then undergoes edge detection and a boundary is then drawn over the letters. This gives it more definition. The image is then processed by the OCR and TTS to give audio ouput.

Optical character Recognition (OCR) is a conversion of scanned or printed text images, handwritten text into editable text for further processing. In this paper, we have presented a robust approach for text extraction and convert it to speech. Testing of device was done on raspberry pi platform. The Raspi is initially connected to the internet through VLAN. The software is installed using command lines. The first setup is to download the

ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

installation script, second command is to convert it to executable form and the last command starts the script which does the rest of the installation work. The webcam is manually focused towards the text. Then, to take a picture, press pushbutton switch. A delay of around 7 seconds is provided, which helps to focus the webcam, if it is accidently defocused. After delay, picture is taken and processed by Raspi to hear the spoken words of the text through the earphone or speaker plugged into Raspi through its audio jack.

OBJECTIVE OF THE PROJECT

As reading is of prime importance in the daily routine (text being present everywhere from newspapers, commercial products, sign-boards, digital screens etc.) of mankind, visually impaired people face a lot of difficulties. Our device assists the visually impaired by reading out the text to them.

AIM OF THE PROJECT

The main aim of this project is to convert the text in the textual image into the speech efficiently. For this project we are using the Raspberry Pi 3 processor, which supports OpenCV libraries and some image processing algorithms. The program was written in the Python scripting Language.

EXISTING SYSTEM

In the earlier days, there are no efficient text to speech recognition techniques available. As the digital image processing techniques evolved, it is possible to convert the textual data in the image into the speech by using the several synthesizers and the image processing algorithms but those not efficient and take time to convert the text into speech and only the letter-wise speech is possible. To avoid these difficulties we are developing the proposed system.

PROPOSED SYSTEM

In the proposed system, we developed an efficient text to speech conversion technique by using the Raspberry Pi 3 processor. When the text image was captured by the camera, the synthesizer used to separate the text from the image and then the Optical Character Recognition algorithm was implemented to recognize the characters in the text and then the Raspberry Pi 3 was the responsible to convert that text into speech by using the OpenCV libraries.

BLOCK DIAGRAM

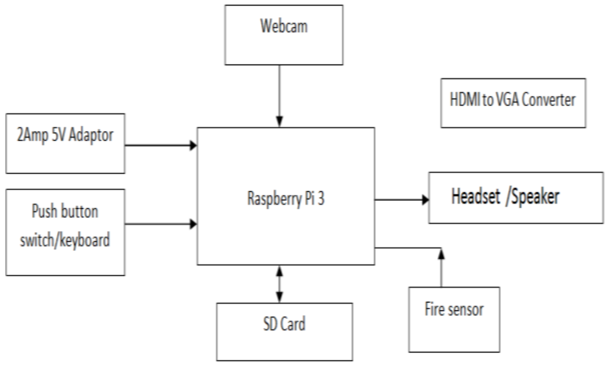

Fig 2.1: Block diagram of proposed system

Initially we should focus the USB camera to the textual image. Whenever you want to hear the text in the image then we should press a button, then in turn the Raspberry Pi 3 processor triggers the USB camera to take the snap shot of that image. The camera will take few seconds to focus and then the capture the image and send to the Raspberry Pi 3. Then the Text to Speech Synthesizer (TTS) will separate the textual data from the image. Then the algorithm, OCR, Optical Character Recognition will recognize that characters in the text and then given to them to the Pi. The Pi will convert that text into the speech and play them through the 3.5mm Audio jack of the Raspberry Pi board.

RASPBERRY PI 3

The Raspberry Pi is a credit-card sized computer that plugs into your TV and a keyboard. It is a capable little computer which can be used in electronics projects, and for many of the things that your desktop PC does, like spreadsheets, word-processing and games. It also plays high-definition video. We want to see it being used by kids all over the world to learn how computers work, how to manipulate the electronic world around them, and how to program.

ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

The Raspberry Pi is a low cost, credit-card sized computer that plugs into a computer monitor or TV, and uses a standard keyboard and mouse. It is a capable little device that enables people of all ages to explore computing, and to learn how to program in languages like Scratch and Python. It's capable of doing everything you'd expect a desktop computer to do, from browsing the internet and playing high-definition video, to making spreadsheets, word-processing, and playing games.

What's more, the Raspberry Pi has the ability to interact with the outside world, and has been used in a wide array of digital maker projects, from music machines and parent detectors to weather stations and tweeting birdhouses with infra-red cameras. We want to see the Raspberry Pi being used by kids all over the world to learn to program and understand how computers work.

There are currently four Raspberry Pi models. They are the Model A, the Model B, the Model B+ and the Compute Module. All models use the same CPU, the BCM2836, but other hardware features differ.

The Model B+

Released in July 2014, the Model B+ is a updated revision of the Model B. It increases the number of USB ports to 4 and the number of pins on the GPIO header to 40. In addition, it has improved power circuitry which allows higher powered USB devices to be attached and now hot plugged. The full size composite video connector has been removed and the functionality moved to the 3.5mm audio/video jack. The full size SD card slot has also been replaced with a much more robust micro SD slot.

The following list details some of the improvements over the Model B.

- Current monitors on the USB ports mean the B+ now supports hot plugging.
- Current limiter on the 5V for HDMI means HDMI cable powered VGA converters will now all work
- 14 more GPIO pins
- EEPROM readout support for the new HAT expansion boards
- Higher drive capacity for analog audio out, from a separate regulator, which means a better audio DAC quality.
- No more back powering problems, due to the USB current limiters which also inhibit back flow, together with the "ideal power diode"
- Composite output moved to 3.5mm jack
- Connectors now moved to two sides of the board rather than the four of the original device.
- Ethernet LED's moved to the Ethernet connector
- 4 squarely positioned mounting holes for more rigid attachment to cases etc.

The power circuit changes also means a reduction in power requirements of between 0.5W and 1W.

General Purpose I/O (GPIO)

General Purpose Input/Output pins on the Raspberry Pi This page expands on the technical features of the GPIO pins available on BCM2836 in general. For usage examples, see the GPIO Usage section. When reading this page, reference should be made to the BCM2836 ARM Peripherals Datasheet. GPIO pins can be configured as either general-purpose input, general-purpose output or as one of up to 6 special alternate settings, the functions of which are pin-dependant.

Power-On States

All GPIOs revert to general-purpose inputs on power-on reset. The default pull states are also applied, which are detailed in the alternate function table in the ARM peripherals datasheet. Most GPIOs have a default pull applied.

Interrupts

Each GPIO pin, when configured as a general-purpose input, can be configured as an interrupt source to the ARM. Several interrupt generation sources are configurable:

- Level-sensitive (high/low)
- Rising/falling edge
- Asynchronous rising/falling edge

ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

Level interrupts maintain the interrupt status until the level has been cleared by system software (e.g. by servicing the attached peripheral generating the interrupt).

The normal rising/falling edge detection has a small amount of synchronisation built into the detection. At the system clock frequency, the pin is sampled with the criteria for generation of an interrupt being a stable transition within a 3-cycle window, i.e. a record of "1 0 0" or "0 1 1". Asynchronous detection bypasses this synchronisation to enable the detection of very narrow events.

Alternative Functions

Almost all of the GPIO pins have alternative functions. Peripheral blocks internal to BCM2836 can be selected to appear on one or more of a set of GPIO pins, for example the I2C busses can be configured to at least 3 separate locations. Pad control, such as drive strength or Schmitt filtering, still applies when the pin is configured as an alternate function.

The block diagram for an individual GPIO pin is given below:

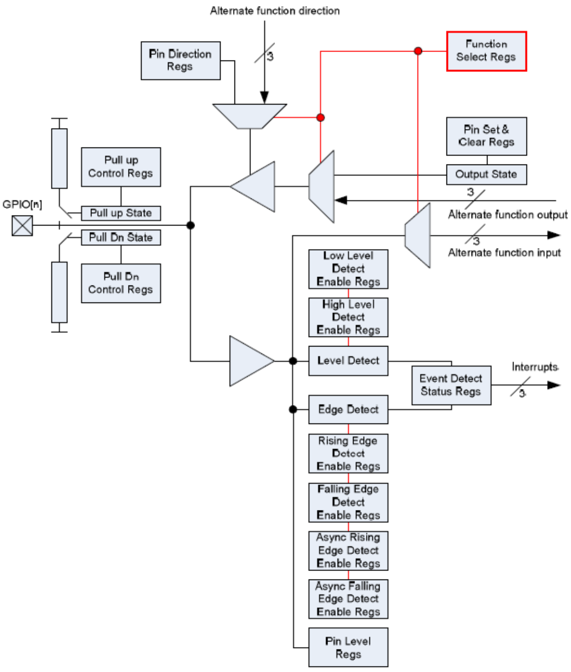

Fig 2.5: GPIO block diagram

RESULTS

The implementation of realization of "Image Text to Speech Conversion Using OCR Technique in Raspberry Pi" is done successfully. The communication is properly done without any interference between different modules in the design. Design is done to meet all the specifications and requirements.

PROPOSED SYSTEM RESULTS

Initially we should focus the USB camera to the textual image. Whenever you want to hear the text in the image then we should press a button, then in turn the Raspberry Pi 3 processor triggers the USB camera to take the snap shot of that image.

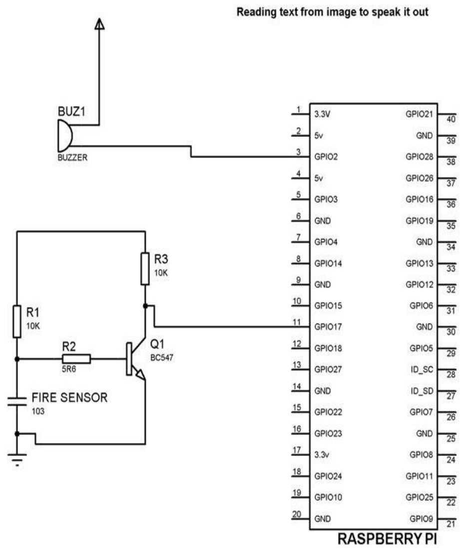

Fig 5.1: Schematic diagram of proposed system

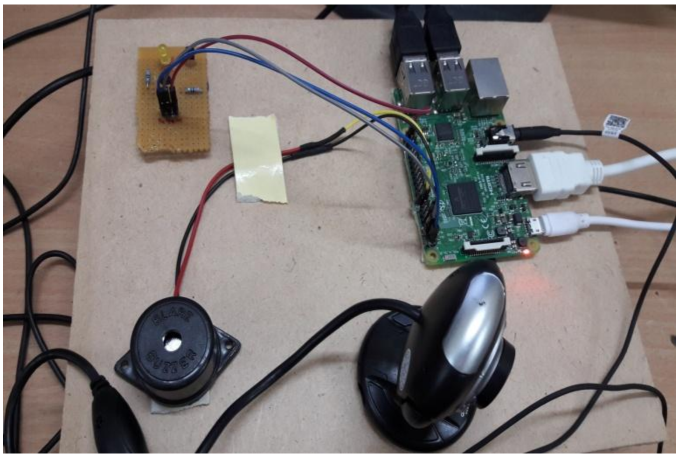

Fig 5.2: Proposed system

ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

The camera will take few seconds to focus and then the capture the image and send to the Raspberry Pi 3. Then the Text to Speech Synthesizer (TTS) will separate the textual data from the image. Then the algorithm, OCR, Optical Character Recognition will recognize that characters in the text and then given to them to the Pi. The Pi will convert that text into the speech and play them through the 3.5mm Audio jack of the Raspberry Pi board.

In this project, we are adding the fire sensor to the Raspberry Pi 3 processor I/O pin. If that sensor gets activates then the Raspberry Pi 3 will play the corresponding audio message through the audio jack. This project is very useful to the visually disabled persons.

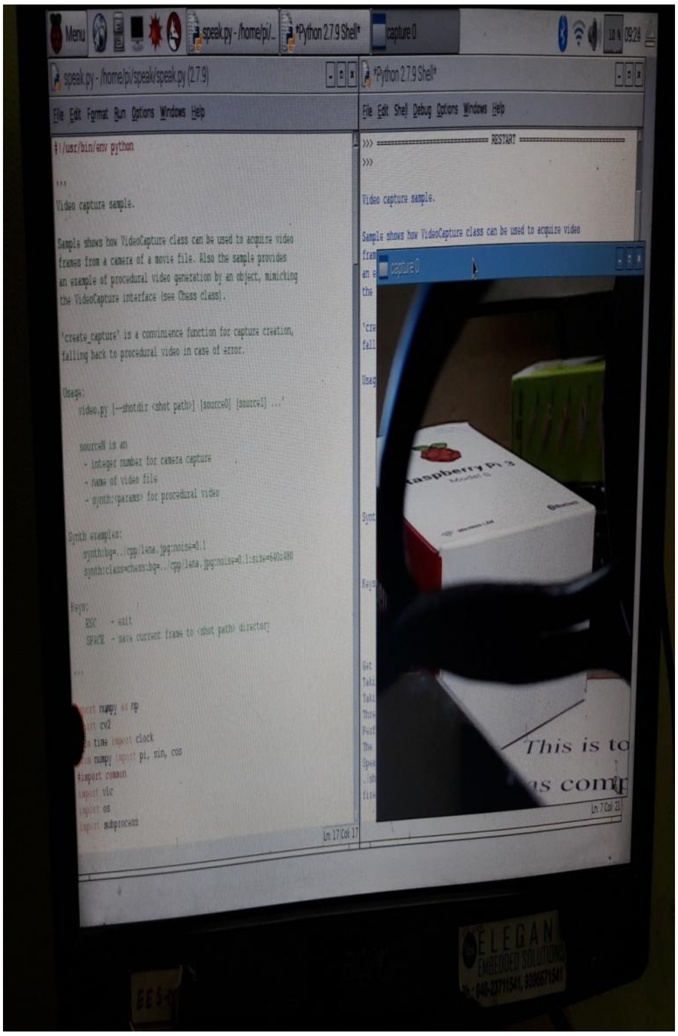

Fig 5.3: Proposed video capture

Fig 5.4: Speech synthesis

Fig 5.5: Interfacing web camera with Raspberry pi

ADVANTAGES AND APPLICATIONS

ADVANTAGES

1. Text is extracted from the image and converted to audio.
2. It recognizes both capital as well as small letters.
3. It recognizes numbers as well.
4. Range of reading distance was 38-42cm.
5. Character font size should be minimum 12pt.
6. Maximum tilt of the text line is 4-5 degree from the vertical.

APPLICATIONS

1. Hospitals
2. Home Security applications
3. All security applications

ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

CONCLUSION

The system enables the visually impaired to not feel at a disadvantage when it comes to reading text not written in braille. The image pre-processing part allows for the extraction of the required text region from the complex background and to give a good quality input to the OCR. The text, which is the output of the OCR is sent to the TTS engine which produces the speech output. To allow for portability of the device, a battery may be used to power up the system.

FUTURE SCOPE

In the future we can use more robust and the efficient algorithms to read the images and separate the text from the images. We the captured image was blurring, and then also we will de-blur the image in less time and can separate the data efficiently to convert them to the speech.

REFERENCES

[1]. D.Velmurugan, M.S.Sonam, S.Umamaheswari, S.Parthasarathy, K.R.Arun[2016]. A Smart Reader for Visually Impaired People Using Raspberry PI. International Journal of Engineering Science and Computing IJESC Volume 6 Issue No. 3.

[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]. Silvio Ferreira, C´eline Thillou, Bernard Gosselin. From Picture to Speech: An Innovative Application for Embedded Environment. Faculté Polytechnique de Mons, Laboratoire de Théorie des Circuits et Traitement du Signal Bˆatiment Multitel - Initialis, 1, avenue Copernic, 7000, Mons, Belgium.

[4]. Nagaraja L, Nagarjun R S, Nishanth M Anand, Nithin D, Veena S Murthy [2015]. Vision based Text Recognition using Raspberry Pi. International Journal of Computer Applications (0975 – 8887) National Conference on Power Systems & Industrial Automation.

[5]. Poonam S. Shetake, S. A. Patil, P. M. Jadhav [2014] Review of text to speech conversion methods.

[6]. International Journal of Industrial Electronics and Electrical Engineering, ISSN: 2347-6982 Volume-2, Issue-8, Aug.-2014.

[7]. S. Venkateswarlu, D. B. K. Kamesh, J. K. R. Sastry, Radhika Rani [2016] Text to Speech Conversion. Indian Journal of Science and Technology, Vol 9(38), DOI: 10.17485/ijst/2016/v9i38/102967, October 2016.

[8]. World Health Organization. 10 facts about blindness and visual impairment. 2015. Available from: http://www.who. int/features/factfiles/blindness/blindness_facts/en/

[9]. http://elinux.org/RPi_Text_to_Speech_(Speech_Synthesis)

[10] ARCHANA A. SHINDE, D. "Text Pre-processing and Text Segmentation for OCR. International Journal of Computer Science Engineering and Technology", pp. 810-812, 2012.

[11] 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.

[12] R. Smith. !An Overview of the Tesseract OCR Engine", USA: Google Inc.

[13] H. Shah and A. Shah. "Optical Character Recognition of Gujarati Numerical", International Conference on Signals, Systems &Automation, pp. 49-53,2009.

[14] Raspberry pi cook book by simon monk.

[15] Text Localization and Extraction in Images Using Mathematical Morphology and OCR Techniques, 2013.