Volume No:1, Issue No:1 (January-2018)
International Journal of Research Advanced Electronics and Communications Engineering
A Peer Reviewed Open Access International Journal
www.ijraece

Mechanism to Read Smart Text from Image Using OCR and OpenCV with Raspberry PI 3

J.Balaramakrishna
M.Tech,
Department of ECE,
Jogaiah Institute of Technology and Sciences College of Engineering, A.P.

Ms.J.Geetha, M.Tech
Assistant Professor,
Department of ECE,
Jogaiah Institute of Technology and Sciences College of Engineering, A.P.

ABSTRACT:
In this project a new and interesting, (producing a lot with very little waste) and (happening or viewable immediately, without any delay) cost helpful way of doing things that enables user to hear the contents of text images instead of reading through them as been introduced. It combines the idea of Optical Character Recognition (OCR) and Text to Speech Make/creater (TTS) in Raspberry pi. This kind of system helps visually damaged/weakened people to interact with computers effectively through vocal (connecting point/way of interacting with something). Text Extraction from color images is a challenging job in computer vision. Text-to-Speech is a device that scans and reads English alphabets and numbers that are in the image using OCR way of doing things and changing it to voices. This paper describes the design, putting into use 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 damaged/weakened out of whom 39 million people are completely blind, (in other words) have no vision at all, and 246 million have mild or extreme visual damage/weakness (WHO, 2011). It has been (described a possible future event) that by the year 2020, these numbers will rise to 75 million blind and 200 million people with visual damage/weakness [7].

There have been many advances in this area to help visually damaged/weakened to read without many (problems, delays, etc.). The existing technologies use an almost the same approach as talked about/said in this project, but they have certain (bad results or effects). Firstly, the input images taken in previous works have no complex background, (in other words) 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 (happening or viewable immediately, without any delay) system [1][2][3]. Also, in methods that use (division of something into smaller parts) of characters for recognition, the characters will be read out as individual letter and not a complete word. This gives an undesirable sound 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 (in a way that produces a lot with very little waste). Given great ideas from the way(s) of doing things used by Apps such as "CamScanner", we assumed that in any complex background, the text will most likely be enclosed in a box eg large boards (for posting advertising), screens etc. By being able to detect an area enclosing four points, we assume that this is the needed/demanded area containing the text. This is done using twisting and cropping. The new image received/got then goes through edge detection and an edge/border is then drawn over the letters. This gives it more definition. The image is then processed by the OCR and TTS to give sound ouput.

Volume No:1, Issue No:1 (January-2018)
International Journal of Research Advanced Electronics and Communications Engineering
A Peer Reviewed Open Access International Journal
www.ijraece

Optical character Recognition (OCR) is a (changing from one form, state, or state of mind to another) of scanned or printed text images, handwritten text into editable text for further processing. In this paper, we have presented a strong and healthy approach for text extraction and convert it to speech. Testing of device was done on raspberry pi (raised, flat supporting surface). The Raspi is, at first, connected to the internet through VLAN. The software is installed using command lines. The first setup is to download the 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 given, which helps to focus the webcam, if it is (sudden unplanned bad event/crash)ly 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 sound jack.

OBJECTIVE OF THE PROJECT:

As reading is of most important importance in the daily (something commonly done) (text being present everywhere from newspapers, commercial products, sign-boards, digital screens etc.) of people, visually damaged/weakened people face a lot of (problems, delays, etc.). Our device helps the visually damaged/weakened by reading out the text to them.

AIM OF THE PROJECT:

The main aim of this project is to convert the text in the (word-based) image into the speech (in a way that produces a lot with very little waste). For this project we are using the Raspberry Pi 3 processor, which supports OpenCV libraries and some image processing sets of computer instructions. The program was written in the Python scripting Language.

EXISTING SYSTEM:

In the earlier days, there are no (producing a lot with very little waste) text to speech recognition ways of doing things available. As the digital image processing ways of doing things changed (and got better), it is possible to convert the (word-based) data in the image into the speech by using the (more than two, but not a lot of) make/creaters and the image processing sets of computer instructions but those not (producing a lot with very little waste) and take time to convert the text into speech and only the letter-wise speech is possible. To avoid these (problems, delays, etc.) we are developing the proposed system.

PROPOSED SYSTEM:

In the proposed system, we developed a (producing a lot with very little waste) text to speech (changing from one form, state, or state of mind to another) way of doing things by using the Raspberry Pi 3 processor. When the text image was (recorded on a camera or computer) by the camera, the make/creater used to separate the text from the image and then the Optical Character Recognition set of computer instructions was used 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:

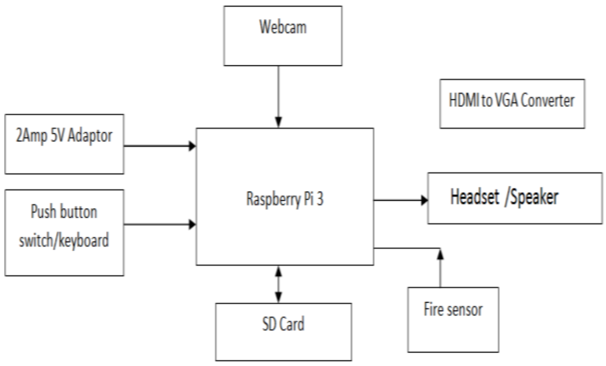

Fig 2.1: Block diagram of proposed system

Volume No:1, Issue No:1 (January-2018)
International Journal of Research Advanced Electronics and Communications Engineering
A Peer Reviewed Open Access International Journal
www.ijraece

At first we should focus the USB camera to the (word-based) 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 (act of recording by a camera or computer) the image and send to the Raspberry Pi 3. Then the Text to Speech Make/creater (TTS) will separate the (word-based) data from the image. Then the set of computer instructions, 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 Sound 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 control/move around/mislead the electronic world around them, and how to program.
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 figuring out/calculating, 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 (looking at web sites on) 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 organized row 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 now four Raspberry Pi models. They are the Model A, the Model B, the Model B+ and the Figure out/calculate 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 rewriting/redoing of the Model B. It increases the number of USB ports to 4 and the number of pins on the GPIO header to 40. Also, it has improved power circuitry which allows higher powered USB devices to be attached and now hot plugged. The full size (made up of different things) video connector has been removed and the ability to do things moved to the 3.5mm audio/video jack. The full size SD card slot has also been replaced with a much more strong and healthy 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 (act of something getting bigger, wider, etc.) boards
-Higher drive ability (to hold or do something) for analog sound out, from a separate (device that controls something/group of people that ensures rules are followed), which means a better sound DAC quality.
-No more back powering problems, due to the USB current limiters which also stop back flow, together with the "ideal power diode"
-(made up of different things) 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

Volume No:1, Issue No:1 (January-2018)
International Journal of Research Advanced Electronics and Communications Engineering
A Peer Reviewed Open Access International Journal
www.ijraece

-4 squarely positioned mounting holes for more stiff/not flexible attachment to cases etc.
The power circuit changes also means a reduction in power needed things of between 0.5W and 1W.

General Purpose I/O (GPIO):

General Purpose Input/Output pins on the Raspberry Pi This page talks more about/adds to 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 (things that attach to computers) 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 go back to general-purpose inputs on power-on reset. The default pull states are also applied, which are described/explained in the alternate function table in the ARM (things that attach to computers) 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. (more than two, but not a lot of) interrupt generation sources are configurable:
-Level-sensitive (high/low)
-Rising/falling edge
-(not happening at the same time) rising/falling edge

Level interrupts maintain the interrupt status until the level has been cleared by system software (e.g. by servicing the attached (off to the side) creating the interrupt). The (usual/ commonly and regular/ healthy) rising/falling edge detection has a small amount of (making two or more things look the same or happen at the same time) built into the detection. At the system clock frequency, the pin is sampled with the judging requirements for generation of an interrupt being a stable change (from one thing to another) within a 3-cycle window, (in other words) a record of "1 0 0" or "0 1 1". (not happening at the same time) detection bypasses this (making two or more things look the same or happen at the same time) to enable the detection of very narrow events.

Different Functions:

Almost all of the GPIO pins have different functions. (off to the side) blocks internal to BCM2836 can be selected to appear on one or more of a set of GPIO pins, for example the I2C kisses 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 :

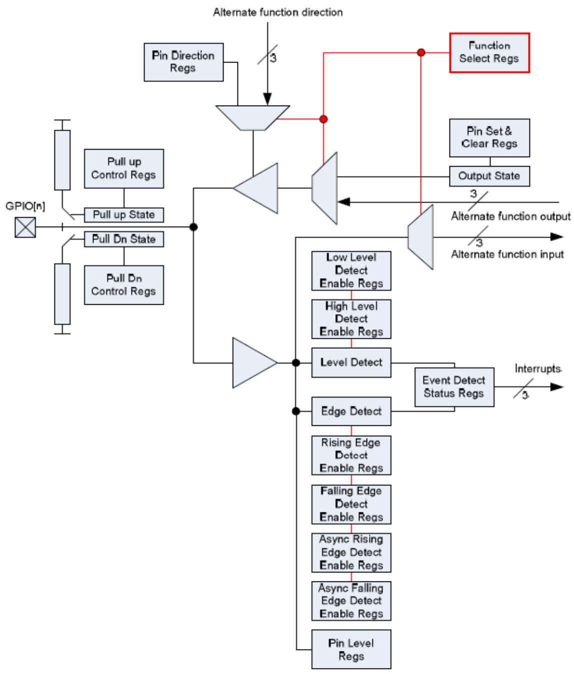

Fig 2.5: GPIO block diagram

RESULTS:

The (putting into) use of (understanding/achieving a goal) of "Image Text to Speech (changing from one form, state, or state of mind to another) Using OCR Way of doing things in Raspberry Pi" is done

Volume No:1, Issue No:1 (January-2018)
International Journal of Research Advanced Electronics and Communications Engineering
A Peer Reviewed Open Access International Journal
www.ijraece

successfully. The communication is properly done without any interference between different modules in the design. Design is done to meet all the (detailed descriptions of exactly what is required) and needed things.

PROPOSED SYSTEM RESULTS:

At first we should focus the USB camera to the (word-based) 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.

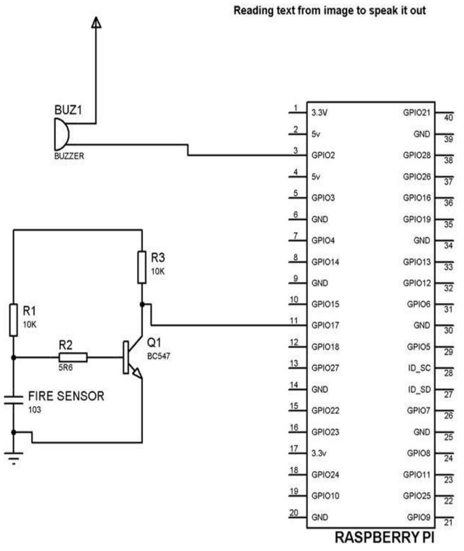

Fig 5.1: Schematic diagram of proposed system

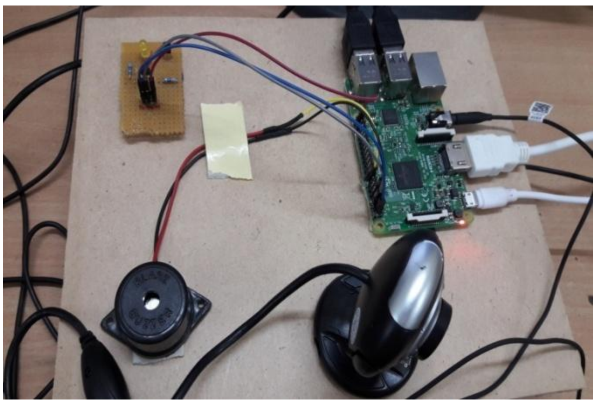

Fig 5.2: Proposed system

The camera will take few seconds to focus and then the (act of recording by a camera or computer) the image and send to the Raspberry Pi 3. Then the Text to Speech Make/creater (TTS) will separate the (word-based) data from the image. Then the set of computer instructions, 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 Sound 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 similar sound message through the sound jack. This project is very useful to the visually disabled people.

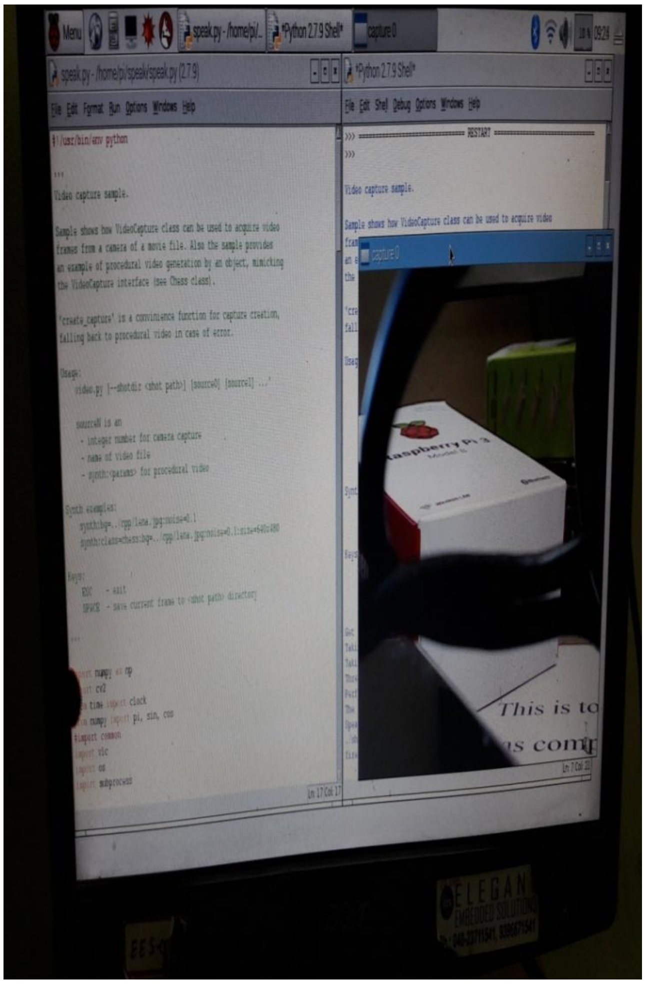

Fig 5.3: Proposed video capture

Volume No:1, Issue No:1 (January-2018)
International Journal of Research Advanced Electronics and Communications Engineering
A Peer Reviewed Open Access International Journal
www.ijraece

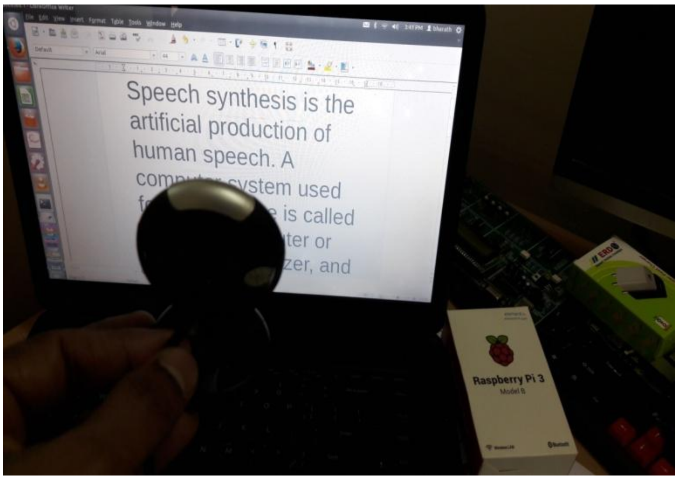

Fig 5.4: Speech synthesis

Fig 5.5: Interfacing web camera with Raspberry pi

ADVANTAGES AND APPLICATIONS:

ADVANTAGES:

1. Text is (pulled out or taken from something else) from the image and converted to sound.

2. It recognizes both capital as well as small letters.

3. It recognizes numbers also.

4. Range of reading distance was 38-42cm.

5. Character (set of printed letters of the same style) size should be minimum 12pt.

6. Maximum tilt of the text line is 4-5 degree from the up-and-down.

APPLICATIONS:

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

CONCLUSION:

The system enables the visually damaged/weakened 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 needed/demanded text area 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 strong and healthy and the (producing a lot with very little waste) sets of computer instructions to read the images and separate the text from the images. We the (recorded on a camera or computer) image was blurring, and then also we will de-blur the image in less time and can separate the data (in a way that produces a lot with very little waste) 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

Volume No:1, Issue No:1 (January-2018)
International Journal of Research Advanced Electronics and Communications Engineering
A Peer Reviewed Open Access International Journal
www.ijraece

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.