OCR Algorithm Using Raspberry Pi3 For Destitute of Vision

Praveena.S[1], Sowmiya.N[2], Thenmozhi.M[3], Mr. Manoharan.L[4]
[1, 2, 3] Dept of Electronics And Communication Engineering
[4]assistant Professor, Dept Of Electronics And Communication Engineering
[1, 2, 3, 4] Jeppiaar Srr Engineering College

Abstract- An OCR (Optical Character Recognition) is a system which is a branch of computer vision and in turn a sub-class of Artificial Intelligence. Optical character recognition is the translation of optically scanned bitmaps of printed or hand-written text into audio output by using of Raspberry pi. OCRs developed for many world languages are already under efficient use. This method extracts moving object region by a mixture-of-Gaussians-based background subtraction method. A text localization and recognition are conducted to acquire text information. To automatically localize the text regions from the object, a text localization and Tesseract algorithm by learning gradient features of stroke orientations and distributions of edge pixels in an Adaboost model. Text characters in the localized text regions are then binaries and recognized by off-the-shelf optical character recognition software. The recognized text codes are output to blind users in speech. Performance of the proposed text is localization algorithm. As the recognition process is completed, the character codes in the text file are processed using Raspberry pi device which recognize the character using Tesseract algorithm and python programming, the audio output is listed.

Keywords- OCR Technology, Raspberry Pi3, HDMI to VGA converter, Tesseract algorithm.

I. INTRODUCTION

Used for the detection and reading of documented text in images to help the blind and visually impaired people. The overall algorithm has a success rate of 90% on the test set as the unread text is significantly small and distant from the camera. We have proposed a technique to extract text from typed documents, convert them into machine encoded text, create the text files and then process them using Digital Image Analysis (DIA0) to convert the text into audio output. Our focus is on enhancing the capabilities of blind people by providing them a solution so that the information can be fed to them in the form of a speech signal. An OCR (Optical Character Recognition) is a branch of computer vision and a sub-class of Artificial Intelligence. Optical character recognition is the translation of optically scanned bitmaps of printed into audio output by using of Raspberry Pi. OCRs

developed for many world languages are already under efficient use. This method extracts moving object region by a mixture-of-Gaussians-based background subtraction method. A text localization and recognition are conducted to acquire text information. The recognized text codes are output to blind users in speech. The recognition process is completed, the character codes in the text file are processed using Raspberry pi device, which recognize character using Tesseract algorithm and python programming, and the audio is listened as the output. This project can also be implemented for the automatic detection of road signs, warning signs, in other terms to improve the blind navigation on larger scale.

II. EXISTING SYSTEM

In existing approach, it is a method to design a Text to Speech conversion module by the use of Mat lab by simple matrix operations. Firstly, by the use of microphone some similar sounding words are recorded using a record program in the Mat lab window and recorded sounds are saved in. wave format in the directory. The recorded sounds are then sampled and the sampled values are taken and separated into their constituent phonetics. The separated syllables are then concatenated to reconstruct the desired words. By the use of various Mat lab commands i.e. wave read, subplot etc. the waves are sampled and extracted to get the desired result. This method is simple to implement and involves much lesser use of memory spaces.

We develop a framework for reconstructing images that are sparse in an appropriate transform domain from polychromatic computed tomography (CT) measurements under the blind scenario where the material of the inspected object and incident-energy spectrum are unknown.

The existing navigation systems for the blind people require a precise GPS map. This make them unusable in region where there are no GPS maps, they are not sufficiently accurate. Algorithm for GPS navigation for the visually impaired along a GPS track, which describe the path as a sequence of waypoints is proposed. The natural voice navigation, adaptive to the velocity and accuracy of the GPS

data, start of the navigation from any waypoint, correlation of the direction of movement if it is necessary, return the user to the route if deviation is deviated, work with and without electronic compass, detection

DISADVANTAGE

- Braille system is very slow and not very practical.
- Existing OCR systems are not automatic.
- Require full-fledged computers to run effective.
- Cell phone and allows the user to read mail, receipts, files.

III. PROPOSED SYSTEM

In proposed system we have described a prototype system to read printed text on hand-held objects for assisting blind persons. In order to solve the common aiming problem for blind users, we have proposed a motion-based method to detect the object of interest, while the blind user simply shakes the object for a couple of seconds. The automatic ROI detection and text localization algorithms were independently evaluated as unit tests to ensure effectiveness and robustness of the whole system. We subsequently evaluated this prototype system of assistive text reading using images of hand-held objects captured by ten blind users in person. Two calibrations were applied to prepare for the system test. First, we instructed blind users to place hand-held object within the camera view. Since it is difficult for blind users to aim their held objects, we employed a camera with a reasonably wide angle. In future systems, we will add finger point detection and tracking to adaptively instruct blind users to aim the object. Second, in an applicable blind-assistive system, a text localization algorithm might prefer higher recall by sacrificing some precision. By using ultrasonic sensor, we will measure the distance between the blind people and obstacle then the distance will be played through ear phones

Optical Character Recognition is the electronic conversion of images of typed, handwritten or printed text into machine-encoded text, whether form a scanned document, a photo of a document, a scene-photo or form subtitle text superimposed on an image. It is widely used as a form of information entry from printed paper. It is a method of digitizing printed texts so they can be electronically edited, searched, stored, more compact, displayed on-line and used in machine process such as cognitive computing, machine translation, text-to-text, key data and text mining is a field of research in pattern recognition, artificial intelligence and computer vision. Early versions needed to be trained with images of each character, and worked on one font at a time. Advanced systems capable of producing a high degree of recognition accuracy for most fonts are now common, and with support for a variety of digital image file format inputs. Some systems are capable of reproducing formatted output that closely approximates the original page including images, columns, and other non-textual component.

ADVANTAGES

- It is inexpensive.
- Automatic system for reading text books will be implemented that not only converts printed books to digital text, but also reads them as an audio output.
- Our proposed algorithm can effectively handle complex background and multiple patterns.
- Extract text information from both hand-held objects and nearby signage.
- Character recognition is an art of detecting segmenting and identifying characters from image.
- The advancement of automation process and improving the interface between man and machine in many applications.

IV. BLOCK DIAGRAM
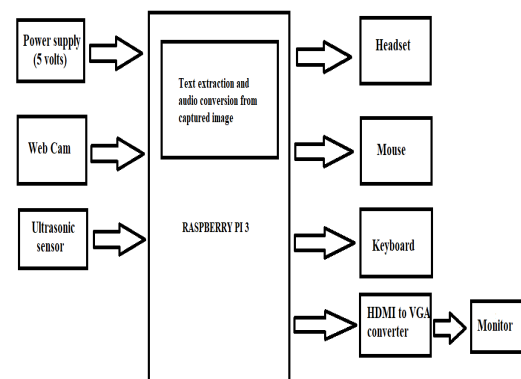
V. WORKING DESCRIPTION

When capture button is clicked this system captures the image placed in front of the web camera which is connected to raspberry pi3 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 this system for OCR technology TESSERACT library were used. Using this 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 raspberry pi3. Now converted text should be converted to voice to hear name as voice through ear phones connected to audio jack port.

Ultrasonic sensors are used for obstacle detection and calculation of distance between the obstacle and the visually impaired person. It consists of transmitter and receiver. The transmitter emits eight 40 kHz pulse, this pulse after hitting the obstacle is received back at the receiver, it records the time taken by the emitted pulse to return back at the receiver end. The system is successful in warning the user about the presence of obstacles in their path. It can detect any object within a pre-specified minimum distance. For out tests, we set the minimum distance value to 0.1m. The system announces the distance calculated in real time in meters or centimeters through headphones.

VI. HARDWARE USED

RASPBERRY PI3

The Raspberry Pi is a series of small single-board computers developed in the United Kingdom by the Raspberry Pi Foundation to promote teaching of basic computer science in schools and in developing countries.
There are two models of it, Raspberry Pi 2 and Raspberry Pi 3. These two are bit similar with few advance features on Pi 3. Compared to the Raspberry Pi 2 it has

- A 1.2 GHz 64-bit quad-core ARMv8 CPU
- 802.11 Wireless LAN
- Bluetooth 4.1
- 4 USB Ports
- 40 GPIO pins
- Full HDMI Port
- Ethernet port
- Camera interface
- Display interface
- Micro SD Card Slot

The Raspberry Pi board comes equipped with an SD card. The SD card is a main storage device for raspberry pi board like a hard disk of a personal computer.

ULTRASONIC SENSOR

Ultrasonic sound vibrates at a frequency above the range of human hearing. Ultrasonic sensor, use a single to send a transducer pulse and to receive the echo. The sensor determines the distance to a target by measuring time lapses between the sending and receiving of the ultrasonic pulse. It consists of four pins.

HDMI TO VGA CONVERTER

HDMI- High definition multimedia interface.VGA- Video graphic array. It is the standard monitor or display interface used in most PC's. HDMI TO VGA CONVERTER is used to connect the monitor and the raspberry kit in order to convert the image into audio output.

TEXT TO SPEECH CONVERTER:

Text to speech, abbreviated as TTS, is a form of speech synthesis that converts text into spoken voice output. Text to speech systems were first developed to aid the visually impaired by offering a computer-generated spoken voice that would "read" text to the user. TTS systems, in contrast, are theoretically capable of "reading" any string of text characters to form original sentences.

MONITOR:

A computer monitor is an output device that displays information in pictorial form. A monitor usually comprises the display device, circuitry, casing, and power supply.

Monitors are connected to the computer via VGA, Digital VisualInterface (DVI), HDMI, DisplayPort, Thunderbolt, low-voltage differential signaling(LVDS) or other proprietary connectors and signals.

HEADSET:

Headsets are a pair of small loudspeaker drivers worn on or around the head over a user's ears. They are electro acoustic transducers, which convert an electrical signal to a corresponding sound. Headphones let a single user listen to an audio source privately, in contrast to a loudspeaker, which emits sound into the open air for anyone nearby to hear.

Headphones are also known as ear speaker, ear phones or, colloquially, cans. Circumaural ('around the ear') and supra-aural ('over the ear') headphones use a band over the top of the head to hold the speakers in place.

WEB CAM:

Webcam is a small digital video camera directly or indirectly connected to a computer or a computer network. A webcam is an input device that captures digital images. These are transferred to the computer. A webcam is generally connected by a USB cable, or similar cable, or built into computer hardware, such as laptops. Webcams are known for their low manufacturing cost and their high flexibility, making them the lowest-cost form of video telephony.

Despite the low cost, the resolution offered at present is rather impressive, with low-end webcams offering resolutions of 320×240, medium webcams offering 640×480 resolution, and high-end webcams offering 1280×720 (720p) or even 1920×1080 (1080p) resolution. The most popular use of webcams is the establishment of video links, permitting computers to act as videophones or videoconference stations. Other popular uses include security surveillance, computer vision, video broadcasting, and for recording social videos. The video streams provided by webcams can be used for a number of purposes, by using appropriate software

VII. SOFTWARE USED

Operating System -Raspbian OS
Language Used   -Python

RASPBIAN OPERATING SYSTEM

Raspbian is an operating system based on Debian optimized for the Raspberry Pi hardware. An operating system contains a set of basic programs which is used to run Raspberry Pi 3. Raspbian OS consist of 35,000packages, pre-compiled software bundled in a nice format for easy installation on Raspberry Pi 3.

Raspbian OS is one of the official Operating systems available for free to download and use. The system is based on Debian Linux and is optimized to work efficiently with the Raspberry Pi computer. Debian is very lightweight and makes a great choice for the Pi. The Raspbian includes tools for browsing, python programming and a GUI desktop.

PYTHON

Python is an interpreter, object-oriented, high-level programming language with dynamic semantics. It's a coding tool which allows you to write, test and debug your code in an easier way. Python programming language is used for developing both desktop and web application. Python encourages program modularity and code reuse.
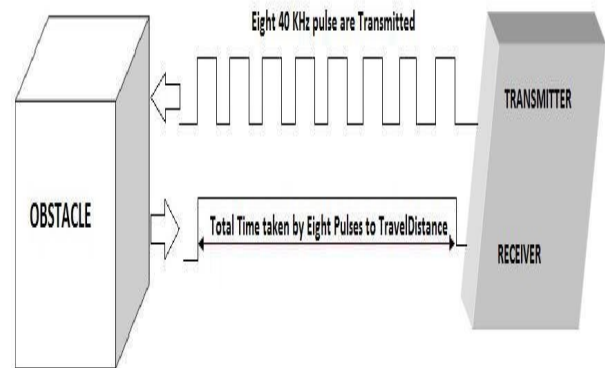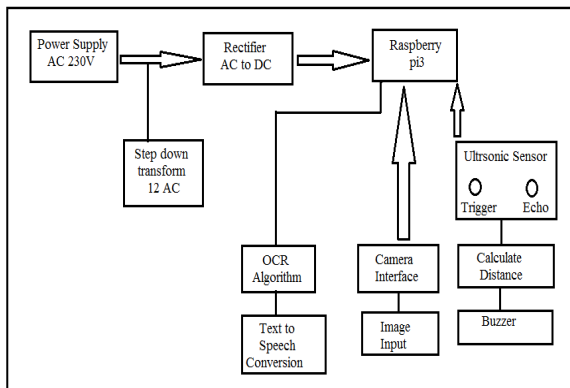
VIII. OCR TECHNOLOGY

IMAGE CAPTURING AND PRE-PROCESSING

The video is captured by using web-cam and the frames from the video is segregated and undergone to the pre-processing. First, get the objects interest is extracted from the camera image and it converted into gray image. Use haar cascade classifier for recognizing the character from the object. The work with a cascade classifier includes two major stages: training and detection. . To extract the hand-held object of interest from other objects in the camera view, ask users to shake the hand-held objects containing the text they wish to identify and then employ a motion-based method to localize objects from cluttered background.

AUTOMATIC TEXT EXTRACTION

In order to handle complex backgrounds, two novel feature maps to extracts text features based on stroke orientations and edge distributions, respectively. Here, stroke is defined as a uniform region with bounded width and significant extent. These feature maps are combined to build an Ad boost based text classifier. The extraction information from audio and image source restricted to information execution from text. The actual transduction of audio and image data into text is the processing of OCR output.

IX. SYSTEM ARCHITECTURE

POWER SUPPLY

Power Supply Unitis the device that supplies electrical energy to the output loads. It gives a well-regulated power supply of +5v

TRANSFORMER

A transformer is a static device that transforms electric power from one circuit to electric power of the same frequency in another circuit. The voltage can be raised or decreased in a circuit, but with proportional variations in the current ratings. In this system it is used to step down 230V AC to 9V AC supply and provides isolation between power grids and circuit.

RECTIFIER:

A rectifier is a device that converts alternating current (AC), which periodically reverses direction, to direct current (DC), which flows in only one direction.
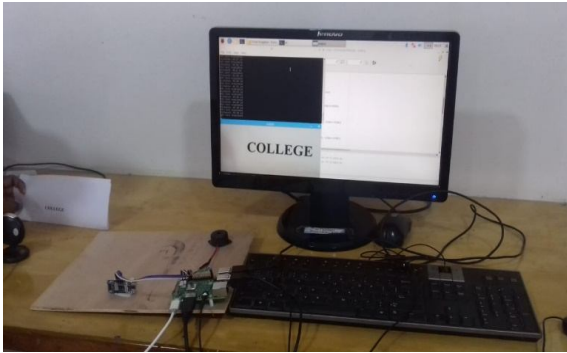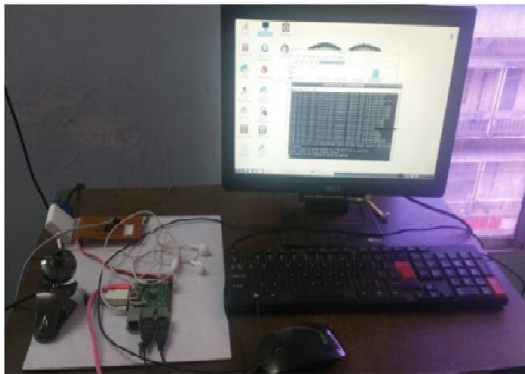
RASPBERRY PI3:

In proposed prototype system, we have used a raspberry pi advanced board model 3 which acts a mini computer. Built with ARM 11, 64bit processor, the new generation Raspberry Pi 3 is advanced and more powerful than its predecessors with built-in wireless and Bluetooth connectivity.

ULTRASONIC SENSOR:

Ultrasonic sensor has been interfaced with the board transmits and receives the pulse signal according to the duty cycle. Distance is measured using the duty cycle.

DISTANCE CALCULATION:

To find the distance between the obstacle and the person, Distance = speed * time

$$OD = \{[Speed\ of\ Sound * Time\ Taken] / 2\}$$

Where, OD*: Distance between an obstacle and the person in meters.

CAMERA INTERFACE:

The camera interface defines an interface between peripheral device (camera) and the processor.

TEXT RECOGNITION AND AUDIO OUTPUT

Text recognition is performed by off-the-shelf OCR prior to output of informative words from the text regions. These text regions are compared with the library and they are sent as audio signals in the audio jack and amplified and delivered as output to the blind users.

X. OUTPUT

STEP1:

STEP2:

STEP3:

XI. CONCLUSION

In this analysis, we've got represented a epitome system to scan written text and handheld objects for helping the blind individuals. To extract text regions from advanced backgrounds, we've got projected a completely unique text localization formula supported models of stroke orientation and edge distributions. The corresponding feature maps estimate the worldwide structural feature of text at each component. Block patterns project the projected feature maps of a picture patch into a feature vector. Adjacent character grouping is performed to calculate candidates of text patches ready for text classification.

REFERENCE

[1] E. Cardillo, V. Di Mattia, G. Manfredi, P. Russo, A. De Leo, A. Caddemi, G. Cerri "An Electromagnetic Sensor Prototype to Assist Visually Impaired and Blind People in Autonomous Walking",IEEEJournal,2017.

[2] Mrs.Shilpa Reddy K, Mounika S.k,Pooja K , Sahana N ,"Text to Speech for the Visually Impaired",IRJCSJournal,2017.

[3] Mallapa D.Gurav, Shruti S. Salimath, Shruti B. Hatti, Vijayalaxmi I. Byakod, ShivaleelaKanade "A Reading aid for the Blind People using OCR and OpenCV", IJSRET Journal 2017.

[4] Suchita Wankhade, MrunaliBichukale, Shruti Desai, Shraddha Kamthe, Archana Borate "Smart Stick for Blind People with Live Video Feed", IRJETJournal, 2017.

[5] Miss. Kirti P. Bhure, Mrs. J. D. Dhande"Object Detection Methodologies for Blind People" IJRIT Journal 2017.