A Novel Approach for Visually Impaired People Using Raspberry Pi

KalpanaDevi.K [1], Pavithra.P [2], Glothungan.M [3], Anusha.P.M[4]
[1, 2, 3] Dept of Electronics & Communication Engineering
[4]Assistant Professor, Dept of Electronics & Communication Engineering
[1, 2, 3, 4] Parisutham Institute of Technology and Science, Thanjavur-613 006,
Thanjavur, Tamilnadu, India.

Abstract- *Visual impairment is one of the biggest limitation for humanity. The device we have proposed aims to help people with visual impairment which converts an image's text to speech. It is implemented using a Raspberry Pi and a USB camera. To capture the image of the object we propose a camera module. These devices will capture the image with the help of raspberry pi and camera modules associated with it. In the extracted region of interest, text finding and recognition are conducted to acquire text information. The captured image undergoes a series of image pre-processing steps to locate only that part of the image that contains the text and removes the background. To automatically understand the text from the image we propose a tesseract Optical Character Recognition (OCR) module which will understand the image. Text obtained from the image is then localized and converted to text format. The recognized text codes are output to users in speech in the form of headphone/speakers with the help of ESPEAK software. USB camera will recognize the human faces which have been stored in the database with the help of haarcascade. With the help of GPS the user location can be identified.*

Keywords- Raspberry pi, Haarcascade, GPS, Tesseract OCR, ESPEAK

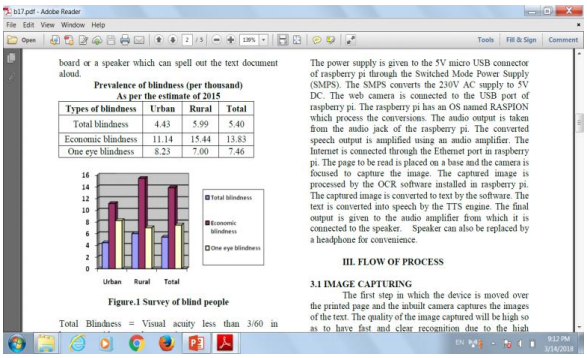
I. INTRODUCTION

In this chapter, we have to discuss about the introduction and problem faced by visually impaired peoples in day-to-day life, regarding this we have derived some objectives and studied some emerging technology, which we are using in our project. And also we got some solution for that problem. We also discussed the organization of this project in this chapter.

II.PROBELM FACED

Computing in crowded environment is a challenge for visually impaired people. Visually impaired people are at disadvantage because they do not have access to any contextual and spatial information around them. According to a survey, as of 2010 there were more that 285 million visually impaired people worldwide, out of which 39 million were blind. So they suffer a lot to move from one place to another place without others guidance. In today's fast paced world, the daily lives of people has been affected by the aid and support offered by technology. People, who are differently abled, now have the option of many devices, which can help them in their day-to-day activities. A lot of devices have been created in this field however, most of them are either not in use or requires a lot of training. A visually impaired individual faces two major challenges in his daily routine. One is being unaware of the path he needs to take to reach from say position A to position B. Another challenge is avoiding obstacles. Unlike any other physically challenged people the blind and visually impaired people are the people facing lots of difficulties in their daily life.

Fig 1: a)Dependent,b)White Cane, c) Accident occurs


board or a speaker which can spell out the text document aloud.

Prevalence of blindness (per thousand)
As per the estimate of 2015

Types of blindness	Urban	Rural	Total
Total blindness	4.43	5.99	5.40
Economic blindness	11.14	15.44	13.83
One eye blindness	8.23	7.00	7.46

Urban Rural Total

Total blindness
Economic blindness
One eye blindness

Figure.1 Survey of blind people

Total Blindness = Visual acuity less than 3/60 in

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. The web camera is connected to the USB port of raspberry pi. The raspberry pi has an OS named RASPION 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. The Internet is connected through the Ethernet port in raspberry pi. The page 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 for convenience.

III. FLOW OF PROCESS

3.1 IMAGE CAPTURING

The first step in which the device is moved over the printed page and the inbuilt camera captures the images of the text. The quality of the image captured will be high so as to have fast and clear recognition due to the high

Table 1: Survey of Blind People

III.OBJECTIVES

- To detect obstacles and provide alternative ways
- To indicate the user through speech (headphone)
- To enhance the mobility of the blind people

- To protect blind people from possible hazards

IV. BACKGROUND STUDY

A. Raspberry Pi

The raspberry Pi is a small, low cost CPU which can be used with a monitor, keyboard and mouse to become an efficient, full-fledged computer. The reason we chose Raspberry Pi micro-computer for our project is that, firstly, it is an easily available, low-cost device. Raspberry Pi uses software which are either free or open source, which also makes it cost-effective. The Raspberry Pi uses an SD card for storage and its small size also gives us the advantages of portability. As a part of the software development, the Open CV (Open source Computer Vision) libraries are utilized for image processing. Each function and data structure was designed with the Image Processing coder in mind.

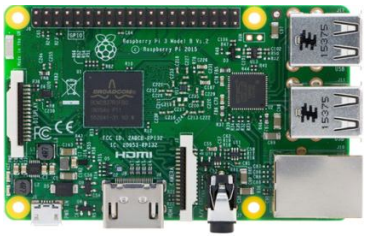


Fig 2: Raspberry pi

B. Tesseract OCR

Tesseract is a free software optical character recognition engine for various operating systems. Tesseract is considered as one of the most accurate free software OCR engines currently available. It is available for Linux, Windows and Mac OS. An image with the text is given as input to the Tesseract engine that is command based tool. Then it is processed by Tesseract command. Tesseract command takes two arguments: First argument is image file name that contains text and second argument is output text file in which, extracted text is stored. The output file extension is given as .txt by Tesseract, so no need to specify the file extension while specifying the output file name as a second argument in Tesseract command. After processing is completed, the content of the output is present in .txt file. In simple images with or without color (gray scale), Tesseract provides results with 100% accuracy. But in the case of some complex images Tesseract provides better accuracy results if the images are in the gray scale mode as compared to color images. Although Tesseract is command-based tool but as it is open source and it is available in the form of Dynamic Link Library, it can be easily made available in graphics mode.

C. ESPEAK

The image capturing element gathers the images which contains text in the form of video or image. In our prototype system images are captured by using a pi camera. Using the camera, image of the object from the cluttered backgrounds or other surroundings are extracted. Text localization algorithm is used for acquiring the text from images. Text recognition is implemented to convert the image based text to readable codes. In the prototype developed a laptop is used for data processing. The audio output is delivered to the blind users through e-speak engine and the audio is presented through headphone or speakers.

V. EXISTING SYSTEM

The image processing method includes capturing of static image with the help of camera. The camera works as an eye for the raspberry-pi. The camera can be connected to the raspberry-pi with the help of LAN Cable, Wi-Fi or Bluetooth. We have primarily used a raspberry pi camera to capture the image. After the successful connection the image is captured with the help of tesseract OCR software. We are using tesseract OCR which is raspberry pi compatible and can understand primarily English language. The teserract OCR is library provided by the Linux community which is open source. The Tesseract OCR is command line OCR which captures the image on the press of button i.e. on execution of instruction. The image can be saved in .jpeg or .png format. The captured image is converted to the .txt format and saved with the same name as an image and the message will be send through headphone.

Disadvantages

It cannot extract information from the complex backgrounds. It will not detect the objects accurately and also it will not recognize the faces.

VI. PROPOSED WORK & DESIGN FLOW

When we enter in ATM (All Time Medicine) dispenser system first we give a fingerprint for authentication and we have to enter the disease symptoms. This information provides to the microcontroller. The proposed system declared a microcontroller and motor based system to dispense the medicines when accessed by the user through an input event, the data pertaining to the medicine storage can be ascertained from the remote area. Basic human parameters like Temperature, Heartbeat rate, blood pressure can also be tested through this machine and the specified medicine will be dispensed based on the patient condition.The system is loaded with the medicines and will be dispensed upon the request of

the users, the total system functioning will be handled by the microcontroller interfaced with the sensors which they detect the ailment according to an information the motors to dispense the medicine. The information related to the storage data of the medicines, and also the user interaction can be done through PC. System is providing certain privileges to the patient to choose the required medicine. Information about the patient and the usage of the medicines by the patient will be tracked.

Advantages

- Prevent medication errors.
- Make the patient and the caregiver feel better.
- Automatic process.
- 24 hours system.
- Increase the human lifetime.
- Decrease the accidents.

A.BLOCK DIAGRAM

Block diagram consists of Raspberry pi, USB camera, Battery bank, Memory storage, Monitor and Headphone. To overcome the problem of existing system, we have used USB camera coded with python. This camera gives us more clarity and text can be extracted easily, so that information in the image was given to the user correctly without any loss.

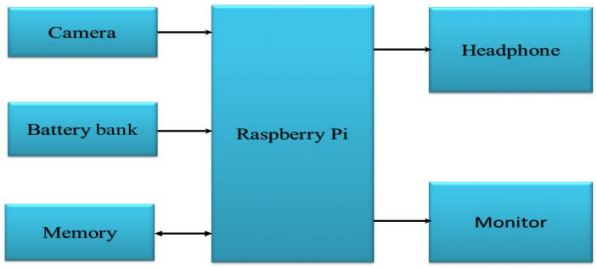


Fig 3: Block Diagram

B. PROCESS DIAGRAM

Our process have the following stages,

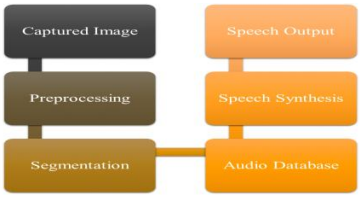


Fig 4: Process Diagram

Capturing Image:

First the image was captured through the camera and stored in the memory. We have also added face recognition along with this to identify the person.

Pre-processing:

Pre-processing is a common name for operations with images at the lowest level of abstraction of both input and output are intensity images. The aim of pre-processing is an improvement of the image data that suppresses unwanted distortions or enhances some image features important for further processing.

Image Processing Method:

The image processing is done with the optical character recognition method. The optical character recognition is a method that captures or scans the images and has an ability to convert the image into readable or text format which can be processed further. The image captured with OCR can be of any resolution but wehave preferred 720p resolution for the images of .jpg or .jpeg format.

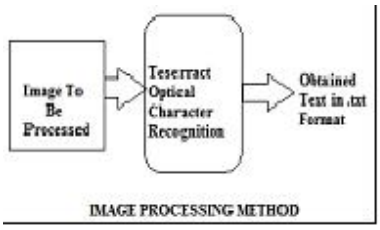


Fig 5:Image Processing Method

The image processing method includes capturing of static image with the help of camera. The camera works as an eye for the raspberry-pi. The camera can be connected to the raspberry-pi with the help of LAN Cable, Wi-Fi or Bluetooth. We have primarily used a raspberry pi camera to capture the image. After the successful connection the image is captured with the help of tesseract OCR software. There are various OCR tools such as gocr, ocrad, ocrfeeder, ocropus, teserract-ocr and cuneiform. We are using tesseract OCR which is raspberry pi compatible and can understand primarily English language. The teserract-ocr is library provided by the Linux community which is open source. The Tesseract-ocr is command line OCR which captures the image on the press of button i.e. on execution of instruction. The image can be saved in .jpeg or .png format. The captured image is converted to the .txt format and saved with the same name as an image.

Voice Processing Method:

In this method our main aim to convert obtained text to speech with the help of coding in raspberry pi. The text to speech module is installed in raspberry pi. Various methods and synthesizers can be used in raspberry pi to convert obtained text to speech. Those modules are pyttsx text to speech, NS Speech Synthesizer, SAPI 5 for Windows XP, Vista and Windows 7, ESPEAK on any platform that can host share library such as Ubuntu and Fedora Linux and Google Text to Speech. We are using pyttsx or festival for converting obtained text to the speech.

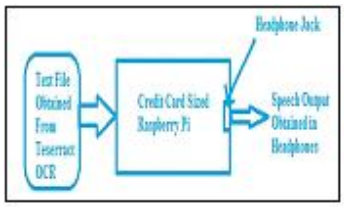
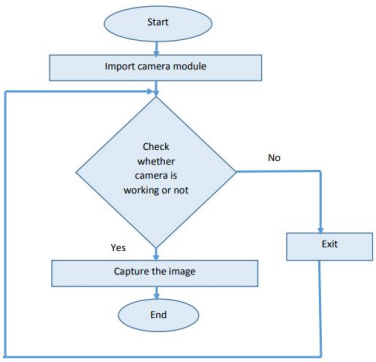


Fig 6: Voice Processing Method

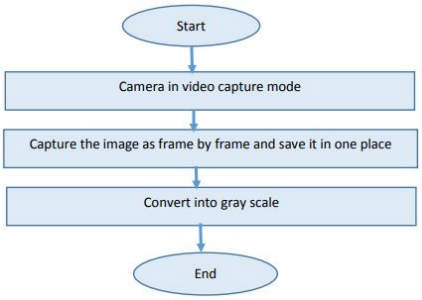
We will provide obtained text file to the text to speech synthesizer with the help of text file input command. We have installed pyttsx in raspberry pi with the help of command sudo apt install pyttsx and festival with command sudo apt-get install festival. These commands will install pyttsx for text to speech conversion in raspberry pi and festival for the same. We can also use Google Text to Speech or pico software for text to speech conversion.

C.FLOW CHART

Process 1:

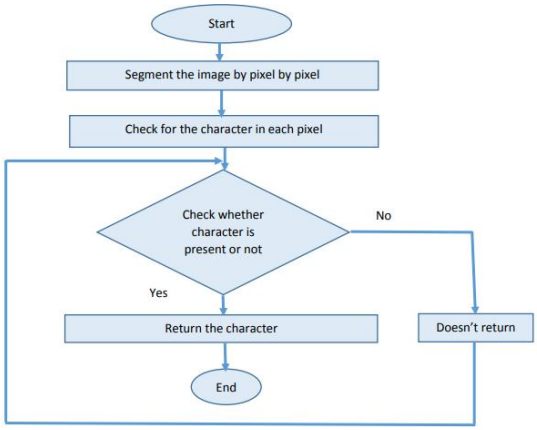


Process 2:

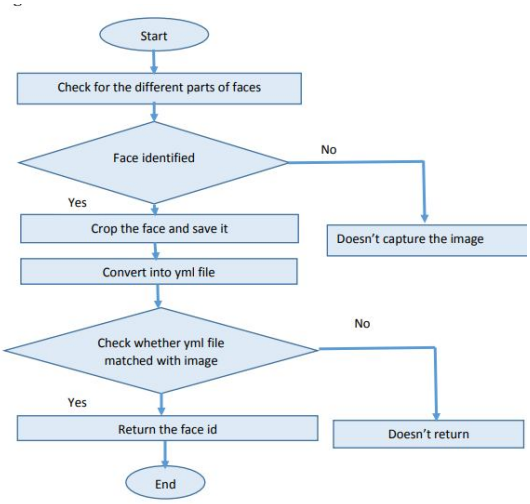


Process 3:

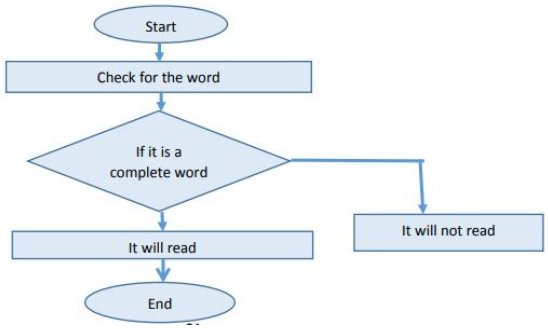
For character



For face recognition



Process 4:



VII. DESIGN IMPLEMENTATION

A.DESIGN AND IMPLEMENTATION

In this chapter, we have to design and implement the visually impaired navigation system using Raspberry Pi.

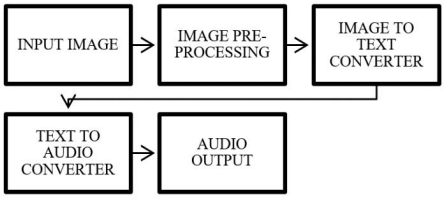


Fig 7: Process involved in image to text conversion

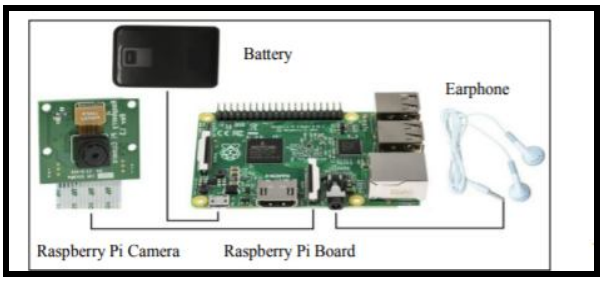


Fig 8: Blind Guidance Module

Figure 5.2 shows the blind guidance module implemented from our project.The project module consists of Raspberry Pi interconnected with power bank, USB camera and head phone.

Face Detection:

The OpenCV face detector is applied to find faces in the incoming images. It is very robust in finding frontal faces, as it is trained with a huge number of training samples. The OpenCV detector can detect the faces (frontal faces) in the poses of looking at the camera using the algorithm based on HaarCascade features. Face detector which is trained by Haarcascade Classifier are applied to find faces in the input image frame. Detected face image is captured, automatically cropped and resized to 222 x 222 pixels. Captured images are used for training database.

Training:

In training stage, for each person enrolled in the database, five face images are collected. All these face images undergoes image transformation using various image filtering techniques such as changing brightness-contrast, titling of an image. After image transformation the number of images increases from 5 to 25 face images. Trained images are stored in the database automatically, path of each image is updated in the file called Comma Separated Value (CSV) file and the number of the people in the database is updated in the file called category file. At each training stage, CSV and category file is updated. When more people have to be enrolled in the database, the incoming face images are transformed using the above mentioned image filtering techniques and then the transformed face images are stored in database.

i) Brightness-Contrast:

An image as in Figure 5.3 shows the proper brightness and contrast for easy viewing. Brightness refers to the overall lightness or darkness of the image. Contrast is the difference in brightness between objects or regions.

Fig 9: Brightness-contrast

ii) Tilt:

A detected face image is tilted to different angles. Figure 5.4 shows the captured image is tilted at various angles.

Fig 10: Image tilted to right and left

B.Hardware Specification:

SoC	Broadcom BCM2835
CPU	700MHz ARM11 ARM1176JZF-S core. **Can be overclocked safely.**
GPU	Broadcom VideoCore IV, OpenGL ES 2.0, OpenVG 1080p30 H.264 high-profile encode/decode
Memory (SDRAM)	512Mib
USB 2.0 ports	2 (via integrated USB hub)
Video outputs	Composite RCA, HDMI (cannot be used simultaneously)
Video input	CSI
Audio outputs	TRS connector / 3.5mm jack, HDMI
Onboard storage	SD / MMC / SDIO card slot
Onboard network	10/100 wired Ethernet RJ45
Low-level peripherals	26 GPIO pins, SPI, I²C, I²S, UART
Power ratings / source	700 mA, (3.5 W) / 5V (DC) via Micro USB type B or GPIO
Size / weight	85.0 x 56.0 x 17 [mm] / 40g

C.Software Description

- Raspbian Jessie
- Python
- OPEN CV
- Haarcascade
- Tesseract OCR
- ESPEAK

IX. RESULTS AND DISCUSSION

In this chapter we are going to discuss about the result we have obtained from our project.

Face Recognition:

Face detection automatically detects a face from a complex background to which the face recognition algorithm can be applied. By the haar cascade process the face can be easily identified. Figure 6.1 shows how the face recognition is done by capturing the image.

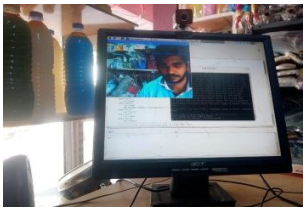


Fig 11: Face Recognition

The image is converted to gray scale as many OpenCV functions require the input parameter as a gray scale image. Noise removal is done using bilateral filter. Edge detection is performed on the gray scale image for better detection of the contours. The wraping and cropping of the image are performed according to the contours. This enables us to detect and extract only that region which contains text and removes the unwanted background. In the end, Thresholding is done so that the image looks like a scanned document. This is done to allow the OCR to efficiently convert the image to text. Then by using E-speak algorithm the message will be given to the headphone/speaker.

Step 1: Capturing the image

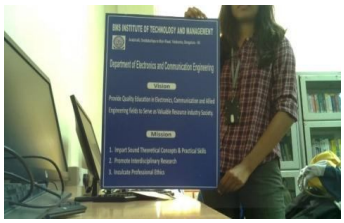


Fig 12: Original image captured from the camera

Step 2: Converting the image into Gray scale

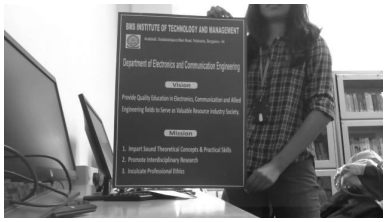


Fig 13: Image converted to gray scale

Step 3: Performing the edge detection

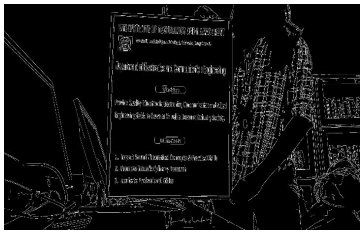


Fig 14: Edge detection

Step 4: Performing the Contour detection

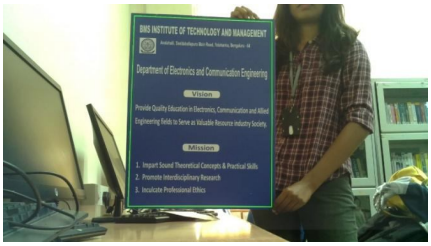


Fig 15: Contour detection

Step 5: Wrapping and cropping the image

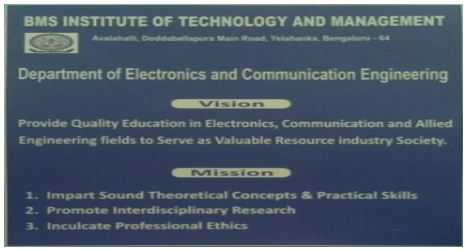


Fig 16: Wrapped and cropped image

Step 6: Sharpening the image

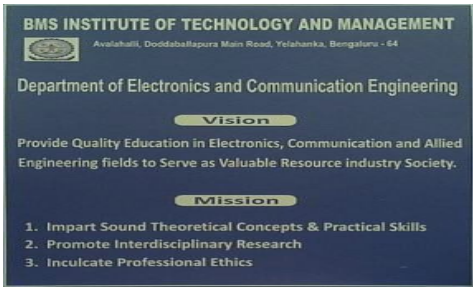


Fig 17: Sharpening the image

Step 7: Converting into gray scale before thresholding

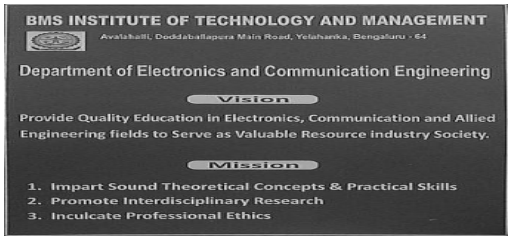
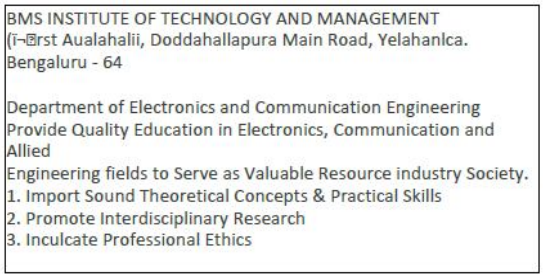


Fig 18: Convert to grayscale before thresholding

Step 8: Performs thresholding

Fig 19: Thresholding



X .CONCLUSION AND FUTURE WORK

The system proposed is of great advantage to the people who are blind and visually impaired. But with the more sufficient investment in technology and human resource the system can be enhanced with artificial eyes called Bionic eyes. Several researches are being conducted by various scientists across the globe to have this system implemented in real time with a much more additional features. These systems will be implement in no longer time.

Advantages

- Person identification.
- Cost efficient.
- Easy to carry
- Avoid accident
- Secured,

Applications

- Using our proposed model the number plate recognition can be done in vehicles.
- The driverless car can be implemented by using this process.

Future scope

In future there is a scope for it can also extend for the security purpose in order to identify the thieves and to enhance the coverage to long distance capturing. Night vision cameras

can be used for 24 hours usage. Further we can also speed up the process through more advanced processors and technique.

REFERENCE

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

[2] K NirmalaKumari, 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´elineThillou, Bernard Gosselin. From Picture to Speech: An Innovative Application for Embedded Environment. FacultéPolytechnique de Mons, Laboratoire de Théorie des Circuits etTraitement 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.