Optical Character Recognition Using Visual Impaired People Shopping Trolley Technology

A Sultan Saleem[1] S.Sasmitha[2] C.Varsha[3]
[1]Associate Professor [2,3]Student
[1,2,3]Department of Computer Science and Engineering
[1,2,3]Dhanalakshmi College of Engineering, Chennai, Tamil Nadu, India

Abstract— Shopping is one of the most essential activities that people have to perform in their day to day life. But there are some people who face many difficulties in shopping such as VIP (visually impaired people). They have to depend on others to shop in unfamiliar places. In order to address these difficulties we provide a device that will help the visually impaired people to shop independently by using different technologies. The proposed system consists of different modules such as product recognition and section information which will process on raspberry pi3 model. Product recognition includes Optical Character Recognition and Text to Speech software which will help user to identify the product. Section information will help user to get the section information by using Radio Frequency Identification (RFID) technology. Ultrasonic sensor will detect the obstacle and alert the user using beep sound.

Keywords: OCR Algorithm -Optical Character Recognition, UV Sensor – Ultra Violet, PCB- Printed Circuit Board

I. INTRODUCTION

In this word visually disabled people are around 285 million and 39 million of them are blind. According to World Health Organization 90% of blind people live in developing country. Blind support devices are expensive and many blind people cannot afford it. This project aims to develop a cheap, blind guidance system for developing countries. The goal of this paper is to develop a low-cost intelligent system for guiding individuals who are blind or partially sighted by providing information about the environmental scenario of static and dynamic objects around them. The main functions of this system are path indication and environment recognition. The system is used to recognize the destination place by using the motors and instruct with the voice command. Whenever the product is shown to a camera it scans the text and converts into audio form. Ultrasonic sensors will provide information about obstacles if it is in the range. Ultrasonic sensors are connected with microcontroller circuit and a microcontroller timer module is used to determine the sensors output and then sent to the motors.

II. LITERATURE SURVEY

E. Cardillo, V. Di Mattia, G. Manfredi, P. Russo, A. De Leo, A. Caddemi, G. Cerri.(2018) "An Electromagnetic Sensor Prototype to Assist Visually Impaired and Blind People in Autonomous Walking."

To assist the autonomous walking of visually impaired and blind users. Microwave radar on the traditional white cane concept.

Mrs.Shilpa Reddy K, Mounika S.k,Pooja K,Sahana N . "Text to Speech for the Visually Impaired". (2017)The main goal is to progress such a system that will peruse the texts from composite backgrounds successfully.

OTSU algorithm and MODI algorithm.

Suchita Wankhade, Mrunali Bichukale, Shruti Desai, Shraddha Kamthe, Archana ." Borate Smart Stick for Blind People with Live Video Feed" (2017). To help the visually impaired people to move in their surrounding environment with mobility and confidence.

GPS (graphic positioning system) Live feed concept.

Miss. Kirti P. Bhure, Mrs. J. D. Dhande ."Object Detection Methodologies for Blind People".(2017) To introduce a proposed system that restores a central function of the visual . system which is the identification of surrounding objects .

Local features extraction concept using SIFT algorithm.

R. K. Katzschmann, B. Araki, and D. Rus ." Safe local navigation for visually impaired users with a time-of-flight and haptic feedback device".(2018) To provide a contactless, intuitive, hands-free, and discreet wearable device that allows visually impaired users to detect low- and high-hanging obstacles, as well as physical boundaries in their immediate environment.

ALVU concept which consists of sensor belt and a haptic strap.

III. PROPOSED METHODOLOGY

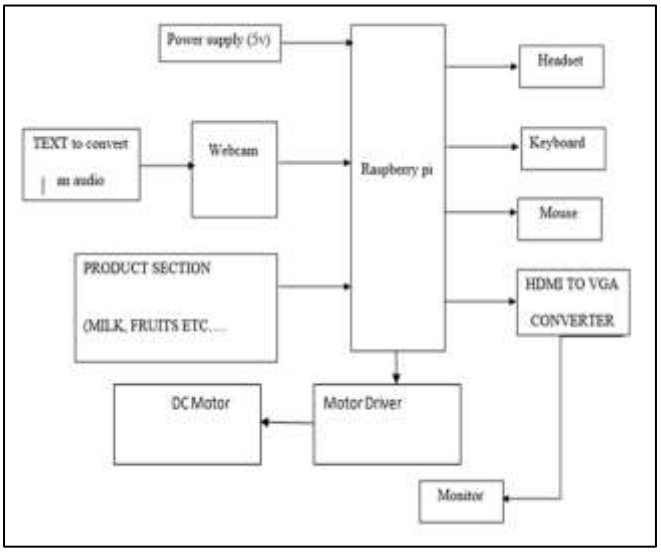
Methodology is a model to explain the methods or techniques used to design, develop or plan a project. This chapter explains about the software and hardware that will be used for developing this project further. The results are going to be analyzed to achieve the objective of this project.

A. System Overview

The proposed system uses a Raspberry Pi board, an ultrasonic sensor and a webcam to recognize the text in the scene. In this, the webcam is focused on the scene. A video streaming is obtained, from which the images are captured frame by frame. The images are refined in order to eliminate any noise that is present in it. A feature called segmentation is used in order to separate each character from other in the text. Graphical details such as icons or logos, if any, are eliminated. Each obtained character is compared with the datasets that are created as a part of the Tesseract library. The Tesseract OCR is the most efficient algorithm available that checks for the obtained character in ten dimensions. Once, the character is recognized, it must be made available as an audio output. For this, we use a software called festival. The festival is used to provide the audio output for the recognized character. Apart from these features, an extra feature is added, that enables the blind to know the type of object that he/she interacts with.(a menu, newspaper and the like). An ultrasonic sensor is included as a part of the project, that makes the project obtain characters only within a particular distance.

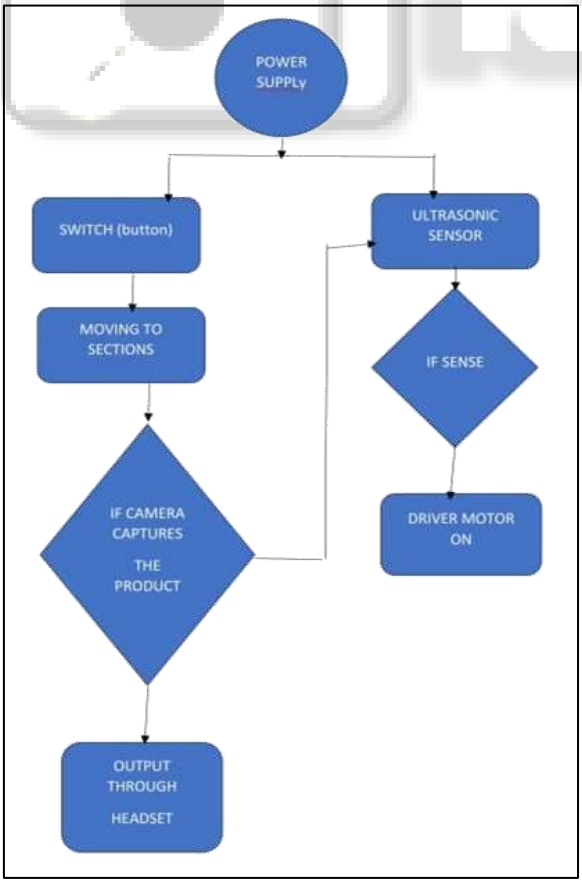
1) Advantages of Proposed System
- Flexible for blind people.
- To convert printed books to digital text.
- Efficient and Effective to handle.
- Text to audio output.

B. Block Diagram



The modules that are based on the work done in the proposed system are as follows.
- Image Capture and Pre-processing.
- Automatic Text Extraction.
- Text Recognition and Audio Output.

C. Flowchart



D. Module Description

1) Image Capturing and Processing

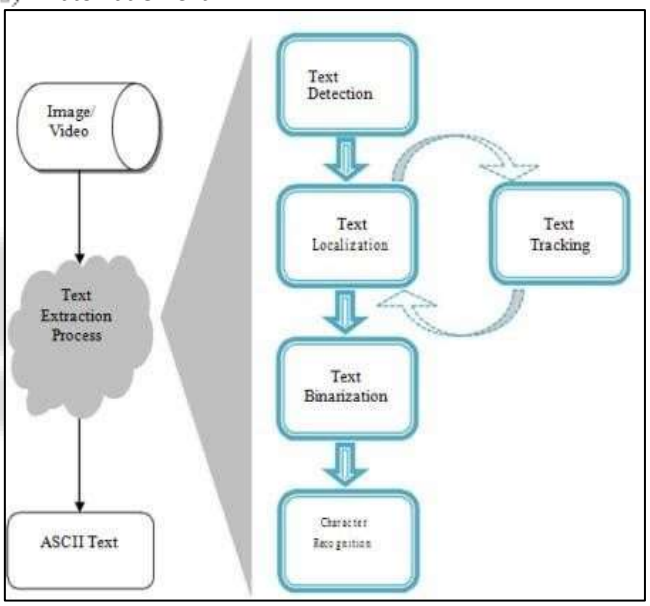
a) Capturing the Image

An active Webcam can capture an image up to 30 frames per second. Capturing of the image can be accomplished from any video device including USB cameras. The capture card is connected with analog cameras and TV boards. When a motion is detected in the motioned area by the program, then the alarm can sound, and it starts broadcasting or record a video. There is an additional feature in the program

b) Processing the Image

The noise introduced during capturing or due to poor quality page has to be cleared before further processing. This can be achieved by processing the image. The pixel density and quality is also adjusted and corrected by image processing. Image processing contains the following steps.
- Filtering
- Edge Detection
- Background Separation

2) Automatic Text

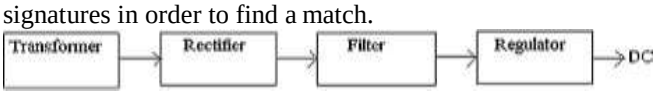


a) Extraction

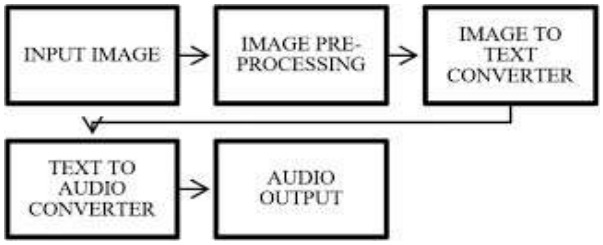
The aim of Optical Character Recognition (OCR) is to classify optical patterns (often contained in a digital image) corresponding to alphanumeric or other characters. The process of OCR involves several steps including segmentation, feature extraction, and classification. In principle, any standard OCR software can now be used to recognize the text in the segmented frames. However, a hard look at the properties of the candidate character regions in the segmented frames or image reveals that most OCR software packages will have significant difficulty to recognize the text. Document images are different from natural images because they contain mainly text with a few graphics and images. Due to the very low-resolution of images of those captured using handheld devices, it is hard to extract the complete layout structure (logical or physical) of the documents and even worse to apply standard OCR systems. For this reason, a shallow representation of the low- resolution captured document images is proposed. In case of original electronic documents in the repository, the extraction of the same signature is straightforward; the PDF or PowerPoint form of

the original electronic documents is converted into a relatively high-resolution image (TIFF, JPEG, etc.) on which the signature is computed.

Finally, the captured document's signature is compared with all the original electronic documents' signatures in order to find a match.



b) Architecture of Text

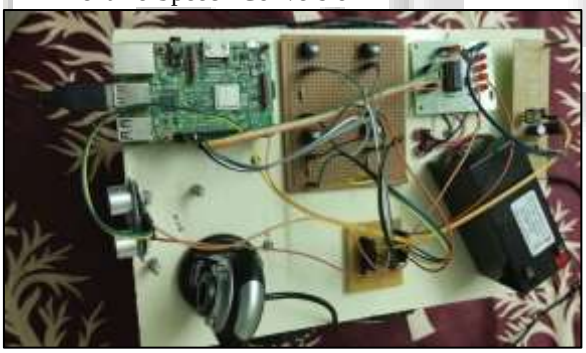


Extraction Process

3) Text Recognition and Audio Output

Text recognition is performed by off-the-shelf OCR prior to output of informative words from the localized text regions. A text region labels the minimum rectangular area for the accommodation of characters inside it. The border of the text region contacts the edge boundary of the text characters. OCR generates better performance text regions which are assigned proper margin areas and binaries to segment text characters. The recognized text codes are recorded in script files. Off the shelf OCR is used to perform word recognition on the localized text regions and transform into audio output for blind users.

a) Text To Speech Conversion



Demonstration of Project

IV. COMPONENTS DETAILS

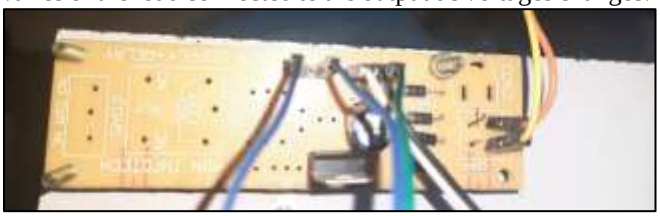
A. Raspberry pi

The Raspberry Pi board contains a processor and graphics chip, program memory (RAM) and various interfaces and connectors for external devices. Some of these devices are essential, others are optional. RPi operates in the same way as a standard PC, requiring a keyboard for command entry, a display unit and a power supply. It also requires 'mass-storage', but a hard disk drive of the type found in a typical PC is not really in keeping with the miniature size of RPi. Instead we will use an SD Flash memory card normally used in digital cameras, configured in such a way to 'look like' a hard drive to RPi's processor. RPi will 'boot' (load the Operating System into RAM) from this card in the same way as a PC 'boots up' into Windows from its hard disk.

RASPBERRY PI3

B. Power Supply (Battery Source)

All electronic circuits work only in low DC voltage. We need a power supply unit to provide the appropriate voltage supply for their proper functioning. This unit consists of transformer, rectifier, filter & regulator. AC voltage of typically 230 volts rms is connected to a transformer voltage which is stepped down to the level of the desired ac voltage. A diode rectifier that provides the full wave rectified voltage that is initially filtered by a simple capacitor filter to produce a dc voltage. This resulting dc voltage usually has some ripple or ac voltage variation. A regulator circuit can use this dc input to provide dc voltage that not only has much less ripple voltage but also remains the same dc value even when the dc voltage varies or the load connected to the output dc voltages changes.



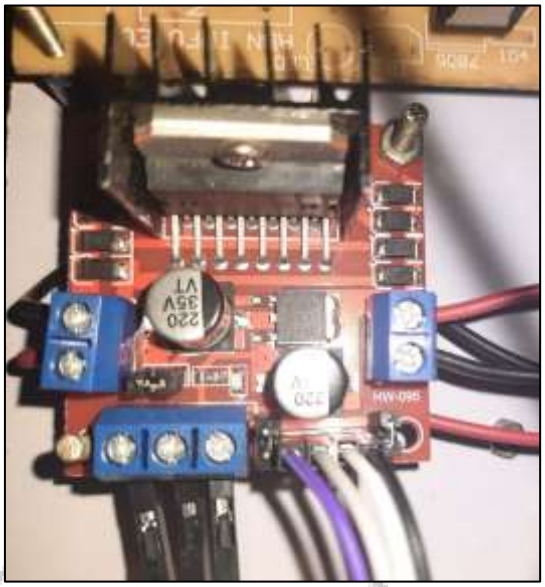
General Block of Power Supply Unit

C. D.C. Motors

An electric motor converts electrical energy into mechanical energy. Most electric motors operate through interacting magnetic fields and current- carrying conductors to generate force, although electrostatic motors use electrostatic forces. Electric motors are found in applications as diverse as industrial fans, blowers and pumps, machine tools, household appliances, power tools, and disk drives. They may be powered by direct (e.g., a battery powered portable device or motor vehicle) or by alternating current from a central electrical distribution grid.

Brushless DC motors use a rotating permanent magnet in the rotor, and stationary electrical magnets on the motor housing. A motor controller converts DC to AC. This

design is simpler than that of brushed motors because it eliminates the complication of transferring power from outside the motor to the spinning rotor. Advantages of brushless motors include long life span, little or no maintenance, and high efficiency. Disadvantages include high initial cost, and more complicated motor speed controllers.



D. Camera

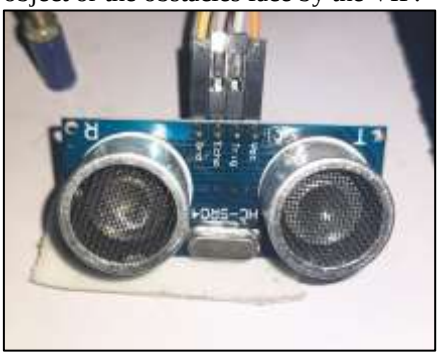
Logitech HD Webcam comes fitted with auto focus wide-angle lens that has smaller focal length to capture more of your subjects.

This type of lens goes well with indoor photography so that you can enjoy amazing images on your notebook or desktop at home or office.

Moreover, with its up to 720p HD video-recording capacity (special PC configuration needed), you view high-quality HD picture on your monitor that is clear, sharp and detailed. Colours also look vivid and beautiful. Further, Logitech Fluid Crystal Technology enhances the video quality by making your chats' display fluid and flawless. Automatic Light Correction feature rectifies brightness-related errors to give you improved viewing experience.

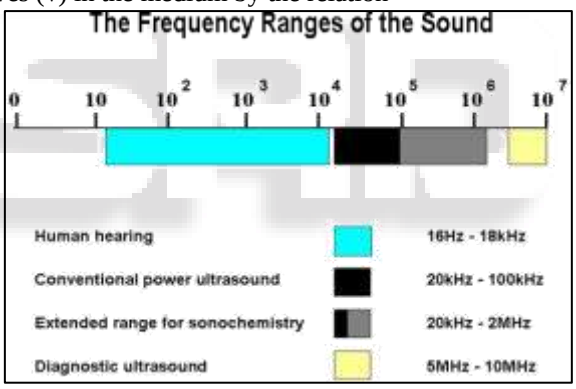
E. Ultrasonic sensor

Ultrasonic sensors measure distance by using ultrasonic wave and receives the wave reflected back from the target. It measures the time required during the emission and reception of the waves to calculate the distance. The device itself works as emitter and receptor. In the proposed system it is used to detect the object or the obstacles face by the VIP.



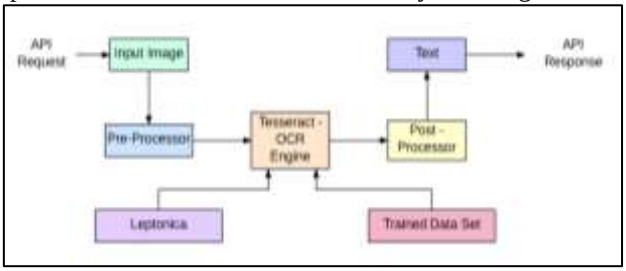
F. UV Sensor

When ultrasonic waves are incident on an object, diffused reflection of the energy takes place over a wide solid angle which might be as high as 180 degrees. Thus some fraction of the incident energy is reflected back to the transducer in the form of echoes and is detected. The distance to the object (L) can then be calculated through the speed of ultrasonic waves (v) in the medium by the relation



V. IMPLEMENTATION

First and foremost process is to capture the image frame by frame and convert into text format. For capturing the image we are using raspberry pi integrated with python and through espeak it is converted into audio form by OCR algorithm




```
import RPi.GPIO as GPIO
import time
import cv2
import numpy as np
import cv2 as cv
GPIO.setwarnings(False)
```



```
GPIO.setmode(GPIO.BOARD)
cap=cv2.VideoCapture(0)
if cap.isOpened()==False:
   cap=cv2.VideoCapture(1)
try:
   while(1):
         ret,img_frame =cap.read()
         cv2.imshow('output',img_frame)
        cv2.waitKey(1)
       time.sleep(1)
       k=cv2.waitKey(10)
      #checking for an esc key press
          if k==27:
             break
except Keyboard Interrupt:
   cap.release()
   cv2.destroyAllWindows()
```

the above code is used to capture the image frame by frame.

For source we need a 5V ground power supply. This trolley is also connected to the printed circuit board which is used to connect the electronic components.

We use buttons for trolley control. The trolley movement is controlled by buttons and the trolley is moved with Geared DC motors. The motors are based upon the essentials things we assigned four buttons in the trolley.

```
import time
import cv2
import numpy
from subprocess import Popen, PIPE
import RPi.GPIO as GPIO
import os
import serial
import numpy as np
import cv2 as cv



TRIG=11
ECHO=13

GPIO.setwarnings(False)
GPIO.setmode(GPIO.BOARD)

GPIO.setup(11,GPIO.OUT)     #Trigger pin
GPIO.setup(13,GPIO.IN)      #Echo Pin
GPIO.setup(37,GPIO.IN)      #switch 1
GPIO.setup(35,GPIO.IN)      #switch 2
GPIO.setup(33,GPIO.IN)      #switch 3
GPIO.setup(31,GPIO.IN)      #  switch 4
GPIO.setup(40,GPIO.OUT)     #Motor 1
GPIO.setup(38,GPIO.OUT)     #Motor 1
GPIO.setup(36,GPIO.OUT)     #Motor 2
GPIO.setup(32,GPIO.OUT)     #Motor 2

cap=cv2.VideoCapture(0)
if cap.isOpened()==False:
    cap=cv2.VideoCapture(1)
```

A button named as switch in the code .

Each button works with the individual function. The first button is used to move forward and turns right.

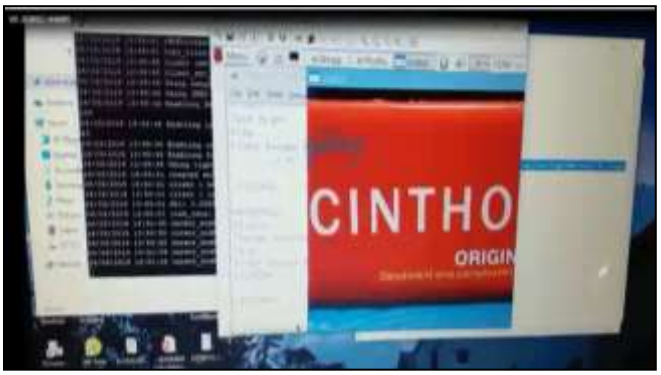
If we press the second button it moves forward and turns left and simultaneously it works.

The Trolley movement is by L293D driver which is connected with Raspberry pi.

It acts as a current amplifier which is used to drive the trolley based on inputs.

When input logic is 00 or 11, the trolley is stopped.

When the input logic is 01 or 10 the trolley moves in the forward or backward direction.



VI. RESULT

In the proposed System, the aim is to design and build a prototype. This has now been implemented. A Smart Shopping Trolley is used to capture the image and convert it to Audio to help the blind. Together the Hardware and Software are well integrated. The UV Sensor is able to detect the obstacles and indicate to the person with the distance. The Motor helps to move the Trolley.

VII. CONCLUSION AND FUTURE ENHANCEMENT

A. Conclusion

The proposed system will enable visually impaired people to shop independently in a supermarket. Being specific gadget for product identification, Section Information and obstacle detection can be designed. It makes the use of RFID and raspberry-pi technologies for providing the smart environment for the visually impaired. To solve the common aiming problem we have implemented motion-based method to detect the objects of interest. Text extractions are done using stroke orientation and distribution of edge pixels. The text characters are recognized using Optical Character Recognition, and the text codes are transformed as speech for blind persons.

B. Future Enhancement

Our future work will extend the text localization algorithm with more features and we will address the human interface issues associated with text reading by the blind user.

REFERENCE PAPERS

[1] Aladrn, G. Lpez-Nicols, L. Puig, and J. J. Guerrero, "Navigation assistance for the visually impaired using

RGB-D sensor with range expansion," IEEE Syst. J., vol. 10, no. 3, pp. 922–932, Sep. 2016.

[2] T. Amemiya and H. Sugiyama, "Orienting kinesthetically: A haptic handheld wayfinder for people with visual impairments," ACM Trans. Accessible Comput., vol. 3, no. 2, 2010.

[3] And, S. Baglio, V. Marletta, and A. Valastro, "A haptic solution to assist visually impaired in mobility tasks," IEEE Trans. Human-Mach. Syst., vol. 45, no. 5, pp. 641–646, Oct. 2015.

[4] S. Bhatlawande, M. Mahadevappa, J. Mukherjee, M. Biswas, D. Das, and S. Gupta, "Design, development, and clinical evaluation of the electronic mobility cane for vision rehabilitation," IEEE Trans. Neural Syst. Rehabil. Eng., vol. 22, no. 6, pp. 1148–1159, 2014.

[5] B. Blasch, R. G. Long, and N. Griffin-Shirley, "National evaluation of electronic travel aids for blind and visually impaired individuals: Implications for design," in Proc. RESNA 12th Annu. Conf., 1989, pp. 133– 134

