Face and Text Recognition for Visually Impaired Person Based On Raspberrypi

Varsha Patil[1] , Shivani Patel[2] , Kajal Patel[3] , Namrata Patil[4] , Gitanjali Chaudhari[5], Prof.V.T.Patil[6]

Abstract- *Independent travel is a well known challenge for blind or visually impaired persons and also the increasing availability of cost efficiency, high performance and portable digital imaging devices has created a tremendous opportunity for supplementing traditional scanning for document image acquisition. We propose a camera based visual assistance framework for text reading, motion of objects and the feelings of persons etc. It converts into a voice output to help blind peoples. Thus we propose this system from a single camera-captured image as well as more number of frames and highlight some sample applications under development and feasible ideas for future development. This paper presents the automatic document reader for visually impaired people, developed on Raspberry Pi. It uses the Optical character recognition technology for the identification of the printed characters using image sensing devices and computer programming. It converts images of typed, handwritten, or printed text into machine encoded text. In this research these images are converted into the audio output (Speech) through the use of OCR and Text-to-speech synthesis. The conversion of printed document into text files is done using Raspberry Pi which again uses Tesseract library and Python programming. The text files are processed by OpenCV library & python programming language and audio output is achieved.*

I. INTRODUCTION

To improve the ability of people who are blind or have significant visual impairments to independently access, understand, and explore unfamiliar indoor and outdoor environments, we propose a new framework using a single camera to detect and recognize the signs, text and obstacles to give audio as an output. we extract and recognize the text information associated with the detected objects. We first extract text regions from indoor signs with multiple colors. Then text character localization and layout analysis of text strings are applied to filter out background interference. The object type, orientation, and location can be displayed as speech for blind travelers. To improve the ability of people who are blind or have significant visual impairments to independently access, understand, and explore unfamiliar indoor environments, we propose a new framework using a single camera to detect and recognize the face, obstacles, signs incorporating text information associated with the detected object. In order to

discriminate similar objects in indoor environments, the text information associated with the detected objects is extracted. This paper is organized as follows: section II. existing methods, section III proposed methods, section IV materials and methods, followed by result in section V and conclusion in section VI. 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. 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 AGRICULTURAL AUTOMATION SYSTEM:

This part illustrates the present method of monitoring and controlling agricultural fields.

2.1 Camera-based analysis of text and documents:

Here they proposed a camera-based assistive framework to help blind persons to read text labels from cylinder objects in their daily life. First, the object is detected from the background or other surrounding objects in the camera view by shaking the object. Then we propose a mosaic model to unwrap the text label on the cylinder object surface and reconstruct the whole label for recognizing text information. This model can handle cylinder objects in any orientations and scales. The text information is then extracted from the unwrapped and flatted labels. The recognized text codes are then output to blind users in speech. Experimental results demonstrate the efficiency and effectiveness of the proposed framework from different cylinder objects with complex backgrounds.

2.2 Context-based Indoor Object Detection as an Aid to Blind Persons Accessing Unfamiliar Environments:

In order to find different rooms (i.e. an office, a lab, or a bathroom) and other building amenities (i.e. an exit or an elevator), they incorporated this door detection with text recognition. First they developed a robust and efficient algorithm to detect doors and elevators based on general geometric shape, by combining edges and corners. The algorithm is generic enough to handle large intra-class variations of the object model among different indoor environments, as well as small inter-class differences between different objects such as doors and elevators. Next, to distinguish an office door from a bathroom door, they extracted and recognize the text information associated with the detected objects. Here first text regions from indoor signs with multiple colours were extracted. Then text character localization and layout analysis of text strings are applied to filter out background interference. The extracted text is recognized by using off-the-shelf optical character recognition (OCR) software products. The object type, orientation, and location can be displayed as speech for blind travelers.

2.3 Eyes of Things :

Responsible Research and Innovation (RRI) is an approach that anticipates and assesses potential implications and societal expectations with regard to research and innovation, with the aim to foster the design of inclusive and sustainable research and innovation. While RRI includes many aspects, in certain types of projects ethics and particularly privacy, is arguably the most sensitive topic. The objective in Horizon 2020 innovation project Eyes of Things (EoT) is to build a small high performance, low-power, computer vision platform that can work independently and also embedded into all types of artifacts. In this paper, we describe the actions taken within the project related to ethics and privacy. A privacy by-design approach has been followed, and work continues now in four demonstrators.

2.4 Blind people guidance system:

This paper describes a guidance system for blind and partially sighted people with the aim of coping in the known and unknown internal and external spaces without the assistance of human guides. This work represents a significant step forward in the application of innovative technological solutions to increase independence and improve the quality of life for people with disabilities. This paper describes the technical and functional architecture of the system for orientation and guidance of a blind person using available modern technology. The described system consists of a digital sensor for determining the location by dead reckoning technique, infrared marks in space and handheld device that provides voice guidance instructions.

2.5 Reading Labels of Cylinder objects for blind :

Here they proposed a camera-based assistive framework to help blind persons to read text labels from cylinder objects in their daily life. First, the object is detected from the background or other surrounding objects in the camera view by shaking the object. Then they proposed a mosaic model to unwarp the text label on the cylinder object surface and reconstruct the whole label for recognizing text information. This model can handle cylinder objects in any orientations and scales. The text information is then extracted from the unwarped and flatted labels. The recognized text codes are then output to blind users in speech. Experimental results demonstrate the efficiency and effectiveness of the proposed framework from different cylinder objects with 3rcomplex backgrounds.

III. PROPOSED METHOD WORKING METHODOLOGY :

In all the above previous system we had found some errors which affects the efficiency of the system and its code. So, we had done very efficient coding in python script which makes the coding side as simple as possible, we also used jessay OS & open cv lib file to make thing more easy.

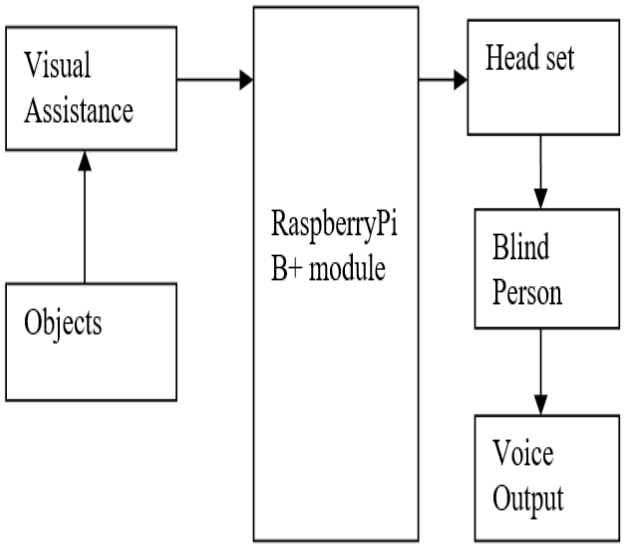

Fig .1 :Block Diagram of the Proposed Method

Fig 2. Example showing road signs

IV. MATERIALS & METHODS

4.1 RASPBERRY PI:

The Raspberry Pi 3 Model B is the third generation Raspberry Pi. This powerful credit-card sized single board computer can be used for many applications and supersedes the original Raspberry Pi Model B+ and Raspberry Pi 2 Model B. Whilst maintaining the popular board format the Raspberry Pi 3 Model B brings you a more powerful processer, 10x faster than the first generation Raspberry Pi. Additionally it adds wireless LAN & Bluetooth connectivity making it the ideal solution for powerful connected designs.

Specifications:

Processor : Broadcom BCM2387 chipset. 1.2GHz Quad-Core ARM Cortex-A53 802.11 b/g/n Wireless LAN

GPU : Dual Core VideoCore IV® Multimedia Co-Processor. Provides Open GL. It is ES 2.0, hardware-accelerated OpenVG, and 1080p30 H.264 high-profile decode. Capable of 1Gpixel/s, 1.5Gtexel/s or 24GFLOPs with texture filtering and DMA infrastructure
Memory : 1GB LPDDR2

OS : Boots from Micro SD card, running a version of the Linux operating system or Windows 10 IoT Dimensions: 85 x 56 x 17mm

Power : Micro USB socket 5V1, 2.5A
Connectors : Ethernet10/100 BaseT Ethernet socket Video Output : HDMI (rev 1.3 & 1.4Composite RCA (PAL and NTSC)

Audio Output : Audio Output 3.5mm jack, HDMIUSB 4 x USB 2.0 Connector

GPIO Connector : 40-pin 2.54 mm (100 mil) expansion header: 2x20 strip Providing 27 GPIO pins as well as +3.3 V, +5 V and GND supply lines

Camera Connector : 15-pin MIPI Camera Serial Interface (CSI-2)

Display Connector : Display Serial Interface (DSI) 15 way flat flex cable connector with two data lanes and a clock lane

Memory Card Slot : Push/pull Micro SDIO

Fig. 3. : RaspberryPi B+Module

Key Benefits:

• Low cost

• Consistent board format & 10x faster processing The main advantage of the raspberry pi module over another processor is raspberry pi is a fully functional linux computer and also compact in size.

4.2 LOGITECH CAMERA : System Requirements is 1GHz = CPU Minimum in which 1.6GHz is CPU Recommended. The Minimum RAM is 1gb whereas the recommended RAM is 2gband also 2.4 GHz Intel Core 2 Duo CPU is Recommended and 2 GB RAM. The Software Support (at release), Logitech Webcam Software 2.0 (LWS).

Camera Specifications:

The Optical Resolution (True)	1280 x 960 1.2MP Image
Capture (4:3 SD)	320x240, 640x480
	1.2 MP, 3.0 MP
Image Capture (16:9 W)	360p, 480p, 720p
Video Capture (4:3 SD)	320x240, 640x480,
	800x600
Video Capture (16:9 W)	360p, 480p, 720p,
Frame Rate (max)	30fps @ 640x480

Fig.4 :Logitech Camera

4.3 HEAD SET:

It converts into a voice output to help blind peoples. Its Audio input is 3.5mmjack. This gets the result from the raspberrypi module and convey to the receiver.

Fig 5: Head set

Fig 5: Head set

V. HARDWARE RESULT

By the efficient programming in the module it recognizes the feeded obstacles, faces , signs and gave it as a required audio output through the earphone. The main purpose of this model is to help blind persons by guiding them using this system design. It recognizes the face, signs, obstacles, humans such as known and unknown persons will be identified using face and text recognition features. It gives the scanned and recognized images in the form of audio output to help and guide the blind person. It is specially designed to blind navigation purpose.

Fig 6 : Hardware output

VI. CONCLUSIONS & FUTURE ENHANCEMENTS

We have proposed a design on face and text recognition based on raspberry pi which is mainly designed for the purpose of blind navigation. Our future work will focus on detecting the emotions of the persons and recognizing more types of indoor objects and icons on signage in addition to text for indoor wayfinding aid to assist blind people travel independently. We will also study the significant human interface issues including auditory output and spatial updating of object location, orientation, and distance. With real-time updates, blind users will be able to better use spatial memory to understand the surrounding environment, obstacles and signs. The future enhancements are given :

Human emotions can be preprogrammed in order to guide blind peoples.

Unknown persons can be identified to great extent.

REFERENCES

[1] Eyes of things(IEEE 2017) by Noelia Vallez ,THALES Communications & Security, 4 Avenue des Louvresses, 92230 Gennevilliers, France .

[2] Camera based analysis of text and documents by Jian Liang, David Doermann . In Proceedings of the IEEE International Conference on Robotics and Automation, 2004.

[3] Context-based Indoor Object Detection as an Aid to Blind Persons Accessing Unfamiliar Environments by Xiaodong Yang, Yingli Tian from The city of newyork, USA. . IEEE Trans PAMI 24(9):1167–1183.

[4] Reading labels of cylinder objects for blind persons by Ze Ye, Chucai Yi and Yingli Tian. Dept. of Electrical

Engineering, The City College of New York,The City University of New York, USA.e-mail: zye01@ccny.cuny.edu.

[5] Blind people guidance system by Ljupko simunovic,Velimir Andelic, Ivan Pavlinusic at Central European conference on information and intelligent system.

[6] Auditory Feedback and Sensory Substitution During Teleoperated Navigation by Rong Liu, Member, IEEE, and Yong-Xuan Wang