Smart Helmet for Visually Impaired

Salman Shaikh[1], Saiyed Farhan Jafar[2], Karan Sosa[3], Pratap Nair[4]

[1,2,3]*Student, Dept. of Computer Engineering, K C College of Engineering and Management Studies & Research, Thane, Maharashtra, India*
[4]*Asst. Professor, Dept. of Computer Engineering, K C College of Engineering and Management Studies & Research, Thane, Maharashtra, India*

---***---

Abstract – *This paper proposes a prototype, Smart Helmet, for aiding visually impaired to identify common real life object such as a person, car, cat, dog, bottle, etc. Our system use Object Detection of Computer Vision to identify objects, Text-to-Speech to provide an audio output about what object is in front of the user and GPS for navigating to home and SOS. The Prototype is implemented using a Raspberry Pi computer along with Pi Camera and GPS Module. The main purpose of this project is to help user identify the object on real-time.*

***Key Words*: Visually Impaired, Computer Vision, Text-to-Speech, GPS, Raspberry Pi, Pi Camera, GPS Module**

1. INTRODUCTION

Visual impairment, also called as vision impairment or vision loss, is the decreased ability to see to a degree that causes complications that are not fixable by usual means, such as glasses. Some also include those who have no access to glasses or contact lenses at earlier stages of poor vision. Visual impairment is not blindness, the term blindness is used for complete or nearly complete vision loss. Visual impairment may cause people difficulties with normal daily activities such as walking, reading, driving and socializing.

Visual impaired rely heavily on sensory organs for object identification which is effective but not efficient. Take the case of identification of a person, a visually impaired can identify a person based on his/her voice, but voice can be mimicked, which can mislead the visual impaired. Our system can identify a person by running object detection and facial recognition.

2. RELATED WORK

2.1. Ground and Obstacle Detection Algorithm[1]

This algorithm is based on 3D depth image obtained from RGB-D camera and attitude angles obtained from attitude angle sensor. Instead of growing threshold, the Sobel edges of image and the boundaries of region are adequate for consideration in the algorithm to improve detection accuracy. Seeds are chosen according to the edges of image, and the stops of growing refers to the Sobel edges of images and growing threshold. The ground and obstacles are roughly detected after region growing, however, all of the regions are not intended. Therefore, regions are combined or excluded according to their boundaries. The results of the algorithm enables the user to differentiate obstacles from the ground.

2.2. Ultrasonic Sensors[2]

This cost efficient guidance system is based on an AT89C52 microcontroller and is developed for facilitating visually impaired. Ultrasonic sensors are used to calculate distance of the obstacles around the user to guide him/her to a safe path. Output is in the form of voice commands which the visual impaired can hear through an audio output e.g., right, left, etc. There's also an advance mode where the system will be able to recognize objects using image processing algorithms.

2.3. IP Cameras for Indoor Navigation[3]

This paper describes a new approach of an ambient navigation system that would aid the visually impaired or blind for indoor navigation (house, office, etc.) without any assistance. The system is composed of IP cameras attached to the ceiling of each room and the smart phone of the subject is used as human machine interface (HMI). Frames are sent to a computer that analyzes the environment, detects and recognizes objects. A computer vision guidance algorithm is designed and used to help the user reach his desired destination with obstacle detection. Feedbacks (alerts, route) are output voice messages by the application to the user. This system provides a reliable solution to assist users in their indoor environment providing them a safe route with obstacle avoidance.

2.4. Computer Vision and Text Recognition using Android Smartphone[5]

The system that uses existing technologies such as the Optical Character Recognition (OCR) and Text-to-Speech (TTS) available on an Android smartphone, to automatically identify and recognize texts and signs in the environment and help the users navigate. The proposed system uses a combination of computer vision and Internet connectivity on an Android smartphone to recognize signs, reconstruct sentences and convert them to speech.

2.5. Virtual White Cane[6]

The Virtual White Cane uses a laser pointer and smartphone to simulate a cane. The laser aligns with the camera to generate a baseline and pan angle.

The camera captures the laser's reflection off the planar surface, which become input to a smartphone application that uses active triangulation to calculate the object's distance from the user. The application then uses the smartphone's vibration to alert the user to the object's proximity.

The user carries an Android Smartphone coupled with a laser pointer. The device simulates a white cane vibrating in the presence of obstacles where the user is pointing to. The metallic structure couples the laser pointer and the user's smartphone. As the laser points to an object near the user, a smartphone application uses active triangulation to measure the object's distance and causes the smartphone to vibrate. As the object gets closer, the vibration intensity increases.

2.6. Vibratory Belt[6]

The Vibratory Belt consists of a Kinect camera connected to an embedded computer, an inertial measurement unit (IMU), and three small vibrating motors. The camera provides depth images, and the computer calculates the distance of the closest obstacles in three positions in front of the user. To separate the objects from the floor, the IMU calculates the camera's orientation.

Obstacle detection by vibrations across waistline through the Belt with three motors; an embedded computer and a Kinect sensor. The Vibratory Belt is easy to use, does not block hearing and detects head-level obstacles. But it is generally used for indoor operations or controlled environments only.

3. PROPOSED METHODOLOGY

3.1. Basic System Flow

The figure 1 shows the basic flow of the Object Detection Processing of the system. The camera captures the live video stream and and send it to the raspberry Pi. The Raspberry Pi inputs the video stream and runs it through the object detection algorithm which detects and recognises the objects as a class and uses text-to-speech to translate the objects as Audio Output.

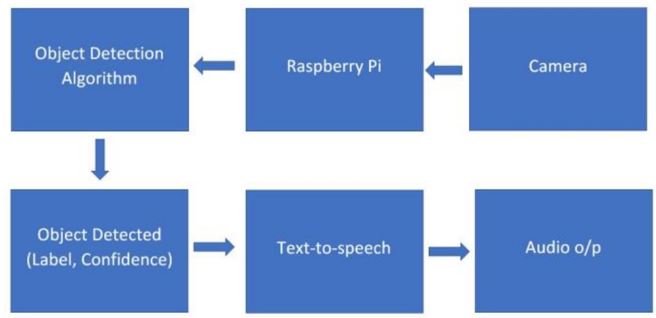

Figure 1

3.2. Bock Diagram

The figure 2 depict the block diagram of the system. The System consist of a Power Supply Raspberry Pi, Pi Camera, GPS Module, 4G Dongle and Stress Button.

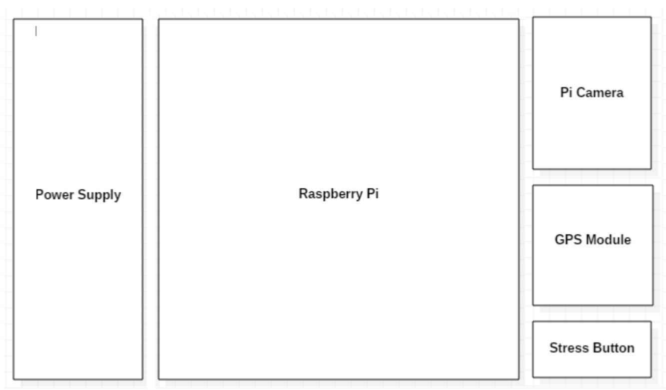

Figure 2

The Raspberry Pi used is a 3rd generation Model B+ with ARM Cortex-A53 1.4GHz Processor and 1GB SRAM. The Power Supply provided to the Raspberry Pi is of 5V 2.5 mAh through a standard USB Power Cable from a Power Bank. The Pi Camera v2 is connected to the Raspberry Pi by the ribbon cable port on the Raspberry Pi through the ribbon cable of the Pi Camera. The GPS Module is provided with a 3.3V power supply from the 3.3 V VCC and GND GPIO pins on the Raspberry Pi. And TX and RX of the GPS Module is connected with the RX and TX GPIO pins on the Raspberry Pi. The 4G Dongle, which will provide the internet connection to the system, is connected to the Raspberry Pi using the USB port. The Stress buttons are implemented by programming the GPIO pins.

3.3. Object Detection Flow Chart

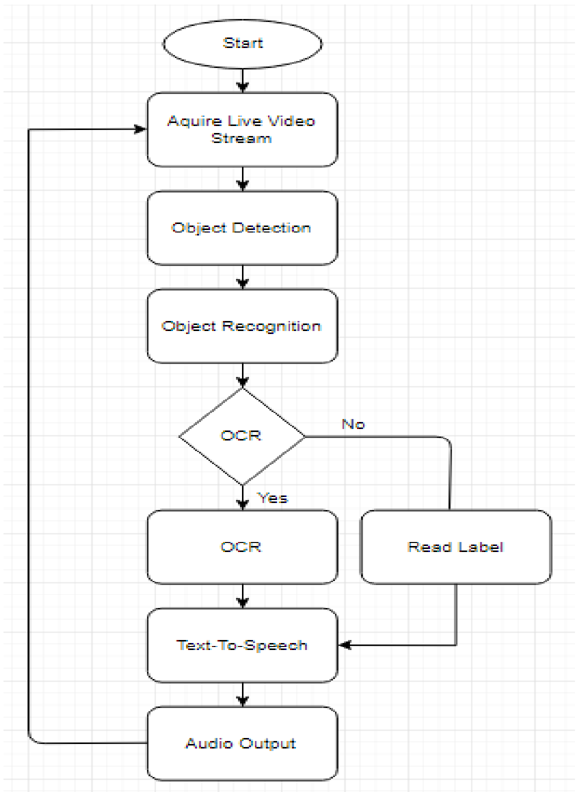

Figure 3

The figure 3 describes the flow of the system. It starts from acquiring the Live Video Stream from the Pi Camera and passing it on to the Object Detection Algorithm, where the objects are detected and recognized. Once the objects are recognized, the system checks if he objects requires OCR processing. The OCR processing are requires for object such as road signs, traffic signs and work ahead signs. If the object detected requires OCR, the objects are processed for text extractions. Once the text is extracted, the Text-to-speech converts the text to audio output. If the object doesn't requires OCR, the System just converts the labels to audio output.

4. TECHNOLOGY USED

4.1. Software

4.1.1. Object Detection

Object detection is a branch of computer vision and image processing that deals with detecting instances of semantic objects of a certain class (such as cats, dogs, humans, buildings, or cars) in digital videos and images. Object detection has applications in many areas of computer vision, including image retrieval and video surveillance. It is widely used in computer vision task such as face detection, face recognition, video object co-segmentation. In our system, object detection is used to detect various classes of objects that can aid visual impaired to carry out mundane tasks.

4.1.2. OpenCV3

OpenCV (or Open source computer vision) is a library aimed at real-time computer vision. We'll be using OpenCV3 along with the deep learning framework Caffe Model for implementing Object Detection.

4.1.3. pytesseract

OCR (or Optical Character Recognition) is the electronic conversion of images of typed, handwritten or printed text into machine-encoded text. We have used the python implementation of tesseract (an OCR engine), pytesseract for OCR processing of object such as roads and traffic signs for text extraction.

4.1.4. GTTS

Google Cloud Text-to-Speech converts text into human-like speech in over 100 voices across 20+ languages and variants. It applies groundbreaking research in speech synthesis called WaveNet and Google's powerful neural networks to deliver high-fidelity audio output. We are using Google's GTTS for providing audio outputs to the user.

4.1.5. GPSD

GPSD is a daemon that receives data from a GPS receiver and provides the data back to multiple applications such as GPS navigation software. It provides a unified interface to receivers of different types. GPSD is used to extract GPS Coordinates (latitude and longitude) from the GPS module for tracking the user.

4.1.6. SMS Gateway

An SMS Gateway enables a system to send and receive SMS messages to and from a SMS capable device over the global telecom network (normally to a mobile phone). An SMS Gateway is used for sending SOS message from the system in case of an emergency.

4.2. Hardware

4.2.1. Raspberry Pi 3 Model B+

The B+ model of the Raspberry Pi 3 is used as its ARM Cortex-A53 1.4GHz Processor and 1GB SRAM is reliable for performing object detection algorithms along with other processing for GTTS, GPS, SMS and more. Also its size makes it suitable for this application as the system is supposed to be mobile.

4.2.2. Pi Camera v2

The Pi Camera v2 having 8 megapixel Sony IMX219 image sensor is customly designed add-on board for Raspberry Pi and is used to provide live video input to the Raspberry Pi for Object detection.

4.2.3. GPS Module

The Neo-6M GPS Module is used for tracking user by extracting the live location from the module and processing it through the Pi.

5. ADVANTAGES

- Object Detection of upto 15 objects per frame
- SOS Message can be sent to registered mobile numbers along with the current location of the user.
- Direction to Home, audio direction to home from current location with both walking and public transit mode depending upon the distance.

6. LIMITATIONS

- The system have a latency of 2-4 seconds due to processing limitations of the Raspberry Pi.
- Neo-6M is accurate to 2.5-10m depending upon the surrounding.

7. CONCLUSION

We have implemented the smart helmet using Object Detection on Raspberry Pi for helping user identify real-life object on real-time. The system also serves as a way-to-home and an SOS messenger for the user. Our system is a stepping stone for a more advance system for visually impaired. The advancements in our system will result in low latency, higher processing speed, higher accuracy, higher efficiency, more features such as live location updates, etc.

8. REFERENCES

[1] Ruiqi Cheng*, Kaiwei Wang*, Kailun Yang*, Xiangdong Zhao*, *State Key Laboratory of Modern Optical Instrumentation, Zhejiang University, China. "A GROUND AND OBSTACLE DETECTION ALGORITHM FOR THE VISUALLY IMPAIRED"

[2] Asim Iqbal, Umar Farooq, Hassan Mahmood, Muhammad Usman Asad, Department of Electrical Engineering, University of The Punjab, Lahore-54590, Pakistan. "A Low Cost Artificial Vision System for Visually Impaired People". 2009 Second International Conference on Computer and Electrical Engineering

[3] Kabalan Chaccour, Georges Badr, "Computer vision guidance system for indoor navigation of visually impaired people". 2016 IEEE 8th International Conference on Intelligent Systems.

[4] Adrian Bhai Rosebrock, PhD Computer Science-Computer Vision and Machine Learning at the University of Maryland, Baltimore County.

https://www.pyimagesearch.com/

[5] Hao Jiang, Thomas Gonnot, Won-Jae Yi and Jafar Saniie Department of Electrical and Computer Engineering Illinois Institute of Technology, Chicago, Illinois, USA. "Computer Vision and Text Recognition for Assisting Visually Impaired People using Android Smartphone"

[6] Juan R. Terven and Joaquín Salas, Instituto Politécnico Nacional, Mexico Bogdan Raducanu, Computer Vision Center, Spain. "New Opportunities for Computer Vision–Based Assistive Technology Systems for the Visually Impaired".