International Journal of Advanced Science and Technology
Vol. 29, No. 7s, (2020), pp. 4636 - 4643

Assisting Visually Impaired People with Entity Detection Using Embedded Systems

[1] P.M.Lavanya,[2] Aishwarya.S,[3] Asifa Ekkem, [4]Guru Prakash Ram, [5] K.Kumaran

[1][2][3][4][5] Assistant Professor, Department of Information Technology, Easwari Engineering college, Tamil nadu,India.

Abstract

Some of the major problems encountered by the blind and visually impaired people are navigation and Entity detection. The sudden spike of technology has led to the invention of several gadgets and devices that are of great use to the visually impaired and blind people. Our proposed system uses Haar Cascade algorithm for object and facial recognition. It uses OCR (Optical Character recognition) for text recognition. USB camera placed on the device takes a picture and detects the object from a database and the name of the device is rendered to the person by speech through their headphones.

Keywords: *Raspberry pi, USB camera, Headphones and Entity Detection.*

1. Introduction

At present, screen free interactive devices are emerging and people are very curious to interact with machines. Instead of depending on interfacing devices like touch screen, button systems we are going to control devices by means of voice. In this system we are going to utilize this technology to help visually impaired people to access the latest technologies without screen.

This device operates in three modes. The first and most important one being object recognition. In this mode, an object in front of the person using the device is identified and it is made known to the user by the means of voice. The second mode is facial recognition, a database is created with the faces of the people important to the user like family or friends. This enables the user to identify the person in front of them. The third mode is text recognition in which plain text can be translated to speech aiding the user further on indoor navigation. The camera used is a USB camera that is attached to the eyeglass of the user and the text is conveyed to the user to with the help of earphones they are required to wear for the proper usage of the device.

For facial and object recognition Haar Cascade algorithm is used. Although facial recognition uses a database with the limited input that is required for users. For text recognition, OCR (Optical Character Recognition) algorithm is used. With the help of this algorithm the words in the page of a text book can be converted to speech which will be conveyed to the user in the form of speech through their headphones.

2. Literature Survey

There are several approaches to helping visually impaired people with entity detection. The detection of the entity, be it a face, object or text it has to be first captured by the USB camera. The detection of objects in images is a technology related to computer vision, image processing and deep learning.

Most of the research papers used two types of sensors, namely haptic sensors and ultrasonic sensors. Haptic sensors also known as touch and feel sensors. Take the input in the form of a touch and gives output as a light signal or an appropriate signal as coded by the maker. Ultrasonic sensors used light and its reflection to find the distance and dimension of objects.

Object Detection with deep learning [1]: A overview published by Zhong-Qiu Zhao, Peng Zheng, Shou-Tao Xu and Xindong Wu in 2019 outlines a comparison of various object detection algorithmsTo improve the performance, typical object detection is discussed along with some modifications and useful suggestions. The tasks discussed here includes object detection, face detection, and pedestrian detection and they have been surveyed. To compare various methodologies experimental analyses have been provided and meaningful conclusions have been drawn.

A Smart walking stick for visually impaired using raspberry pi [12] published in the year 2018 by M.Vanitha, A. Rajiv, K. Elangovan and S.Vinoth Kumar. The smart walking stick will aid blind people to navigate self-sufficiently. This intelligent walking stick contains four ultrasonic sensors, one camera and an earphone. 3 of the sensors that are placed in the stick are used for obstacle detection and the last one of the sensors is used for pothole detection. Camera is used for Identifying the text and object detection. The output will be from an earpiece.

An electronic walking stick for blinds [11] published in the year 2014 by Shashank Chaurasia and K.V.N. Kavith. The white cane has been the main travel aid used by blind people and these days efforts are made to make it better by adding a remote sensor. This way they won't have much of a problem navigating the streets and when they encounter stairs because the improved cane has sharp haptic sensitivity.

Assistive technologies for people with disabilities [10] published in the year 2018 by Nierling Linda. This paper targets mainly on assistive technologies (ATs) that are of great help to individuals with disabilities. Three case studies were taken out: A survey that was carried out over the wide-ranging European scale addressing individuals with impaired visual face to face. Similarly, two more studies were taken with the help of the specialists on ATs. The research conducted targeted on 1) autism spectrum disorders (ASD); 2) Individuals with impaired visual and blindness; 3) Individuals with Impaired auditory and deafness.

Design and Construction of Navguide for Visually Impaired People [9] published in the year 2018 by R. Kiran Rakshana1, C. Chitra2 states the need for these advanced Technologies to make the humans life comfortable. The demand for making lives of disabled people easier, new areas of research has emerged which makes use of electronic mobility aid designed for the blind. Very little of these kinds of Smart systems accessible in the market that makes use of these electronic sensors set up on the cane which also has a lot of fall backs. The Nav Guide arranges the data details based on the neighboring circumstances and prioritizes the data in such a way that it does not lead to an overdose of data. The prioritized data is then given to the user via vibrations and sound responsive manner. NavGuide is used for locating the impediments, wet floors and ascending staircases. It also detects fire and current passing and gives the line following guidance.

Object Detection and Identification for Blind People [7] in Video Scene published in the year 2017 by Hanen Jabnoun, Faouzi Benzarti, Hamid Amiri. this paper introduces a system which restores a visual system for identifying the surrounding objects and this acts as the central function. It uses the concept of local features concept. It uses the SIFT algorithm and key points are matched for the Maximum accuracy of the detecting objects. Thus, the paper presents the concept of a visual replacement system dependent on attribute withdrawal and match up to identify and discover items in pictures.

Designing Media for Visually-Impaired Users of Refreshable Touch Displays [8]: Possibilities and Pitfalls published in the year 2015 by Sile O'Modhrain, Nicholas A. Giudice, John A. Gardner, and Gordon E. Legge. Some of the human factors such as tactile, haptic, vibrotactile, and integrative methods of restoring illustration, charts and prototypes that are based on the effectiveness of strength and weakness. The work is based on the visually impaired research on how to find and prepare the display pictures for the two, paper copy and technology arbitrated demonstration of Braille and palpable graphics.

In all related works the application part was very unique. The compilation of the understanding of all these research work lead to the pondering of a much-required novelty that would make the device a little better.

3. Proposed System

3.1 Overview

The processing time of the voice commands took too long for a given task. It made them feel quite complicated when it comes to subject of visually challenged people. Here, the OCR technology is inherited from existing system and it is combined with facial recognition. It is a portable handy device used to help visually impaired people to identify known persons and helps to read the book. Here, Cascade algorithm was used to identify the persons, pre-trained Mobile net-SSD model used to detect indoor object detection and OCR algorithm helps to read the book or documents.

3.2 System Design

Given diagram depicts the usage of devices such as headphones and USB camera using raspberry pi and cascade algorithm.

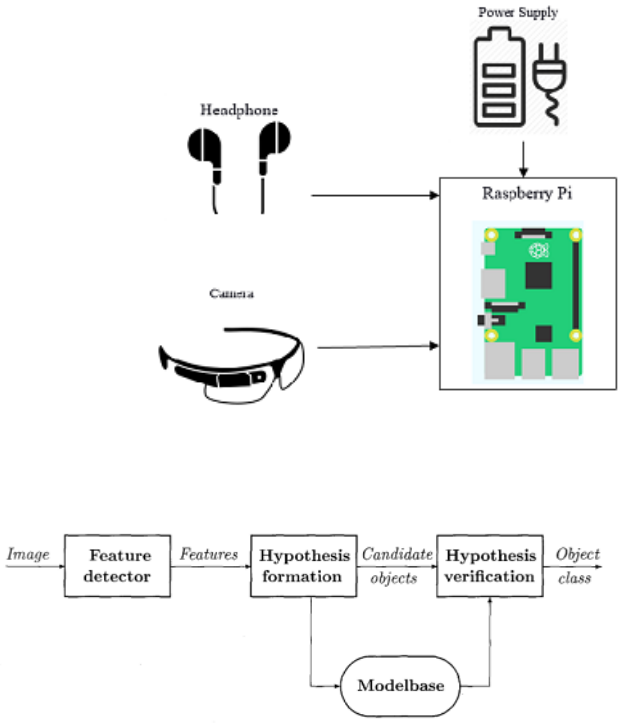


Fig 1: Overall System Flow

3.2 Workflow

- In this Block Diagram, Headphone and USB camera connected with Raspberry Pi. When the system identifies faces or objects it can intimate
- user via attached headphone and uses android mobile to read books aloud.
- The USB camera placed on the device takes a picture and detects the object from a database and the name of the device is rendered to the person by the speech through their headphones.

• It uses cascade algorithm for object and facial recognition. It uses OCR (Optical Character recognition) for text recognition. This algorithm helps the user to convert a whole page of text into speech which is rendered to the user by means of headphones

3.4 Algorithm Used

The algorithm used is Haar Cascade,
1)Initialize the list of class labels MobileNet SSD
2) Detect and generate a set of bounding boxes,
Classes= {Trained dataset}
3) espeak is used as speech synthesizer,
cmd_beg = "espeak "
cmd_out=" -ven+m1 -s 150 -a 300"
4) Iniciate id counter
5) Select the mode of operation,
Button1: Face detection Mode
Button2: Object Detection Mode
Button3: Book Reading Mode
6) Grab the frame from the threaded video stream boxes = [(y, x + w, y + h, x) for (x, y, w, h) in rects]
7) Attempt to match each face in the input image to our known, name = data["names"] [i]
8) update the list of names
9) For object, grab the frame and resize it.
10) Grab the frame dimensions and convert it to a blob blob = cv2.dnn.blobFromImage(cv2.resize(frame, (300, 300)), 0.007843, (300, 300), 127.5)
11) Pass the blob through the network and obtain the detections
12)Filter out weak detections by ensuring the confidence is greater that minimum, confidence > 0.8
13)Show the output frame,
cv2.imshow("Frame", frame)
14) Do a bit of clean up,
vs.release()
cv2.destroyAllWindows()

a. Expected Outcomes

Detecting the face:

Detecting the eyes:

Facial recognition:

Fig 2: Face Recognition

4. Entity Detection

The first module in this Implementation is the detection of the Entity, be it face, object or text it has to be first captured by the USB camera. The detection of objects in images is a technology related to Computer Vision, Image processing and Deep learning. There are two main ways in which the Image is processed when object detection is classified. They are known as HAAR Cascades,

- Pictures that are trained to be identified by the Framework can be known as the Positive picture.
- Pictures that are not trained to be identified by the Framework can be known as the Negative picture.

5. Entity Recognition

The system can only recognise the objects that are trained Dataset and the same goes for Images. Before the process of Training, the images have to be gathered. To make classifiers efficient, it will be better to collect as many as varying versions of the object such as different backgrounds

and Varying lighting effects as possible so that the system can detect the objects efficiently. Images can be taken with camera or downloaded from elsewhere.

For the process of Entity Detection, Haar cascade algorithm is being implemented. The cascade operation should be trained with a couple of sample images. There are four steps when it comes to Haar cascade,

- Feature Selection
- Generate Integral Images
- Adaboost Training
- Classifiers

The Haar attributes considers the close by rectangular area of a particular destination in an Identified square box, adds up all of the picture element magnitudes in every area and figures out the variability between the summations

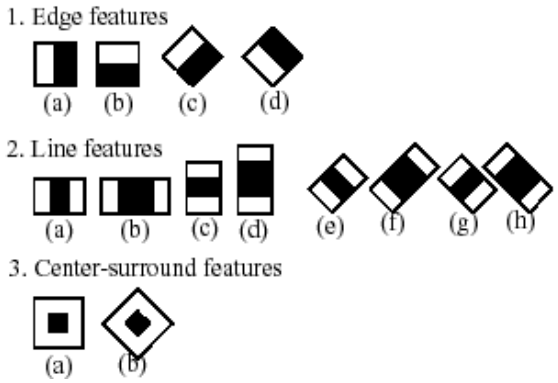
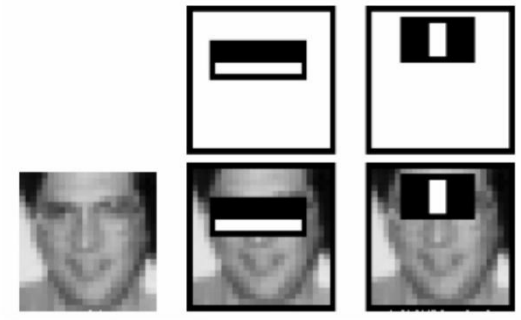


Fig 3: Feature Selection

In the picture below, the topmost row indicates the best quality of image. The first picture points out that the area of the eyes is generally darker than that of the area of the nose and cheeks. Next picture points out that the area of the eyes is always darker than that of the nose bridge. When these rectangular boxes are applied on any other area such as cheeks then it becomes unrelated.



. Fig 4: Generate Integral Images

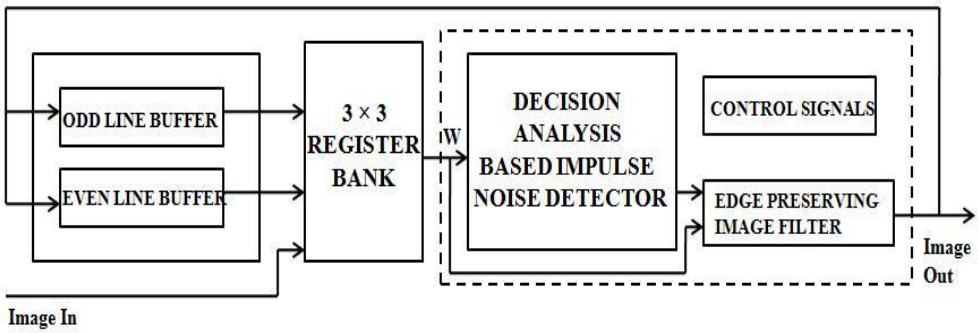
In order to select the best features of them all, we use Adaboost that can be used to capture the finest attributes and takes advantage of the classifier which are trained by it. A square box of Intended length is rolled over the Inserted picture, for every session of picture Haar attributes are determined. In order to describe the features with enough accuracy they are organized into cascade classifiers.

There are different stages involved in Cascade classifiers in which every phase is an accumulation of fragile beginner. These beginners that maybe known as classifiers can be known as Decision stumps. Training of every phase is implemented by making use of a method known as Boosting in cascade algorithm. Boosting can be known as the method that is used to coach a classifier by accumulating the sum of all the conclusions formed by the fragile beginners. Based

upon that current location of sliding square box the classifier labels a region as one of the two, Positive or Negative.

Positive states that the item was detected, if it is negative it states that no item was detected. After a region is complete, the radar glides the square box to the forthcoming destination. When the outcome is positive the classifiers moves the section to the forthcoming stage. Details of an Item detected is then mentioned in existing square box location by the radars.

There has to be at least a couple of negative and positive samples to train the classifiers. There must be a couple of positive pictures with domain of regards mentioned to be adopted as positive samples. Image Labeller can even be used to Label objects. One can provide set of negative samples of image for which the function produces negative samples automatically. For a better accuracy of detection, we can allocate the number of phases, variety of operation and attributes.



6. Conversing The Result

There are two ways in which the audio can be played. It can be played by HDMI or by using head phone Jack. We can switch between these modes anytime we want to. The HDMI cable can be used when TV or HDMI monitors have built in speakers. We can also switch it into Headphone jack when the output is not in HDMI default. We use espeak as the speech synthesis engine.

```
cmd_beg = "espeak "
cmd_out=" -ven+m1 -s 150 -a 300"
```

7. Conclusion And Future Work

By creating a database containing only the faces of the people known to the user, the system helps them identify strangers and therefore saves them from a potential unsafe situation. The text reading function will help the user when the situation arises. In a nut shell, the system will help the user be more independent which is an invaluable innervation every person treasure.

The system can be upgraded to detect and differentiate hazardous and non-hazardous objects. This can be accomplished by creating a dataset containing hazardous objects. The system can be upgraded to detect fire by possibly inserting some smoke detectors. It could be made to send an automatic message to the fire department in case of fires. With the introduction of a panic button which send an automatic message to the police or ambulance in case of emergencies would increase the efficiency of the system immensely.

References

1. Jinqiang Bai 1,* , Zhaoxiang Liu 2 , Yimin Lin 2 , Ye Li 2, Shiguo Lian 2 “Wearable Travel Aid for Environment Perception and Navigation of Visually Impaired People“2019
2. Chung Hyuk Park**:** Telerobotic Haptic Exploration in Art Galleries and Museums for Individuals with Visual Impairments: 2015

3. Kavitha. G1, Varshitha:R2 Survey on Design and Construction of Electronic Aid for Visually Impaired People:2019
4. N.Saranya1, M.Nandinipriya2, U.Priya: Real Time Object Detection for Blind People:2018
5. Sile O'Modhrain, Nicholas A. Giudice, John A. Gardner, and Gordon E. Legge Designing Media for Visually-Impaired Users of Refreshable Touch Displays: Possibilities and Pitfalls: 2015
6. Prof. Seema Udgirkar1, Shivaji Sarokar2, Sujit Gore3, Dinesh**:** Object Detection System for Blind People:2016
7. Hanen Jabnoun1, Faouzi Benzarti1, Hamid Amiri1**:** Object Detection and Identification for Blind People:2017
8. Sile O'Modhrain, Nicholas A. Giudice, John A. Gardner, and Gordon E. Legge**:** Designing Media for Visually-Impaired Users of Refreshable Touch Displays: Possibilities and Pitfalls:2015
9. R. Kiran Rakshana1, C. Chitra2**:** Design and Construction of Navguide for Visually Impaired People:2018
10. Kiruthika, S. U., S. K. S. Raja, and R. Jaichandran. "IoT based automation of fish farming." Journal of Advanced Research in Dynamical and Control Systems 9 (2017): 50-57.
11. Nierling Linda**:** Assistive technologies for people with disabilities:2018
12. Shashank Chaurasia and 2K.V.N. Kavitha**:** An electronic walking stick for blinds:2014
13. M.Vanitha, 2 A. Rajiv, 3 K. Elangovan and 4 S.Vinoth Kumar: A Smart walking stick for visually impaired using Raspberry pi:2018