Object Detection and Human Identification using Raspberry Pi

Dr. Sunitha M.R.
Professor, Computer Science and Engineering
Adichunchanagiri Institute of Technology
Chikkamagaluru, Karnataka, India.
hemayajeyan@gmail.com

Fathima Khan
Bachelor of Engineering: Computer Science and Engineering
Adichunchanagiri Institute of Technology
Chikkamagaluru, Karnataka, India.
fathimah812@gmail.com

Gowtham Ghatge R
Bachelor of Engineering: Computer Science and Engineering
Adichunchanagiri Institute of Technology
Chikkamagaluru, Karnataka, India.
kiranghatge02@gmail.com

Hemaya S.
Bachelor of Engineering: Computer Science and Engineering
Adichunchanagiri Institute of Technology
Chikkamagaluru, Karnataka, India.
sunitharmr2310@gmail.com

***Abstract*—The aim of this paper is to investigate the development of a navigation system for blind and visually impaired people along with human identification. It is based on a microprocessor with synthetic speech output. This solution is designed to give details about human whomever they encounter. The proposed system detects the any obstacle in the path of the blind via stereoscopic sonar system using ultrasonic waves and sends back vibro-tactile feedback to inform the blind about its localization. And if there is an obstacle and then the second system kicks into to check whether the obstacle is human or not, and if proven human, it matches with the existing system database to find out the person details and provide it to the user. Face Detection is an important step in any face recognition systems, for the purpose of localizing and extracting face region from the rest of the images.**

Keywords—Raspberry Pi, Ultrasonic Sensor, Camera Module, Harr Classifier, Adaboost, Object detection, Facial Recognition

I. INTRODUCTION

Since the development of technology, the application for latest technology has been growing exponentially. But even with modern technologies, we sometimes fail to see the basic requirement for many sections of the society. Navigating tools for blind is one among the many. But recent trends show various development in blind navigation-based system for people and wearables are making an entry into the market.

Countries like US, Australia, Canada, New Zealand are still using guide dogs for partially impaired or fully blind people, where they use dogs to guide them to the destination. With Cloud computing and Machine learning improvising every day, it can pathway to build latest gadget for blind to navigate freely. Most of the blind navigation system involves ultrasounds [1].

All blind navigation system as far as now work on the principle of sound reflection using ultrasounds. Sonic Pathfinder, Mowat-Sensor, and Guide-Cane [2] are called obstacle detectors or clear path indicators since the blind can only know whether there is an obstacle in the path ahead. These devices are used to search for obstacles in front of the blind person, and they operate in a manner similar to a flashlight, which has very narrow directivity.

SonicGuide and NavBelt [3], however, are called an environment sensor since it has wide directivity enabling it to search for several obstacles at the same time.

The motivation of this project was to develop a portable navigation aid for blind pedestrians coupled with human identification. The most widely used primary mobility aid today is the long cane. This has several limitations such as a range limited sensor, typically one pace ahead of the user, difficulties detecting overhanging obstacles, and difficulties storing in public places [4].

In this paper, the suggested navigation system involves a microprocessor with speech output. It is a self-contained portable electronic unit. It can supply the blind person with assistance about walking routes by using spoken words to point out any obstacles or names of the person standing in front of them.

On the other hand, and in order to overcome the imperfections of existing electronic travel aids, the proposed not only looks out for objects but also does a facial recognition of people and try to map them to existing database for identification.

II. Methodology

The solution as shown in Figure 1 consists of a high definition Raspberry pi camera module, Raspberry pi Zero W running Raspbian version 9 with Python 3.7.2, HC – SR04 Ultrasonic sensor, Haptic feedback speaker.

Initially, Ultrasonic is used for object detection, which uses ultrasound for that purpose. Upon finding an object, that object is then captured by the camera (triggered from ultrasonic after object detection) for further analysis. These analyses include facial recognition, where the captured image is analyzed for facial features and if found to be human face, then it is compared with the database for finding a facial match and to reveal the username.

The part where obstacles are detected include, an ultrasonic receiver and transmitter operating in 40 KHz frequency signal. These signals are strong enough to detect objects within 2 meter and as low as 2 cm. It starts off by sending a pulse wave of ultrasound.

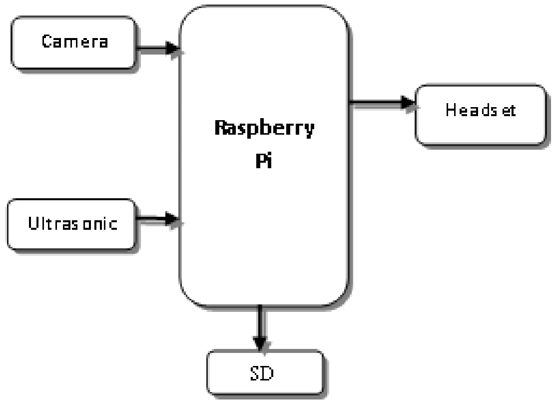

Figure 1 Block Diagram

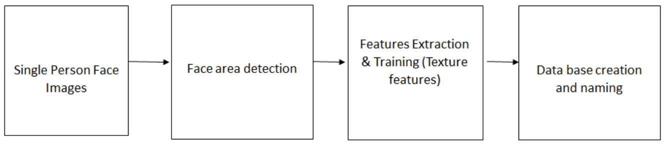

Figure 2 Face Detection

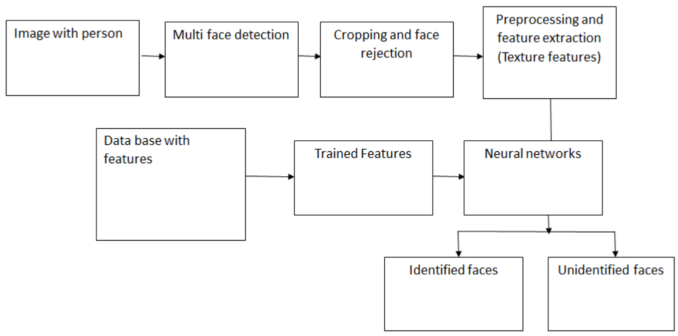

Figure 3 Detailed Face Detection

Eventually the wave is reflected from a solid object in the path of the wave. Upon receiving the reelected wave, we can calculate the distance based on the time taken using the Speed, Distance, Time formula. This information is then used to trigger the camera for detecting the object and analyzing it further.

Once an object is detected, then it is subjected for image analysis to determine whether the object is human or not. This is carried out using Haar classifier and Adaboost.

Haar wavelets are single wavelength square waves which consists of one high interval and one low interval waves. In a 2–dimensional environment, a square wave is a combination of 2 rectangles one dark and one light placed next to each other to form a square as shown in Figure 4. This is the binary method of identification that sets apart from others:

The original rectangle combinations which are used for detecting object are not Haar wavelets, but they are changed combination of rectangle which are suited for visual recognitions. This is the main reason, why we call Harr features as Harr – Like features [5].

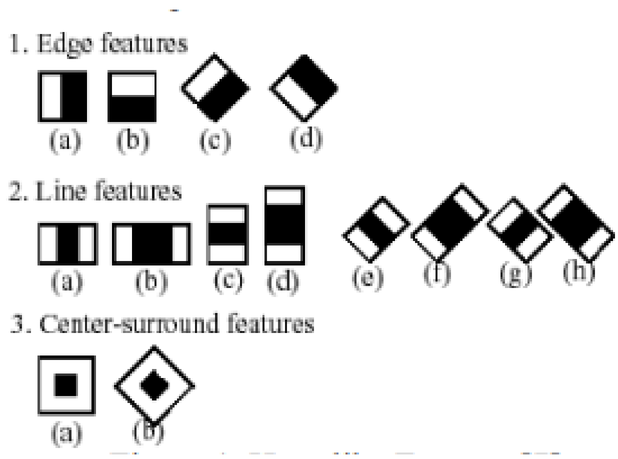

Figure 4 Haar like Features

Harr – Like features are calculated by basically taking the whole average value of the light pixel region and subtracting the dark pixel region from it [6]. If the difference is above the threshold (set during learning), that feature is said to be present. Examples of Haar-like features are shown in Figure 5.

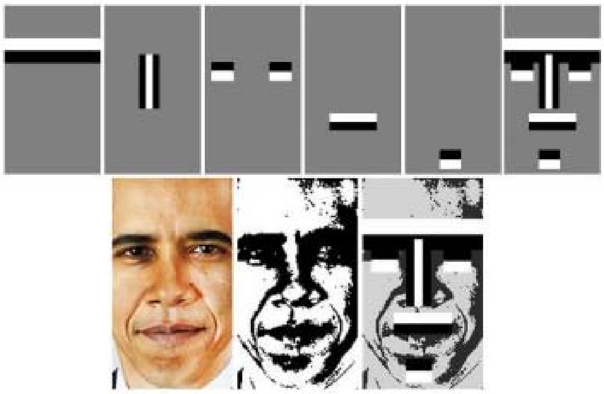

Figure 5 Implementation of Haar like Features

The complete human face detection procedure is shown in Figure 3 briefly, It includes capturing of image, multi face detection and other preprocessing techniques.

To detect a human face using Harr – Like feature, it can be done using only one feature per image or it can be hundred of features per image, all depending upon the complexity of the image, occlusion, structure of face, skin tone, image exposure to light, facial depth etc, All these processes are mentioned in Figure 2.

To find out whether the presence or absence of all the Harr – Like feature are present in any given image, within the image, at a higher scale of depth, we use Integral Image [7]. On the whole, we will be adding the pixel values generated on smaller portion of the image, until we obtain the whole image integral value.

The integral value for the smaller region is basically the sum of all pixels which are above it and to its left. This begins from the left top corner, and then moving towards the right and then moving to the bottom, till the last right bottom most pixel is reached. Each section is taken and integrated seperately and then a combined to form a bigger section and then integrated and then finally the whole image is integrated. [6].

The goal is to get all the integrals from the beginning to the end and save them for further processing. This is carried out by using a summed area table for the image, and then having processed section. This process is called computing integral image. In any given integral image, the rectangular region area can be processed by computing only or any 4 array sets [8].

To improve Haar – Like classification's accuracy and also get the most in real – time performance, we use a different algorithm called AdaBoost – Adaptive Boost Learning Algorithm, that can adaptively select each feature and boost it in each step and them combine them to form a stronger classifier.

As shown in Figure 6,Adaboost learning based training algorithm uses a given set of positive samples, which has various objects of interest, in our usage it's the face, and a set of negative samples, which doesn't have the objects in them. During the training process, selective Haar – Like features are obtained and processed with the image containing positive samples to get the best output in each stage [9].

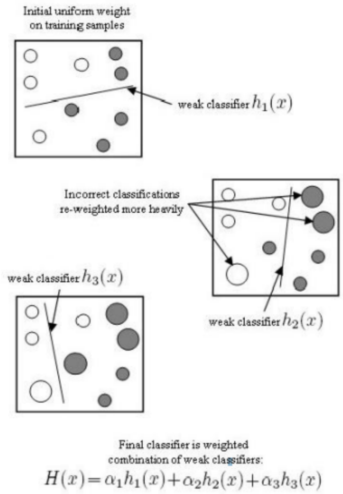

Figure 6 Adaboost Learning Algorithm

III. Technical Description

A. Microprocessor – Raspberry Pi

The Raspberry Pi 3 Model B+ the latest int he line of Raspberry Pi products contains a wide range of features and improvements that allow anyone to integrate Pi into their products. Pi 3B+ is bigger in size when compared to Pi Zero W (smallest in size), but it also has better performance when compared to Pi Zero W. Here are some of the new Pi's specs [10]:

- SoC: Broadcom BCM2837B0 quad-core A53 (ARMv8) 64-bit @ 1.4GHz
- GPU: Broadcom Videocore-IV
- RAM: 1GB LPDDR2 SDRAM
- Networking: Gigabit Ethernet (via USB channel), 2.4GHz and 5GHz 802.11b/g/n/ac Wi-Fi
- Bluetooth: Bluetooth 4.2, Bluetooth Low Energy (BLE)
- Storage: Micro-SD
- GPIO: 40-pin GPIO header, populated
- Ports: HDMI, 3.5mm analogue audio-video jack, 4x USB 2.0, Ethernet, Camera Serial Interface (CSI), Display Serial Interface (DSI)
- Dimensions: 82mm x 56mm x 19.5mm, 50g

B. HC – SR04 Ultrasonic Sensor

The HC-SR04 is an ultrasonic distance sensor. The range of these sensor vary from 400 cm to 2 cm which is of non–contact measuring capability and with the topmost accuracy of 3mm. Every HC–SC04 is equipped with a Ultrasonic receiver and an Ultrasonic transmitter and a logical control unit [11].

C. Raspberry Pi Camera Module

The Raspberry Pi camera module can be used to take high-definition video, as well as stills photographs. It can be used for time-lapse, slow-motion etc. We can also use the libraries that are bundled with the camera to create effects.

It attaches via a 15cm ribbon cable to the CSI port on the Raspberry Pi. It can be accessed through the MMAL and V4L APIs, and also by Picamera Python library [12].

- 5MP sensor
- Wider image, capable of 2592x1944 stills, 1080p30 video
- 1080p video supported
- CSI
- Size: 25 x 20 x 9 mm

IV. Results

Usual face detection algorithms use LBP method for face detection. This means Haar has to perform better in facial recognition for better results.

The algorithms were tested on Raspberry Pi running Raspbian with OpenCV library. This testing running on Broadcom BCM2837B0 quad-core A53 (ARMv8) 64-bit @ 1.4GHz, 1GB LPDDR2 SDRAM and SanDisk Extreme Pro 32GB. The algorithm has been tested using the settings 1.1 scale factor with 4 neighbors' threshold and not setting up any flag, with a minimum of 25x25 detection scale. Haar and LBP had already been tested by using multiple sets of face database:

- The Color FERET Database, NIST USA [13] with 1127 images.
- MIT CBCL Database [14] with 2000 images
- Taarlab Face Database, Brown University [15] with 759 images.

All data sets that have been selected are focus on frontal image and with 30–degree occlusion to right and left.

Table 1 Haar Classifier Result

Algorithm	Dataset	Detected Face	Hit Rate	Detection Speed (ms)
Haar Classifier	Color FERET	976/1127	86.6%	235.4117
	MIT	1670/2000	83.5%	255.5048
	Taarlab	647/759	85.2%	231.5865
Overall		**3293/3886**	**84.7%**	**241**

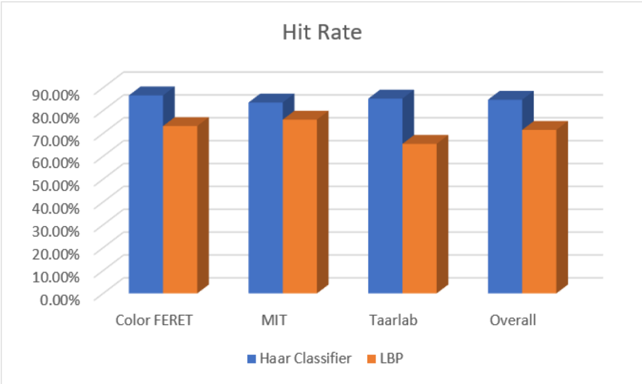

Figure 7 Hit Rate Graph comparing Haar Classifier and LBP

Table 2 LBP Result

Algorithm	Dataset	Detected Face	Hit Rate	Detection Speed (ms)
LBP	Color FERET	826/1127	73.2%	285.5892
	MIT	1520/2000	76.0%	315.5848
	Taarlab	497/759	65.4%	298.0355
Overall		**2843/3886**	**71.5%**	**299.7365**

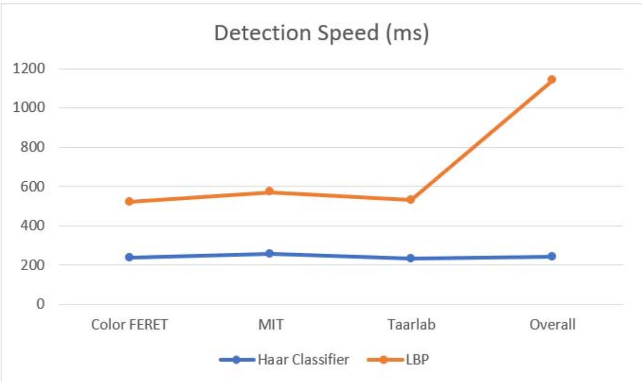

Figure 8 Detection Speed comparing Haar Classifier and LBP

The Figure 7 and 8 show the comparison between the regular LBP method and Haar Classifier along with Adaboost algorithm.

Haar classifier has shown an increased performance in face detection under all three data set as depicted in Table 1 with Taarlab being the largest different ~ 20%. The detection speed is also higher than the regular method. Haar saw an increase of 13.4% with Color FERET data set, 7.5% with MIT data set and 19.8% with Taar lab data set with the overall increase of 13.2% hit rate. The Table 2 shows the result of LBP method.

There is also considerable amount of time reduction with the overall time being 658 ms lesser than the regular time taken.

V. Conclusion

The proposed system has been developed and its main goal is to increase the capability of blind individuals. The technique used here is a well know name in imaging industries, where they reduce the errors by using more advance cameras and algorithms.

With the Ultrasonic, the person would be able to detect each and every obstacle and using camera the person would be able to find out who is exactly who and converse rather easily than be hesitant to ask who whom is. Although the current system can detect objects in the near by location, it can't detect blind spots and other hard corners yet. This is limited with the sensor used in the system and also limited with the principle of ultrasound's reflection.

The results obtain during testing are satisfactory to provide conclusive evidence that Haar – Like classifier with a good Adaboost algorithm, can detect faces more quick than conventional methods.

VI. References

[1] P. Mihajlik, M. Guttermuth, K. Seres, and P. Tatai, "DSP-based Ultrasonic Navigation Aid for the Blind," in Proc. IEEE Instrumentation and Measurement Technology Conference, Budapest, Hungary, May 21-23, 2001, pp.1535-1540.

[2] I. Ulrich, and J. Borenstein, "The guide caneApplying mobile robot technologies to assist the visually impaired," IEEE Transaction on Systems, Man, and Cybernetics-Part A. Systems and Humans, vol. 31, no. 2, pp. 131-136, 2001.

[3] J. Barth, and E. Foulhe, "Preview: A neglected variable in orientation and mobility," Journal of Visual Impairment and Blindness, vol. 73, no. 2, pp. 41-48, 1979.

[4] Y.J. Kim, C.H. Kim, and B.K. Kim, "Design of auditory guidance system for the blind with signal transformation from stereo ultrasonic to binaural sound," in Proc. Int. Sym. Robotics, 19-21 April 2001.

[5] R. Lienhart and J. Maydt. An Extended Set of Haar-like Features for Rapid Object Detection, 2004.

[6] R. Hewitt, 2007 "Seeing with OpenCV: How Face Detection Works", webpage on Cognotics, Resources for Cognitive Robotics [Online]. Available: http://www.cognotics.com/opencv/servo_2007_series/part_2/sidebar.html

[7] Paul Viola and Michael Jones, Rapid Object Detection using a Boosted Cascade of Simple Features. Accepted Conference on Computer Vision and Pattern Recognition, 2001.

[8] C. Souza, 2012, "Haar-feature Object Detection in C#", webpage on CODE PROJECT [Online]. Available: http://www.codeproject.com/Articles/441226/Haar-feature-ObjectDetection-in-Csharp

[9] G. Bradski, A. Kaehler, V. Pisarersky, "Learning-based computer vision with Intels open source computer vision library," Intel Technology Journal, Vol. 9, No. 2, pp. 119-130, 2005

[10] "Raspberry Pi Model 3B+ Spec", webpage on Raspberry Pi [Online]. Available: https://www.raspberrypi.org/magpi/raspberry-pi-3bplus-specs-benchmarks/

[11] "Ultrasonic Distance Sensor – HC–SR04", webpage on SparkFun Electronics [Online] Available: https://www.sparkfun.com/products/13959

[12] "Raspberry Pi Camera Module", webpage on Geee Tech [Online]

Available:
https://www.geeetech.com/wiki/index.php/Raspberry_Pi_Camera_Module
[13] "The Color FERET Database", webpage on NIST [Online].
Available: http://www.nist.gov/itl/iad/ig/colorferet.cfm/
[14] 2004, "Stimulus images courtesy" of M. J. Tarr, Center for the Neural Basis of Cognition and Department of Psychology, Carnegie Mellon University (Brown University), webpage on The CNBC Wiki [Online].
Available:
http://cbcl.mit.edu/software-datasets/heisele/facerecognitiondatabase.html/
[15] 2008, "The MIT CBCL Face Recognition Database", webpage on MIT CBCL [Online]. Available: http://wiki.cnbc.cmu.edu/Face_Place/

978-1-7281-3241-9/19/$31.00 © 2019 IEEE