

# Development of Smart Healthcare System for Visually Impaired using Speech Recognition

## Smart Healthcare System

Khlood Almutairi<sup>1</sup>, Samir Abdlerazek<sup>3</sup>  
Hazem Elbakry<sup>4</sup>

Information Systems Dept. Faculty of Computers and Info.  
Mansoura University, Mansoura, Egypt

Ahmed Ismail Ebada<sup>2</sup>

Research and Development Dept.  
GalaxyTech, Munich, Germany

**Abstract**—This paper presents a solution for the Visually Impaired (VI) based on wearable devices. VI people need support or a guide to support them to locomote from one place to another. Using Wearables help users to achieve a great understanding of the surrounding environment. The proposed system is based on wearable smart glasses to support VI to locomote. It provides a solution integrated with speech recognition to get the destination name and look for the routes. The proposed system is based on Google maps with speech recognition to work as user assistance. The results of the research results proved that the system works with high accuracy of 99% and can help the person as an effective tool for localization guidance. The system can assist VI people to move and have a better life quality.

**Keywords**—Personal assistance; speech recognition; visually impaired person assistant; smart wearable device; smart sensory system

### I. INTRODUCTION

There are about 285 million people who are visually impaired (VI) worldwide based on World Health Organization (WHO). From them, there are about 246 million have low vision and 39 million are totally blind. About 90% of the world's visually impaired live in low-income conditions [1]. Most of VI depend on other people to do their daily activities [2]. This problem is one of the challenges in Europe and USA, because most of VI and disabled live by themselves [3]. One of the biggest obstacles for visually impaired people is that feeling figure. Adapting to experience failure is difficult and takes time. Some people see that the supervisor or support group will help them learn to accept their experience loss so they can move forward toward a richer experience. The significant factor in this often involves developing the skills required to live independently [4]. People who are visually impaired may be born with experience failure or create a visual disability later in life as the result of the accident or heart illness. Some other statements are used to describe the varying degrees of experience loss the individual may have. The terms "visually impaired" and "visual disability" are used to allow all people with reduced experience, irrespective of the severity of experience failure or blindness. However, the following blindness statements and descriptions offer a better explanation of the individual's practical experience [5].

One of the greatest applications she utilizes for mobility is Blind Square, the self-voicing GPS app for those visually impaired that announces crossings and levels of curiosity. Blind Square works with Foursquare, then favorite restaurants, cafés, shops, and other jobs that user frequently go into are flagged. That gives blind users an unbelievable amount of independence, as they don't get to take fellow pedestrians which stores or landmarks they are returning. Blind Square gets them learn [6].

This paper presents a smart wearable sensors system with speech recognition to assist VI people to locomote safely and independently. The proposed solutions employed sensors to track the distance in front of the VI user, alert them, and keep their family the possibility to track them or assist them.

### II. RELATED WORK

There are many researchers used accelerometers and vibration feedback based on gesture recognition as a successful simple user interaction system [5]. A. Ismail et al [6] proposed a healthcare system based on a speech recognition system for smart homes and smart hospitals to help disabled people to control the appliances and depend on themselves on daily based activities.

In [7-19] authors proposed different solutions by integrating speech recognition solutions, a Global positioning system (GPS), and the Internet of Things (IoT) to provide VI with a simple solution that can help them to understand the surrounding area.

Recently, there are concentrations of research on using speech recognition applications like Amazon Alexa in chat bot's application as the interaction between users and devices to find the places, services, and trigger actions [20]. The paper presents a solution that may help visually impaired people to locomote using speech recognition and Google API. We used the research from T. Ashwell et al. [21] as a base to build a speech recognition system for visually impaired people. They used the ASR system to recognize the speech commands from Japanese students who learn English.

In our proposed research, the speech recognition system with the Google API is used to recognize a command from an English speaker based only on some predefined locations by the user. The target from recognition using Google API to find the previously saved locations with a speech command and get

the direction path descriptions as a verbal voice. The system helps the user with localization and with object detection to reach the destination. Besides that, it can detect the obstacle that VI cannot see and the system can ask for video assistance with a caregiver on-demand when it is necessary.

### III. PROPOSED SMART GLASSES

#### A. Proposed System Features

The proposed smart glasses track the street in front of the VI user up to 400 cm and it is worn on the user's eyes. The system works as an alert system that uses an ultrasonic sensor and camera for object detection ahead of the user. The system helps users who wear smart glasses to avoid stumble overdue to the alarm system which is based on the obstacle detection system. The obstacle detection system assists VI users to get alarms using visual recognition and ultrasonic sensors. The caregivers of VI users play only a role when they get a help request from the VI user using a speech recognition command "help" or when a free fall is detected, so they can start monitoring the user and watch a streaming video from the installed camera.

The proposed smart glasses provide a smart solution for VI users to locomote using visual recognition using a camera, a speech recognition system to help the user to provide the system with the target destination. The system uses a sound guide using Google APIs to provide the user with the instruction. It supports 4G connections using a cellular module and GPS modules to find the routes. The proposed system is on sleep mode until it detects a speech command from the user. The proposed smart glasses are lightweight and inexpensive.

#### B. Architecture

The smart glasses use a rechargeable battery as the power source. As shown in Fig. 1, the system structure is based on sensors as inputs, a controller (Raspberry Pi), and outputs as alerts.

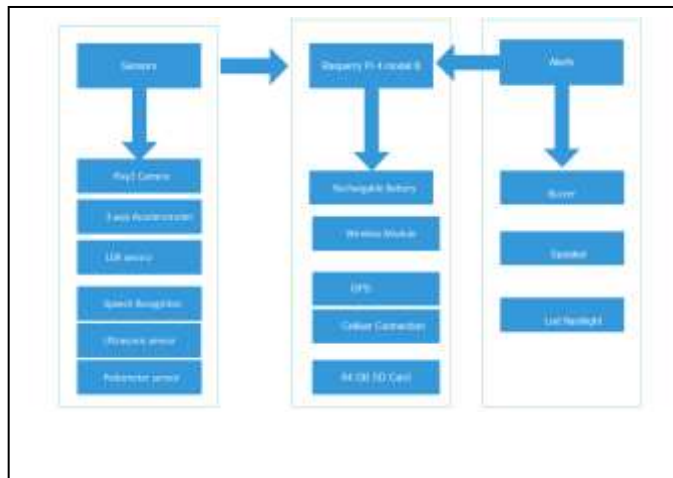


Fig. 1. The Proposed System Block Diagram.

The system used multiple sensors on the proposed smart glasses to work as a guide system for VI users. The system provided two different web APIs on the cloud one for caregivers and the second one for VI users as shown in Fig. 2. The caregiver API accepts provide them with a view of alerts, a possibility to monitor VI user, and a sending voice instruction. The VI user can add new destinations, add new caregivers, ask for help, and receive instructions. The sensors are connected to Raspberry Pi (Fig. 3) which sends the alerts and connects the smart glasses to the internet using the GSM module. The data from smart glasses are saved on the user profile on the cloud.

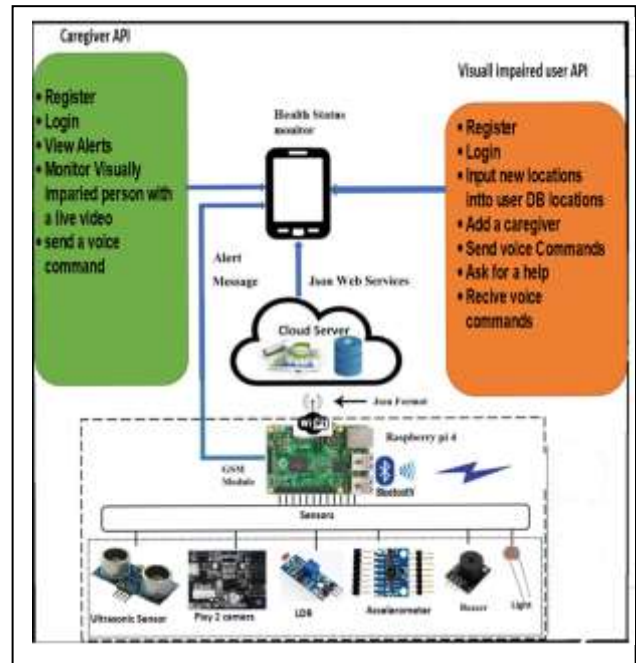


Fig. 2. The Proposed System Architecture.

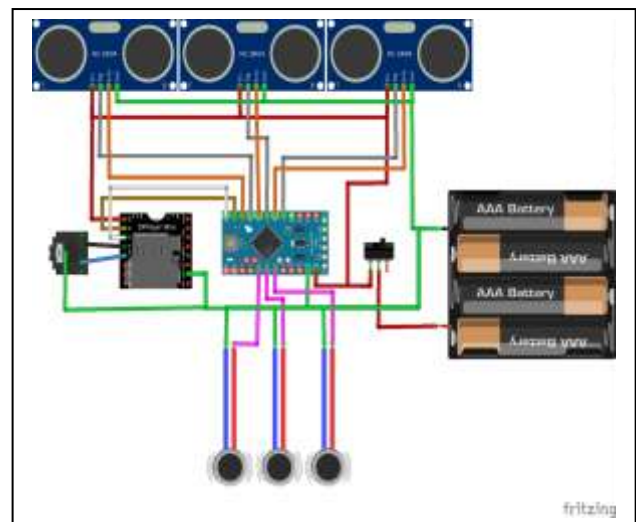


Fig. 3. The Sensor's Connections with the Raspberry Pi Board.

1) *The system controller:* The system is based on the Raspberry pi4 board (Fig. 4). This board has a 64-bit quad-core processor, it is possible to have RAM up to 4GB, USB 3.0, dual-band 2.4/5.0 GHz wireless LAN, Bluetooth 5.0, and Gigabit Ethernet. The board is used rather than Arduino because the proposed system needs a powerful system to handle speech recognition and visual recognition processing.

2) *Sensors:* The proposed smart glasses contain different types of sensors to work as a guide tool. They contain an ultrasound sensor, camera, accelerometer, LDR sensor, and GPS. The overall system depends on the interaction between the different multi-sensors signals to decrease the false alarms as shown in the block diagram in Fig. 5.

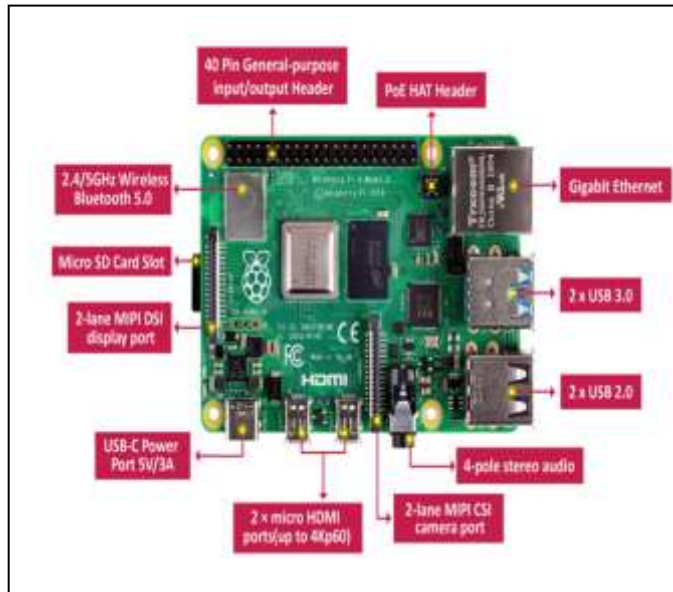


Fig. 4. Raspberry Pi Model 2.

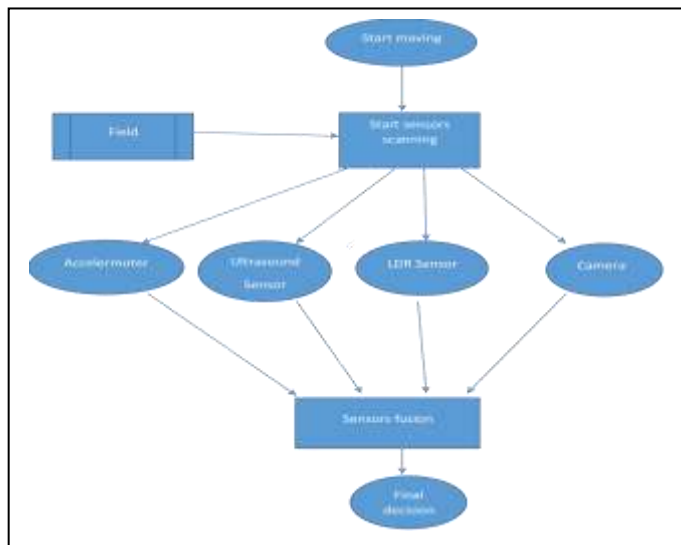


Fig. 5. The Sensory Decision Support System.

The proposed system components are detailed in the following:

a) *Ultrasound Sensor (US):* The proposed system employs US to detect the obstacles in front of the VI user. US detect the distance using ultrasonic waves as shown in Fig. 6. The sensor head emits an ultrasonic wave that is reflected from the target object. Ultrasonic sensors detect the distance to the target object by measuring the time between sending and receiving the ultrasonic wave. An optical sensor has a transmitter and a receiver, while an ultrasonic sensor uses the same ultrasonic element to transmit and receive. In a reflection type ultrasonic sensor, the ultrasonic waves are alternately transmitted and received by a single oscillator. This means that the sensor head can be kept very small. The distance can be calculated using the following formula:

$$\text{Distance } L = 1/2 \times T \times C$$

where L is the distance, T is the time between transmission and reception, and C is the speed of sound (the value is multiplied by 1/2 because T is the time it takes the wave to travel there and back).

The ultrasound sensor sends the sum of the distance between the sensor and the nearest obstacle and the distance can be calculated by dividing the total distance by 2 because the total distance is the distance from the sensor up to the obstacle forward and backward. The sensor sends the distances in centimeters when it is connected to the Arduino board as shown in Fig. 7.

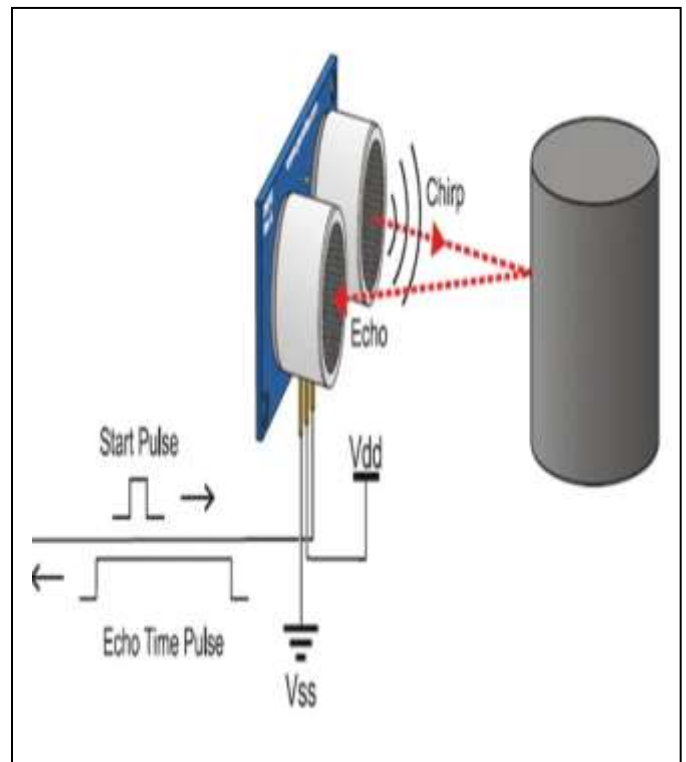


Fig. 6. The Ultrasound Sensor.

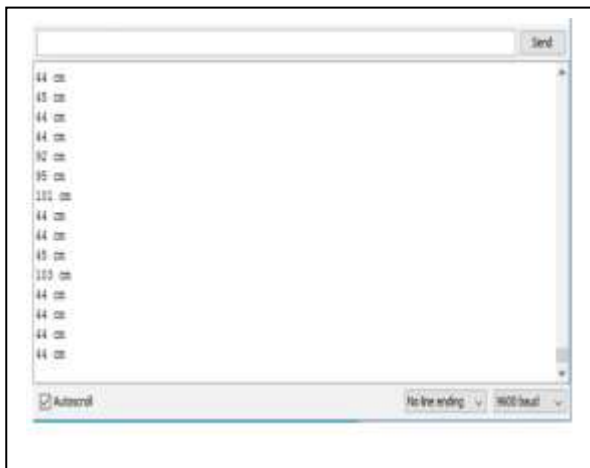


Fig. 7. The Measurements of the Ultrasound Sensor.

b) *Camera*: The proposed system is provided with a Pixy2 Camera (Fig. 8) which has two functionalities. The first is to detect obstacles by the recognition of the obstacles and recognizing the streets to guide VI users as a human eye because GPS has some errors in defining the exact location. The second function of the camera is to work as a watching person to provide the caregiver a possibility to watch the VI user on demand. The Pixy 2 camera is a great solution for visual recognition applications because it is easy to train the camera to recognize the objects. The camera is used to help VI user to recognize when there is a hole, Lighting poles, or dangerous places.

c) *LDR Sensor*: LDR sensor is used to detect darkness to turn a led on the smart sunglasses to provide a clear video streaming for caregivers when their help is needed by the VI users. As shown in Fig. 9, the led is turned on only when the LDR detects the darkness and then the circuit for the led will be closed to turn it on.

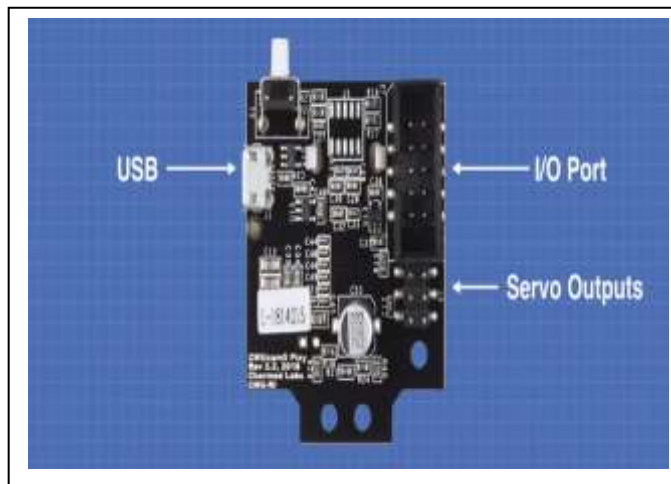


Fig. 8. Pixy2 Camera for Image Recognition.

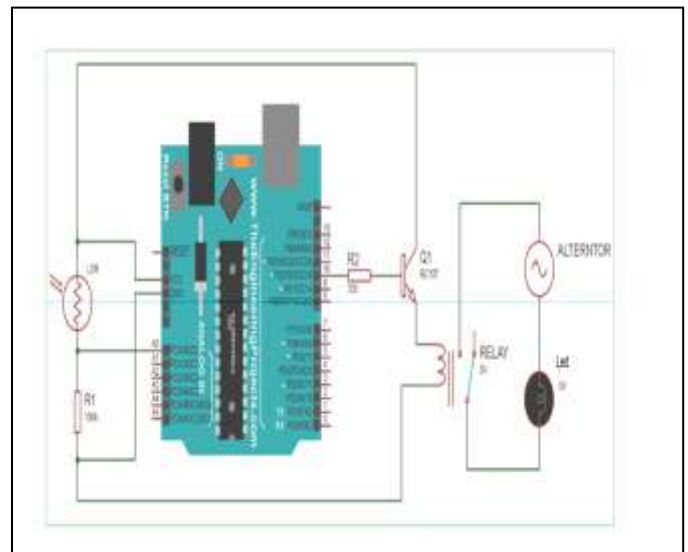


Fig. 9. The Led Connection with LDR Sensor

### C. Speech Recognition

The paper presents smart glasses with a speech recognition module, which is trained to recognize the speech of the user. As shown in Fig. 10 the user starts to give an input speech as the name or location name. Then applying feature extraction to find a similarity score for the command based on a database of the speaker model with the user-defined commands.

The proposed system used the speech recognition module in Fig. 11 installed on the smart glasses. This module is a compact and easy-to-use speech recognition board. It is a loudspeaker dependent speech recognition module. It supports up to 80 voice commands in total. A maximum of 7 voice commands can work at the same time. Each tone can be trained as a command. The users must first train the module before they can recognize any voice command. It has 2 control options: serial port (full function), general input pins (part of the function). Common output pins on the board can generate different types of waves while recognizing appropriate voice commands. The V2 speech recognition module supports 15 commands in all and only 5 commands at the same time. In V2, voice commands are divided into three groups as you practice them. And only one group (5 commands) could be imported into Recognizer. It means that only 5 voice commands are effective at the same time. On V3, voice commands are stored in a large group like a library. All 7 voice commands in the library can be imported into the recorder. Parameters: Voltage: 4.5 - 5.5 V, current: <40 mA, digital interface: 5 V TTL level for UART interface and GPIO, analog interface: 3.5 mm mono channel microphone connection + microphone pin interface. Dimensions: 31 x 50 mm. Detection accuracy: 99% (in ideal surroundings). Features: Supports a maximum of 80 voice commands, with each voice 1500 ms (speak one or two words). A maximum of 7 voice commands effectively simultaneously. Arduino library is included. Easy control: UART / GPIO. User-controlled general pin output.



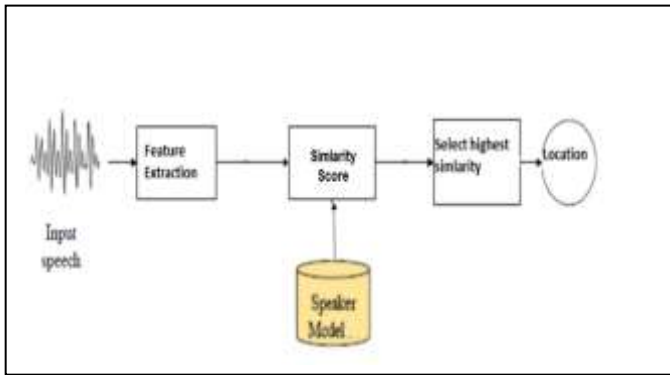


Fig. 10. System Overview.



Fig. 11. The Speech Recognition Module.

The proposed system as shown in Fig. 12 starts with the recognition process by matching the input command against the command templates by using the Dynamic Time Warping (DTW) algorithm which is based on measuring similarity between two temporal sequences. The DTW is used for authentication purpose if the user is authorized, then the system stat to receive the command which contains the destination name as pre-recorded in the speaker database. If it is found in the database, then the user can get a description of the route to the destination.

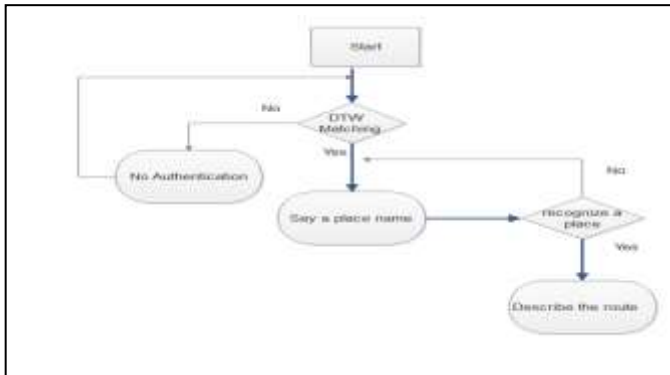


Fig. 12. The Proposed System Flowchart.

The system is based on Google Map APIS to get the route description to describe the route.

#### D. Visual Recognition

The camera in the proposed system is used for two purposes, first as an object recognition tool to detect an obstacle in front of the VI user or as an on-demand streaming tool to show the path to the caregivers to describe the path for the VI people as shown in Fig. 13.

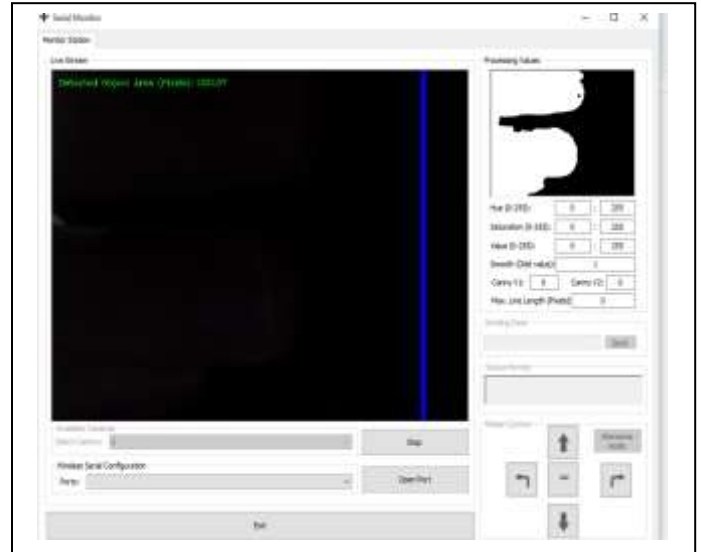


Fig. 13. The Camera Interface on the Caregiver Side.

#### E. System Implementation

The proposed system is comprised of Raspberry Pi and multiple sensors. The software is compiled on the board. Every sensor has a code to send the parameters to the board. The Raspberry pi board and the sensors are integrated with the open-source code community. Functions such as communications and GPS readings are available as libraries, which simplify the code and implementation. The proposed system is supported by a battery for the components and the controller. The battery can work for 2 days and the user can extend the battery time with a power bank if it is needed. The sensors, which were connected to the board, send signals that trigger the predefined events. The ultrasonic sensor gets the power directly from the battery because it is active most of the time while the system is running.

The board requires 5.5 V and has a low power consumption. The alarms, including sound, and camera to transmit video to the caregivers are triggered by the sensor events and therefore consume power during short periods. The proposed system proposed a complete coverage algorithm to define the path based on Google maps. The system uses an obstacle avoidance algorithm to cross through obstacles. The pseudo code of this algorithm is:

```

Pseudo code for the user's motion
if (the camera recognized an obstacle && Ultrasonic
sensor detected an obstacle)
{
    Stop
    send an alert signal
}
}
if (the accelerometer detected a free fall of the
user)
{
    If the Pedometer did detect motion for 30
seconds
    {
        Alert
        Send a call for a help
    }
}
}
    
```

If the sum of all probabilities of overall the system is more than one, the systems give a sound command with the direction and the exact location of the obstacle to overcome it. The aim of this algorithm to detect the obstacles and then check whether they are high, or not. If the camera detected an obstacle that is a danger sign or any predefined object, it sends a signal to the caregiver as shown in Fig. 14 and 15.

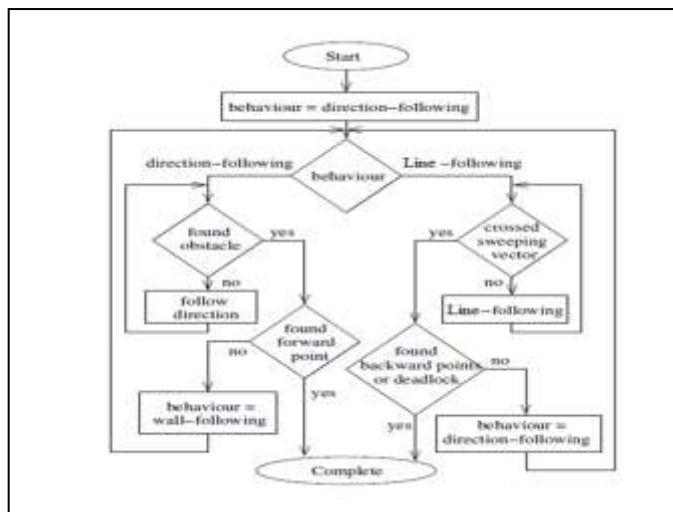


Fig. 14. The Obstacle Avoidance Flowchart.

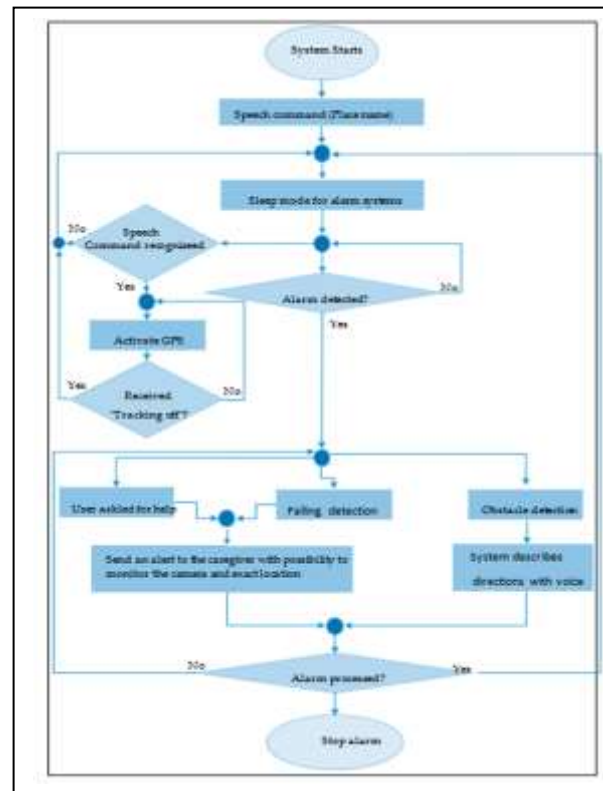


Fig. 15. Basic Flowchart of the System Operation.

If  $(P(S1) + P(S2) + P(S3)) \geq 1$

Then: There is a mine

P (S1) is sensor probability of obstacle existence by ultrasonic sensors.

P (S2) is a sensor probability of darkness existence by the LDR sensor.

P (S3) is a sensor probability of danger existence by the camera.

The proposed smart glasses are as in Fig. 16, a light tool that works like a small computer for visual recognition, speech recognition, and an interaction tool between VI users and their caregivers.

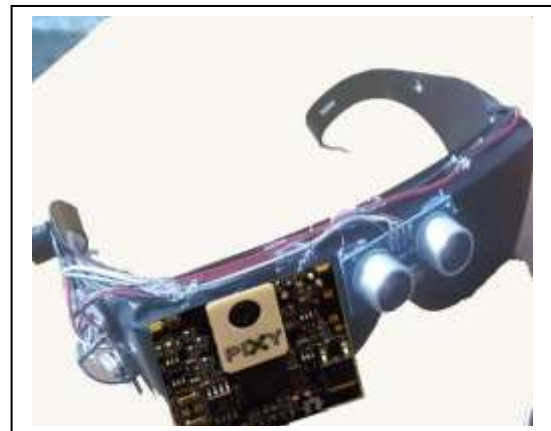


Fig. 16. The Smart Glasses Prototype.

#### IV. TESTS AND RESULTS

The proposed system is tested by users with different accents with an age range between 15 and 30. The test is based on speech commands and locomotion. The proposed system showed efficiency and adaptability with different situations because it gives the user possibility to add new commands to new destinations, new objects to be recognized by the camera. The proposed system was very comfortable for the user while testing because they need to wear only glasses. The system has 15 commands for testing the first section of usually used items designed to elicit performance on 15 grammatical features (possessives, plural -s, 3rd person -s, questions, comparative adjectives, relative clauses, conditionals, modal verbs, relative adverbs, verbs. Multiple instances of each feature appeared in the test. Items ranged in length between 2 and 4 words. The 15 items in each section were arranged in random length order so that the items would be not relatively to find out the performance of the recognition according to a limited number of matching items.

The test items are formulated as orders that contain in between one of the preserved entries as target destinations. The user must record the commands as the name of the destination as a separate entry. The entries as shown in Table I should be short and has a defined name with the user speech and a saved spot on the Google map account. The proposed application is designed and developed with the Xamarin tool; it is a tool from Microsoft to develop cross-platform applications. The application can work on android, windows, and iOS platforms. The application is developed as simple as possible to be used easily by visually impaired people. The application just expects a word of some words from the user to start looking for the target destinations.

TABLE I. USED SENTENCES IN THE SYSTEM TEST

The Sentence	The order
Go Home	Home
Go to my Work	Work
Go to Metro station	Metro
Go to the train station	Train
Go to Restaurant	Restaurant
Go my friend Ali	Ali
Go my sister Lana	Lana
Entertainment near to me	Entertainment
I go now	Green
Call police	Police
Call Ambulance	Ambulance
Go to a concert	Concert
Go to the nearest Supermarket	Market
Call emergency	Call Emergency
Where am I	Where

After testing the system with those registered 15 sentences, the system worked with a success rate of 99%. We tested even changing the order of the sentences with the system with native speaker input to test the system accuracy and the system could match them successfully. The best practice with the speech orders to try to use short sentences with a place or an alias name of the destinations or friends' names. The use of a camera mounted on the smart glasses provided the system with real-time images from the environment to provide the caregivers with the possibility to help VI people in difficult situations.

#### V. DISCUSSION

The proposed system proposed a solution with speech recognition and visual object identification to help VI people in navigation. We developed a system that improved the proposed system from RAMADHAN [18] by improving the accuracy of speech recognition and we added more details of the detection of the objects in front of the VI person by using video recognition by an installed camera on the smart glasses. The system gives the VI people a better experience because it is lightweight smart glasses that is more convenient to wear than a wrist wearable device. The system provides a tracking system for the VI person that can be accessed by a family person to lead them in hard situations. There were some challenges while developing the system such as the design size and battery life. The system provided an optimal low-cost solution with most of the features that can help the VI to navigate by themselves.

#### VI. CONCLUSION

The paper presented a successful solution for VI users to help them to overcome the obstacle they face while the walk alone from point A to point B. The proposed system in this paper is smart glasses that contain a Raspberry Pi board, camera, ultrasound sensor, speech recognition module, and LIDR sensor. The proposed system provided a successful solution based on Google Maps APIs to describe the path to the users based on the pre-recorded commands of the destination the VI user usually visit. The system proposed an applicable solution based on machine learning and IoT [22-27] that can help VI people to have a better life and locomote.

#### ACKNOWLEDGMENT

- The article authors declare there is no interest conflicts in the article.
- Consent for publication: Informed consent was obtained from all individual participants included in the study.
- Data availability: Data is available on request.
- The article did not get funds for the research.
- We accept all ethics approval and consent to participate in the scientific research area.
- All Authors worked on the proposed research.
- Acknowledgments: We would like to thank GlaxyTech, Munich, Germany for supporting devices and sensors and all needed requirements for implementations.

REFERENCES

- [1] R. Jothi, M. Kayalvizhi, K. Sagadevan, "Smart walking stick for visually challenged people". *Asian J. Appl. Sci. Technol.* 2017, 1, 274–276.
- [2] World Health Organization. Available online: <http://www.who.int/en/> (accessed on 9 October 2020).
- [3] T. Volkel, R. Kuhn, G. Weber, "Mobility Impaired Pedestrians Are Not Cars: Requirements for the Annotation of Geographical Data. In *Miesenberger*", Springer, Heidelberg, vol. 5105, pp. 1085–1092, 2008.
- [4] Sangami, A.; Kavithra, M.; Rubina, K.; Sivaprakasam, S. "Obstacle detection, and location finding for blind people". *Int. J. Innov. Res. Comput. Commun. Eng.* 2015, 3, 119–123.
- [5] Kher Chaitrali, S.; Dabhade Yogita, A.; Kadam Snehal, K.; Dhamdhare Swati, D.; Deshpande Aarti, V. An intelligent walking stick for the blind. *Int. J. Eng. Res. Gen. Sci.* 2015, 3, 1057–1062.
- [6] Ismail, A., Abdlerazek, S., & El-Henawy, I. M. (2020). "Development of Smart Healthcare System Based on Speech Recognition Using Support Vector Machine and Dynamic Time Warping", *Sustainability*, 12(6), 2403.
- [7] M. Aziz, T. Owens, U. Zaman, "Received Signal Strength Indicator-Based Localization of Bluetooth Devices Using Trilateration: An Improved Method for the Visually Impaired People", *International Journal of Electronics and Communication Engineering*, 2019.
- [8] N. Ramaprasad, P. Narayanan, "Volunteered Geographic Information System and Its Contribution in Service Sector Employment", In *Geographic Information Systems*. IntechOpen, 2019.
- [9] S. Kotyan, N. Kumar, P. Sahu, V. Udutalapally, "Drishtikon: An advanced navigational aid system for visually impaired people", *arXiv preprint arXiv*, pp. 19040.10351, 2019.
- [10] A. Montuwy, B. Cahour, A. Dommès, "Using Sensory Wearable Devices to Navigate the City: Effectiveness and User Experience in Older Pedestrians", *Multimodal Technologies and Interaction*, vol 3(1), 2019.
- [11] Khatri, A. Assistive vision for the blind. *Int. J. Eng. Sci. Invent.* 2014, 3, 16–19.
- [12] Gayathri, G.; Vishnupriya, M.; Nandhini, R.; Banupriya, M.M, "Smart walking stick for visually impaired", *Int. J. Eng. Comput. Sci.* 2014, 3, 4057–4061.
- [13] Rao, B.; Deepa, K.; Prasanth, H.; Vivek, S.; Kumar, S.N.; Rajendhiran, A.; Saravana, J. Indoor navigation system for the visually impaired person using GPS. *Int. J. Adv. Eng. Technol.* 2012, 3, 40–43.
- [14] Chandana, K.; Hemantha, G.R. Navigation for the blind using GPS along with portable camera-based real-time monitoring. *SSRG Int. J. Electron. Commun. Eng.* 2014, 1, 46–50.
- [15] Kumar, M.N.; Usha, K. Voice-based guidance and location indication system for the blind using GSM, GPS, and optical device indicator. *Int. J. Eng. Trends Technol.* 2013, 4, 3083–3085.
- [16] Q. Rabbani, G. Milsap, N. Crone, "The potential for a speech brain-computer interface using chronic electrocorticography", *Neurotherapeutics*, vol 16(1), pp. 144–165, 2019.
- [17] José, J.; Farrajota, M.; Rodrigues, J.M.; du Buf, J.M.H. The SmartVision local navigation aid for blind and visually impaired persons. *Int. J. Digit. Content Technol. Appl.* 2011, 5, 362–375.
- [18] RAMADHAN, A. Wearable smart system for visually impaired people. *Sensors*, 843, 2018, 18.3.
- [19] N. Giudice, W. Whalen, T. Riehle, S. Anderson, S. Doore, "Evaluation of an Accessible, Real-Time, and Infrastructure-Free Indoor Navigation System by Users Who Are Blind in the Mall of America", *Journal of Visual Impairment & Blindness*, 2019.
- [20] A. Nassif, I. Shahin, I. Attila, M. Azzeh, K. Shaalan, "Speech Recognition Using Deep Neural Networks: A Systematic Review", *IEEE Access*, vol 7, pp. 19143–19165, 2019.
- [21] T. Ashwell, J. Elam, "How Accurately Can the Google Web Speech API Recognize and Transcribe Japanese L2 English Learners?" *Oral Production? Jalt Call Journal*, vol 13(1), pp. 59–76, 2017.
- [22] Ahmed Ismail, Samir Abdlerazek & Ibrahim M. El-Henawy, "Applying Cloud-Based Machine Learning on Biosensors Streaming Data for Health Status Prediction", *The 11th International Conference on Information, Intelligence, Systems and Applications 15 – 17 July 2020 Piraeus, Greece, IEEE*.
- [23] Ahmed Ismail, Samir Abdlerazek & Ibrahim M. El-Henawy. "BIG DATA ANALYTICS IN HEART DISEASES PREDICTION." *Journal of Theoretical and Applied Information Technology* 98, no. 11 (2020).
- [24] A. Ismail, A. Shehab, I. El-Henawy, "Healthcare Analysis in Smart Big Data Analytics: Reviews, Challenges, and Recommendations", In *Proceedings of Springer*, 2018.
- [25] A. Ismail, A. Shehab, I. El-Henawy, "Quantified Self Using IoT Wearable Devices", In the *International Conference on Advanced Intelligent Systems and Informatics 2017. AISI 2017. Advances in Intelligent Systems and Computing*, vol 639. Springer, 2017.
- [26] A. Ismail, M. Elmogy, H. ElBakry, "Landmines Detection Using Low-Cost Multisensory Mobile Robot", *Journal of Convergence Information Technology*, Volume 3, Issue 4, November 2015.
- [27] A. Ismail, M. Elmogy, H. ElBakry, "Landmines Detection Using Autonomous Robots: A Survey", *International Journal of Emerging Trends & Technology in Computer Science (IJETTCS)*, Volume 3, Issue 4, 4, July-August 2014.