Blind Echolocation Device with Smart Object Detection

Sidra Abid Syed*, Taha Mushtaq Shaikh, Fatima Aijaz, Nimra Mehmood, Shujaat Ahmed

Department of Biomedical Engineering, Faculty of Engineering Sciences and Technology, Ziauddin University, Karachi, Pakistan.
*Corresponding author: sidra.agha@zu.edu.pk

Abstract

Blindness is one of the world's biggest issues. According to the recent data of World Health Organization (WHO), 39 million of world's population is blind. World Health Organization (WHO) developed many polices, techniques and strategies to prevent blindness. Another organization is developed with the name World Access for the Blind under the leadership of Daniel Kish who is also a blind person but uses a technique of echolocation called flash sonar. Our idea is based on sensory substitution in which one sensory loss is substituted by the other. We proposed an innovative rehabilitation prototype device that uses a sonar sensor and a bone conduction speaker to perform and enhance the ability of human echolocation. The rehabilitation period can be longer and can vary from person to person. Furthermore we have also installed a smart system that can perform object detection and image to speech synthesis using Python with the help of smartphone camera during a live video streaming. Hence, this prototype device can navigate, echolocate and detect object in its environment. The proposed system is verified through different environmental inputs and objects. The proposed technique allows the blind individuals to perceive through ears and let them perform echolocation at the same time.

Keywords—Blindness, blind aid, echolocation technique, rehabilitation, object detection, artificial intelligence, Python.

1 Introduction

Approximately 1.3 billion people across the globe are estimated to live with some form of distant or near vision impairment. 188.5 million people have a mild impedance of vision with respect to remote vision, 217 million people have medium to severe impedance of vision, and 36 million people are visually impaired. The estimates are expected to reach 200 million visually impaired people by 2020 [1].

With regard to this disease, two famous blind individuals Daniel Kish and Ben Underwood have practiced human echolocation and made the environment friendly for them through echolocation. Kish, who was blind from birth, taught that he experiences the sensations like images, and developed a technique of flash sonar. In his organization, called "World Access for Blind", he trains other people to use the same technique. Kish basically works as a rehabilitation therapist for blinds. This therapy requires time instances which depends on a person's age the optic nerve's ability to integrate sound [2][3]. Hearing for blind people provides a way to access environmental data. Hearing gives the blind individual information about sound sources and the artifacts that represent sounds in the ambient environment [4].

Echolocation is the ability to use reflected sound to provide inter-spatial environmental information. Just as certain types of bats and marine mammals make their own sound emissions, the same technique can be used by people to echolocate. In fact, some blind people have trained themselves to echolocate using mouth clicks. The beam pattern of mouth clicks made by blind echolocators shows a gradual 5 dB decrease in intensity as a function of angle from straight ahead to 90 degrees to the side. However, click energy is attenuated more heavily at other angles [5]. An important and practical use of echolocation is to detect and locate objects in space, as this enables the echolocator to safely navigate their environment. There are a number of physical properties that can affect how echolocation is used, such as deciding whether an object is in front or not, how big it is, how far away it is, and what kind of material it is made of [6]. Typically, it involves the echolocator creating a sonic emission (e.g. an oral click) that personifies the environment and induces the echoes reflected back to the echolocator. In particular, some people with vision loss are known to develop

ISSN: 2523-0379 (Online), ISSN: 1605-8607 (Print)

echolocation skills at an exceptional level often without formal instruction, and usually use a tongue click as their preferred type of sound emission to achieve this. Blind people can thus access many properties of distal objects in the environment that would otherwise be accessed through vision, such as distance position, size, and shape and distal object material [7].

Bone conduction was discovered by Beethoven, the prolific 18th century composer who was almost completely deaf. By adding a wire to his piano and clenching it in his jaws, Beethoven found a way to hear the sound of the piano by his jawbone. As vibrations were passed from the piano to his jaw, he obtained awareness of the sound. Research has shown that, apart from eardrums, sound will enter our auditory system through another medium and that our bones can serve as the other medium. The bone conduction devices (such as headphones) play the role of your eardrums in bone conduction. Such machines translate and transform sound waves into vibrations that the Cochlea can feel directly so that the eardrum never gets involved [8].

Object recognition is a classic computer vision problem which is concerned with deciding if an image contains a specific object of interest [9]. We want to discuss the possibility of using the sense of hearing to perceive visual items in our research. The sense of vision and the sense of hearing have a remarkable similarity: it is possible to identify both visual objects and audio. Humans are able to recognize a sound source's spatial position by hearing it with two ears. We developed a real-time object detection system with the objective of informing the user of the surrounding object. Some existing technologies are capable to detect movement and obstacles, but blind people still face significant problems in recognizing their environment and components. The identification and location of objects is the need for blind people's mobility through travel assistance and navigation assistance. For this reason, we developed an effective algorithm to identify and locate objects for fast and robust computer vision [10].

2 Problem Statement

Human echolocation has been practiced for years by many blind persons. This technique also causes fatigueness which affects the accuracy of echolocation. Furthermore, there is also a need of a smart objection detection system that can help the blind persons in detecting and recognizing the objects which he navigates through using the same device during echolocation.

3 Methodology

Figure 1 shows the flowchart of our implemented algorithm. First of all the smartphone camera is initialized and the video streaming starts. We use the app named "IP WebCam" to capture video frames which are sent to the computer through an IP address. During the video streaming, smartphone camera detects objects. By using Python programming language, the IP address of the camera is called in the code.

The frames detection is done by the function cv2.VideoCapture() which reads the incoming frames. The .read() method is a blocking operation, therefore, the main thread of our Python script is completely blocked until the frame is read from the camera device and returned to the script. We can improve our FPS (frames per second) simply by creating a new thread that does nothing but captures new frames while our main thread handles processing the current frame. The smartphone camera keeps on taking images of its environment and the images of object are then converted into grayscale and binary within the code so that it may form a matrix for the detected object and may be read by the computer easily. The pre-processing stage involves the reading and resizing of the image which is then checked by the inception V3 which is a 42-layer deep learning network [11]. Subsequently, the recognize object is converted to speech in the form of an mp3 audio file. We have used following libraries with Python:

- Tensorflow-gpu 1.0,
- OpenCv 3.2
- Numpy 1.12
- GTTS 1.1
- Pygame 1.9

Every time we face a complex system, we have to divide it into simpler sub-models. In our case, we have the following sub-systems.

- Image Classification
- Text to Speech

For image classification with Inception-V3 model, we use Tensorflow deep learning library. First, we download the deep learning model and create a node lookup class to get results in understandable form. We define a NodeLookup class that processes the labels and returns a string for each classification. The function creates Tensorflow graphs in memory. We use Google TTS (Text-to-Speech) API through GTTS, and save each result in a look-up table so that we can get a better real time performance. To reproduce the size of the audio file, we used Pygame.

Figure 2 shows the complete flowchart of the proposed

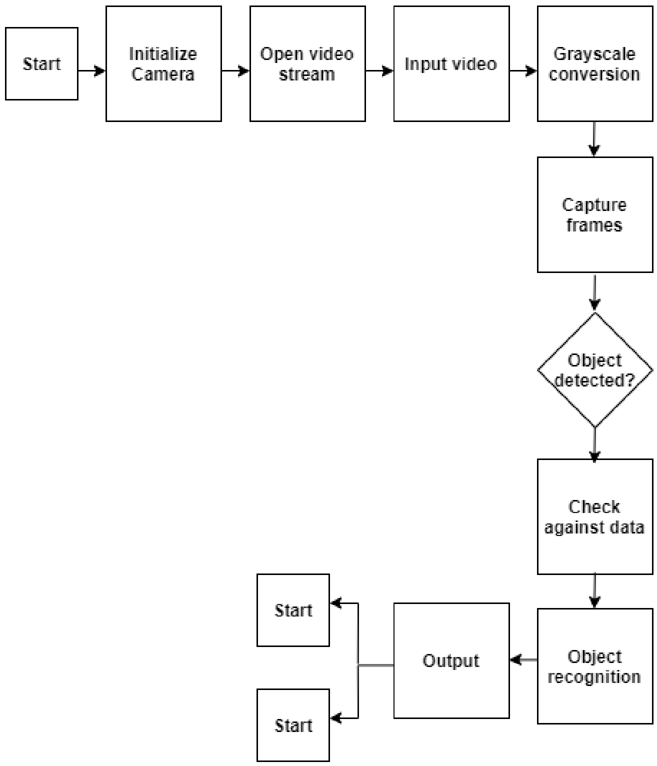


Fig. 1: Flowchart of algorithm of python programming

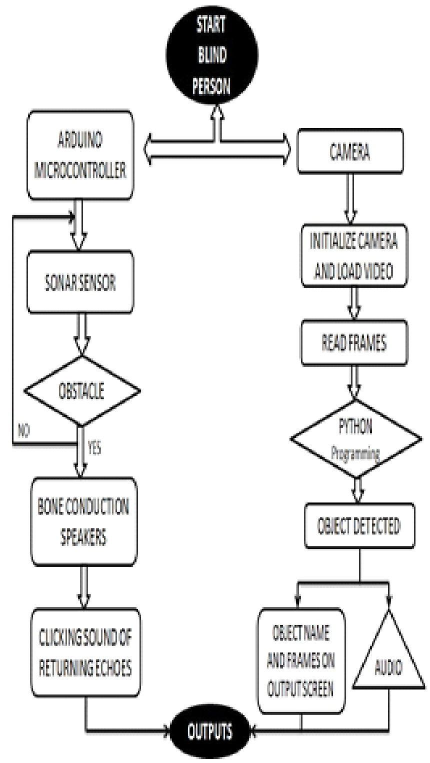


Fig. 2: Flowchart explaining the programming of device

technique. In this device, there are two functions that occur at the same time, making the flowchart parallel. On the left side, the micro-controller is connected to the sonar sensor and the sonar sends sound pulse of high frequency. When this pulse hits the obstacle, it reflects echoes of sound with the level of proximity of an object. These echoes enter the bone conduction speakers where they are converted into vibrations and bypass the normal airway through bone conduction. The vibrations directly enter cochlea where they provide the perception about an object's proximity and helps the blind person in echo locating. On the right side of Figure 2, the algorithm is for object detection in which objects are detected by the smartphone camera where the camera is connected through Wifi and we have called the IP of smartphone camera in our Python code. The image which is detected by the smartphone camera during live video streaming goes into the Python code where all the pre-processing in the code happens. After all the pre-processing, the image is converted into speech which further creates an mp3 file of speech. This speech is then sent to the blind person through Bluetooth air pods.

4 Results

For calibration, we tested LV-EZ MB1000 Ultra Sonar sensor at every 50 inch to check the range of the sensor

SNO.	Distance (inches)		Distance (cm)	
	Sonar	Tape	Sonar	Tape
1.	50	49.5	130	125
2.	100	98.03	254	249
3.	150	148.8	391	378
4.	200	198	510	505
5.	250	252.5	647	642

TABLE 1: Calibration of LV-EZ MB1000 ultra sonar sensor

and then also verified it through measuring tape to remove any linear error. It is concluded from Table 1 that the difference in distance from sensor on serial monitor and distance from measuring tape is approximately 2 cm (0.787 inch) which is a linear error that can be removed from coding algorithm. In Figure 3, graphical representation of readings that have been obtained from sonar sensor have been presented. On X-axis, we have taken time because time is an independent quantity. On Y-axis, we have taken distance. It can be seen that the graph is linear. From the graph,d we can conclude that the increasing distance will also result in the increase of time of returning echoes. Echolocation is a technique that could take months and even years to develop and to properly perceive and practice it in real life whether achieve it through human practice or through device. When we enter in testing phase due to

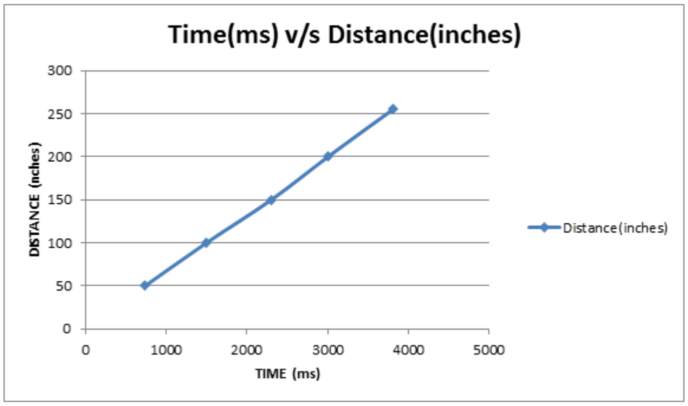
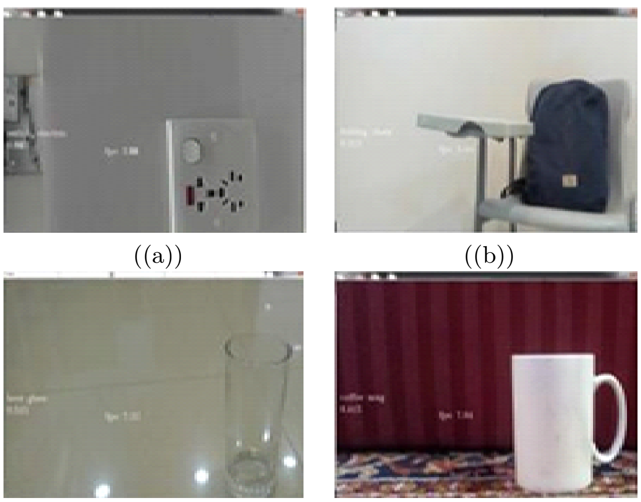


Fig. 3: Graphical representation of sonar readings obtained from serial monitor.



((a)) ((b)) ((c)) ((d))

Fig. 4: (a) Screenshot is attached of our algorithm detecting electric switch, (b) while detecting folding chair, (c) detecting beer glass, and (d) detecting coffee mug.

such a long commitment, we found only one suitable blind person (age = 14, gender = female). Up till now, five months have been completed of the rehabilitation sessions and the results that we have achieved are following:

- The 14 years old subject can easily navigate through the environment and can detect obstacle of what come in her way.
- For recognizing obstacles, our installed object detection system successfully recognizes objects and converts them into speech which is heard by the blind person through an earphone.

5 Discussion

Human echolocation has been practiced for many years by many blind persons which also causes fatigue, affecting the accuracy of echolocation as blind people cannot produce continuous clicking. Furthermore, there is also a need of smart objection detection system that can help blind persons in detecting and recognizing the objects through which they navigates using this device during echolocation.
The objectives achieved through our proposed system are as follows:

- A person can navigate through the environment using this device.
- Ability of echolocation is significantly enhanced.
- Integration of a smart device with the system allows detecting and recognizing objects with a smart phone camera in a live video stream and converting it into speech.

The better we train our code by putting different objects in front of the smartphone camera, the better results will be achieved during real-time object detection.
The features of our current study is that with the help of this device a blind person should be able to navigate and echolocate in the environment. It can help in training a blind person to echolocate. Through this smart object detection system, a blind person should be able to detect any object in a live video streaming through smartphone camera that is also attached with the device. The significance of the project is as follows.

- To enhance the mobility of blind people.
- Protecting blind people from possible hazard.
- Less costly product will be affordable for many people.

6 Conclusion

In this paper, we have proposed a smart phone based echolocation system for blind. The results presented in this paper show that our proposed system is able to detect objects efficiently in real time through the sensors readings acquired within an interval of every 30 seconds. However, the detection accuracy is constrained by the image quality and camera resolution. Our proposed system provides a great support to blind people becoming independent and increase their mobility with safety.

References

[1] "Vision Impairment and Blindness", Online: https://www.who.int/news-room/fact sheets/detail/blindness-and-visual-impairment, [accessed 2 October 2019].

[2] "Visioneers.Org-Our Organization", https://visioneers.org/our-organization, [accessed 2 October 2019].

[3] "Visioneers.Org - We Teach Blind People to See with SonarVision", https://visioneers.org, [accessed 2 October 2019].

[4] B. Schenkman and V. K. Gidla, "Human echolocation in different situations and rooms: Threshold values," presented at the 173rd Meeting of Acoustical Society of America and 8th Forum Acusticum, Boston, Massachusetts, pp. 015002, 2017.

[5] L. Thaler, R. De Vos, D. Kish, M. Antoniou, C. Baker, and M. Hornikx, "Human echolocators adjust loudness and number of clicks for detection of reflectors at various azimuth angles," Proc. R. Soc. B Biol. Sci., Vol. 285, pp. 20172735, 2018.

[6] A. G. Hagargund, S. V. Thota, M. Bera, and E. Fatima, "Image to Speech Conversion for Visually Impaired," Vol. 03, No. 06, pp. 7, 2017.

[7] Norman, L. J., and L. Thaler. "Human echolocation for target detection is more accurate with emissions containing higher spectral frequencies, and this is explained by echo intensity." Perception, Vol. 9, No. 3, 2018.

[8] "Are bone conduction headphones good enough yet? - The Verge.", Online: https://www.theverge.com/2016/10/24/13383616/bone-conduction-headphones-best-pair-aftershokz. [Accessed: 02-Oct-2019].

[9] H. Jabnoun, F. Benzarti, and H. Amiri, "Object detection and identification for blind people in video scene," In 15th International Conference on Intelligent Systems Design and Applications (ISDA), Marrakech, pp. 363–367, 2015.

[10] K. G. Krishnan, C. M. Porkodi, and K. Kanimozhi, "Image recognition for visually impaired people by sound," In International Conference on Communication and Signal Processing, Melmaruvathur, India, pp. 943– 946, 2013.

[11] S.-H. Tsang, "Review: Inception-v3 — 1st Runner Up (Image Classification) in ILSVRC 2015," Medium, 23-Mar 2019, Online: http://medium.com/@sh.tsang/review-inception-v3-1st-runner-up-image-classification-in-ilsvrc-2015-17915421f77c. [Accessed: 21-Jan-2020].