International Journal of Scientific Research in Science, Engineering and Technology
Print ISSN: 2395-1990 | Online ISSN : 2394-4099 (www.ijsrset.com)
doi : https://doi.org/10.32628/IJSRSET20765

A Survey on Visionary Eye for Visual Impairment

Ruchir Shah[1], Dhaval Tamboli[1], Ajay Makwana[1], Ravindra Baria[1], Kishori Shekokar[2], Dr. Sheshang Degadwala[3]

[1]UG Scholar, Sigma Institute of Engineering, Department of Computer Engineering, Sigma Institute of Engineering, Vadodara, Gujarat, India

[2]Assistant Professor, Sigma Institute of Engineering, Department of Computer Engineering, Sigma Institute of Engineering, Vadodara, Gujarat, India

[3]Associate Professor, Sigma Institute of Engineering, Department of Computer Engineering, Sigma Institute of Engineering, Vadodara, Gujarat, India

Article Info
Volume 7 Issue 6
Page Number: 33-39
Publication Issue :
November-December-2020

Article History
Accepted : 01 Nov 2020
Published : 10 Nov 2020

ABSTRACT

In this survey paper, we have discussed a proposed system that can be a visionary eye for a blind person. A common goal in computer vision research is to build machines that can replicate the human vision system. For example, to recognize and describe objects/scenes. People who are blind to overcome their real daily visual challenges. To develop a machine that can work by the vocal and graphical assistive answer. A machine can work on voice assistant and take the image taken by a person and after an image processing and extract the result after neural networks.

Keywords: Computer Vision, Navigation, Artificial Intelligent, Human-Computer Interaction, Visual Impairment, Text to Speech, Speech to Text.

I. INTRODUCTION

There are more than 50 million people live with visual impairment (blindness). They cannot do small daily tasks such as reading newspapers, cannot watch a movie, read the information on daily transportation, they face difficulty in the public area, looking in the mirror is among the daily challenges that they face on their routine. They have to ask questions about their surroundings to people. This makes them depend on other people.

A common aim in our research is to build methods and machines that can replicate the human vision with more advanced attributes and quality. Like, to recognize and describe objects and scenes. Towards this aim, we introduce our project on developing algorithms for assistive technologies with help of artificial intelligence and computer vision. In particular, we hope this work will guide more people and realize them the technological needs of people who are blind while developing assistive technologies that eliminate their accessibility barriers.

II. PROPOSED SYSTEM

One of the biggest challenges for blind people is when they are moving outdoors where there are uneven terrains, obstacles such as cars, rocks, and holes and not knowing where to go without the ability to see. Traditionally, blind people use walking sticks or guide dogs to help them move around and they are limited to move only around familiar places. Navigation is the science (or art) of directing the course of a mobile robot as it traverses the environment. Inherent in any navigation scheme is the desire to reach a destination without getting lost or crashing into another object. Put simply, the navigation problem is to find a path from start to goal, which meets the task constraints, and to traverse that path without collision.

A method performed by a mobile device, for assisting blind or visually impaired users to navigate a room or a new and unfamiliar environment. The method includes blind users acquiring one or more images using a device and invoking algorithms. Processing algorithms include one of view from motion, whereby algorithms construct a 3D representation of the environment being imaged. Further algorithms are applied to identify and assign attributes to objects in the imaged environment. The environment is presented to the user via a device screen, enabling the user to virtually explore the environment using text-to-speech and vice versa. Here the proposed system will ask the user what output he/she want to focus on. And then the proposed system will process and give the desired result via speech.

These type of proposed systems follows on 2 aims either navigation or object detection. We have tried to make things in one with the help of our proposed systems a blind person not only can navigate around but also find useful information from objects he needs.

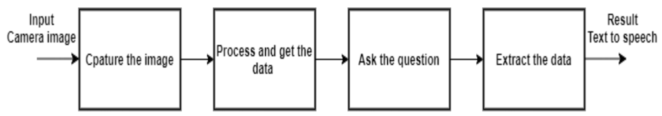


Figure 1.0

A. **Capture The Image**

The user first opens a camera and takes a picture of an object to get the information. Then the process starts as follows.

1. The picture/image is taken.
2. The proposed system will check that if the image is taken properly or not i.e. blur, the object is not visible, etc.
3. If the image taken by the user is not in proper format then, the proposed system will tell the user to take a picture again.
4. Removes the noise in the image.
5. Process the picture in the proper frame and corrects the color.

B. **User Ask Question**

The user asks the question regarding the information that the user needs clearly and formally.

C. **Processing Data**

As per users question the process starts to defragmenting the words in speech given by user. Proposed system will focus on words where it can get more information. i.e. “what”, “where”, “price”, “calories”, “amount”, “who”, “bill”, “is it”, “name”, “year”, “date”, “time”, “Weather”, “address”, “website”, “link”, “weight”, “direction”, etc

D. **Extracting The Data**

The next step is to extract the data as per the user's needs. Which is done by Computer Vision and character reorganization.

E. **Result**

The result will be given after processing the image and question by the user. The result will be given to the user in an audio manner using text to speech approach.

III. CATEGORIZATION

The classification as well as object localization makes it one of the most challenging topics in the domain of computer vision. In simple words, the goal of this detection technique is to determine where objects are located in a given image called object localization and which category each object belongs to, which is called object classification. detecting instances of semantic objects of a certain class (such as humans, buildings, or vehicles) in digital videos and images. Object detection has proved to be a prominent module for numerous important applications like video surveillance, autonomous driving, face detection, etc. Feature detectors such as Scale Invariant Feature Transform and Speeded Up Robust Feature are good methods that yield high-quality features but are too computationally intensive for use in real-time applications of any complexity. Based on the normalized corner information, support vector machine and back-propagation neural network training are performed for the efficient recognition of objects.

Here is a flowchart that shows the entire process in an abstract way.

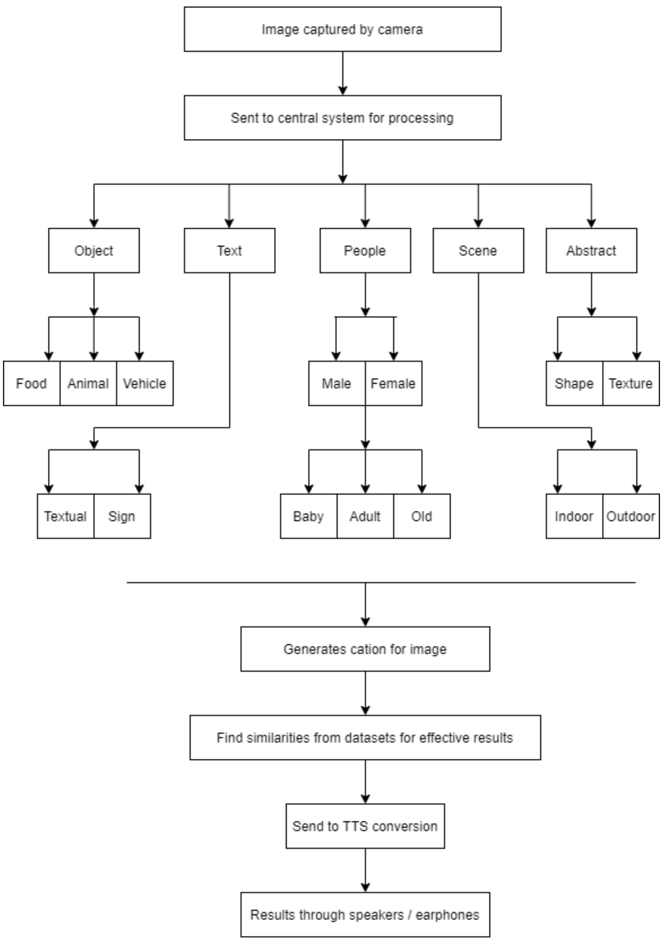
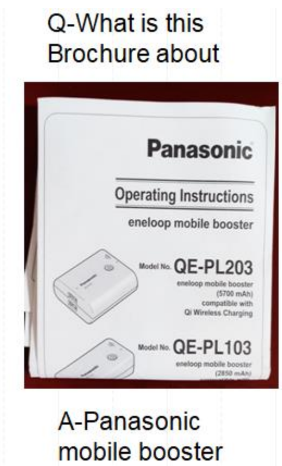


Fig – 2.0 Flowchart

It is a flowchart illustrating the method of operation of the proposed system for the visually impaired. According to the present device, The camera captures the images of the objects which are present in the surroundings of the user. The images which are captured by the camera are automatically sent to the processor through Wi-Fi, which analyses the objects present in the images, identifies them, and gathers all the information about the captured image in collaboration with the different layers of processing. The information about the objects in the captured image that are processed by the interface is converted into the audio file. The audio output is sent to the user which can be heard through the earphones or speakers.

IV. ALGORITHMS AND TECHNIQUES

The modern world is enclosed with gigantic masses of digital visual information. To analyze and organize this devastating ocean of visual information image analysis techniques are major requisites. In particular, use would be methods that could automatically analyze the semantic contents of images or videos. The content of the image determines the significance in most of the potential uses. One important aspect of image content is the objects in the image. So there is a need for object recognition techniques.



Object recognition is an important task in image processing and computer vision. It is concerned with determining the identity of an object being observed in an image from a set of known tags. Humans can recognize any object in the real world easily without any efforts; on the contrary machines by itself cannot recognize objects. Algorithmic descriptions of recognition tasks are implemented on machines; which is an intricate task. Thus object recognition techniques need to be developed which are less complex and efficient.

Algorithms	Speed	mAP
R-CNN	Slow	66.00%
Fast R-CNN	Medium	66.90%
Faster R-CNN	Fast	69.00%
SSD	Fast	71.00%
YOLO	Fast	76.80%

Table 1.0

The different algorithms are explained for suitable for different types of systems and scenarios.

Many successful approaches that address the problem of general object detection use a representation of the image objects by a collection of local descriptors of the image content. Global features provide better recognition. Color and shape features can also be used. Various object recognition techniques are presented in this paper. Difficulties may arise during the process of object recognition; a range of such difficulties are discussed in this paper.

1) R-CNN

It stands for "Region-based Convolutional Neural Networks" and also known as R-CNN The Region-based Convolutional Network method (RCNN) is a combination of region proposals with Convolution Neural Networks (CNNs). R-CNN helps in localizing objects with a deep network and training a high-capacity model with only a small quantity of annotated detection data. It achieves excellent object detection accuracy by using a deep ConvNet to classify object proposals. R-CNN can scale to thousands of object classes without resorting to approximate techniques, including hashing.

2) Fast R-CNN

Written in Python and C++ (Caffe), the Fast Region-based Convolutional Network method or Fast R-CNN is a training algorithm for object detection. This algorithm mainly fixes the disadvantages of R-CNN and SPPnet, while improving their speed and accuracy. The advantages of Fast R-CNN are that it gives higher detection quality(accuracy) (mAP) SPPnet. Training is single-stage, using a multi-task loss Training can update all network layers. No disk storage is required for feature caching.

3) Faster R-CNN

Faster R-CNN is an object detection algorithm that is similar to R-CNN. This algorithm utilizes the Region Proposal Network (RPN) that shares full-image convolutional features with the detection network in a cost-effective manner than R-CNN and Fast R-CNN. A Region Proposal Network is a fully convolutional network that simultaneously predicts the object bounds as well as objectness scores at each position of the object and is trained end-to-end to generate high-quality region proposals, which are then used by Fast R-CNN for detection of objects.

4) SSD

The SSD Stands for "Single Shot Detector". As per the name suggest Single Shot Detector (SSD) is a method for detecting objects in images using a single deep neural network. The SSD approach discretizes the output space of bounding boxes into a set of default boxes over different aspect ratios. After discretizing, the method scales per feature map location. The Single Shot Detector network combines predictions from multiple feature maps with different resolutions to naturally handle objects of various sizes. The advantage of SSD is that SSD eliminates proposal generation and subsequent pixel or feature resampling stages and encapsulates all computation in a single network. Easy to train and straightforward to integrate into proposed systems that require a detection component. SSD has competitive accuracy to methods that utilize an additional object proposal step, and it is much faster while providing a unified framework for both training and inference.

5) YOLO

YOLO is a new and novel approach to object detection. It stands for "You Only Look Once". Prior work on object detection repurposes classifiers to perform detection. YOLO frames object detection as a regression problem to spatially separate bounding boxes and associated class probabilities. A single neural network predicts bounding boxes and class probabilities directly from full images in one evaluation. Since the whole detection pipeline is a single network, it can be optimized end-to-end directly on detection performance. Unlike sliding window and region proposal-based techniques, YOLO sees the entire image during training and test time so it implicitly encodes contextual information about classes as well as their appearance. Fast R-CNN, a top detection method, mistakes background patches in an image for objects because it cannot see the larger context. YOLO makes less than half the number of background errors compared to Fast R-CNN.

V. NAVIGATION

Geographical orientation is a far space activity that determines the location in the geographical space of the travel task such as how to get from point A to point B. This is called "Navigation". Navigation is a purposeful process which comprises the traveler updating their orientation and position and can accomplish in three ways

1. Position based navigation
2. Velocity based navigation and
3. Acceleration based navigation.

In Position based navigation, the user updates on external signs within an environment. Position-based navigation depends on landmark-based navigation and map-based navigation where landmark-based navigation is sometimes considered as a sub-part of map-based navigation. Map aided navigation is to get the benefit from the prior information contained in maps or building plans. Landmark-based methods rely on the detection of landmarks. A landmark is a feature in the environment that can be detected by sensory data. If landmarks are detected, they are matched against a priori information of the environment. Classifications of landmarks are as artificial vs natural. Color markers, BLE (Bluetooth Low Energy) beacons attached to the places with modern interior designs such as Airports, Hospitals and considered as artificial landmarks, and doors, windows, or ceiling lights in indoor environments are considered as natural landmarks. Velocity based navigation relies on exterior and interior signals. Dead Reckoning (DR) is an example of velocity-based navigation, in which one advanced a current position by using path progression and speed from the known previous position and predict future position. Acceleration based navigation is also called inertial navigation.

Indoor Pointing Techniques	Accuracy	Range	Power Consumption	Real-Time Location
Wi-Fi	5-15m	<150 m	High	Yes
BLE	1-3m	<30m	Low	No (Delay)
NFU	10cm	20cm	Low	Yes

	(close range)	or less		
UWB	5-10cm (very close)	Few	Low	Yes

Visually impaired people to navigate between indoors and outdoors. The most common sensor-based method for outdoor navigation relies on GPS, but it is not precise enough for the purposes of the present research, especially in indoor environments. For example, a proposed system for supporting wayfinding by the visually impaired was found to be critically limited by a lack of adequate data for both indoor and outdoor environments. Another recent solution BlindNavi and NavCo3 use beacons in combination with a smartphone app. However, beacon localization also faces some constraints. First, complete route navigation necessitates the deployment of many beacons, which carries considerable installation and maintenance costs. Second, a signal collision may occur if the distances between beacons are too short, and result in incorrect location information. Lastly, this approach cannot easily obtain orientation information or detect dynamic obstacles at all.

VI. CONCLUSION

In an exemplary way, a method, performed by a mobile device includes acquiring an image of the environment, acquiring information about mobile device spatial orientation, and its relative distance to the objects in the environment, wherein the mobile device comprises one or more of phone or tablet with an embedded camera, accelerometers, speakers. The method can further include providing image and sensor data acquisition in response to a user request, processing data to present a plan of the image mapped to the device, and providing a verbal description of the objects in the image responsive to the user. But to navigate indoor and outdoor locations there are more parameters that are needed to be taken care of.

VII. ACKNOWLEDGEMENT

We are sincerely thankful to Prof. Kishori Shekokar Ma'am for her precious guidance and suggestion for our project work Visionary Eye for Visual Impairment.

VIII. REFERENCES

[1]. Q. Zhou, L. Ma, M. Celenk, and D. Chelberg, "Object Detection and Recognition via Deformable Illumination and Deformable Shape," 2006 International Conference on Image Processing, Atlanta, GA, 2006, pp. 2737-2740, doi: 10.1109/ICIP.2006.313113.

[2]. F. Puente-Mansilla, G. Boza-Quispe, G. Lapa-Velasquez, C. Matos-Avalos, and J. Rosales-Huamani, "A wearable UV sensor and accessible smartphone application for blind people," 2016 IEEE International Symposium on Consumer Electronics (ISCE), Sao Paulo, 2016, pp. 81-82, doi: 10.1109/ISCE.2016.7797382.

[3]. J. Intrieri and K. Healy, "Combining Lidar And Radar Measurements To Derive Cirrus Cloud Effective Radii: In Situ Comparison And Simplistic Model Results," Proceedings of IEEE Topical Symposium on Combined Optical, Microwave, Earth and Atmosphere Sensing, Albuquerque, NM, USA, 1993, pp. 119-122, doi: 10.1109/COMEAS.1993.700199.

[4]. R. A. Minhas and A. Javed, "X-EYE: A Bio-smart Secure Navigation Framework for Visually Impaired People," 2018 International Conference on Signal Processing and Information Security (ICSPIS), DUBAI, United Arab Emirates, 2018, pp. 1-4, doi: 10.1109/CSPIS.2018.8642718.

[5]. H. Gada, V. Gokani, A. Kashyap, and A. A. Deshmukh, "Object Recognition for the Visually Impaired," 2019 International Conference on Nascent Technologies in Engineering (ICNTE), Navi Mumbai, India, 2019, pp. 1-5, doi: 10.1109/ICNTE44896.2019.8946015.

[6]. C. Cerrada, S. Salamanca, E. Perez, J. A. Cerrada and I. Abad, "Fusion of 3D Vision Techniques and RFID Technology for Object Recognition in Complex Scenes," 2007 IEEE International Symposium on Intelligent Signal Processing, Alcala de Henares, 2007, pp. 1-6, doi: 10.1109/WISP.2007.4447604.

[7]. A. Dionisi, E. Sardini, and M. Serpelloni, "Wearable object detection system for the blind," 2012 IEEE International Instrumentation and Measurement Technology Conference Proceedings, Graz, 2012, pp. 1255-1258, doi: 10.1109/I2MTC.2012.6229180.

[8]. Sisidiya, Suyog. (2020). Enhancing Vision for Visually Impaired. International Journal for Research in Applied Science and Engineering Technology. 8. 668-672. 10.22214/ijraset.2020.5104.

[9]. A. Zeb, S. Ullah, and I. Rabbi, "Indoor vision-based auditory assistance for blind people in semi-controlled environments," 2014 4th International Conference on Image Processing Theory, Tools and Applications (IPTA), Paris, 2014, pp. 1-6, doi: 10.1109/IPTA.2014.7001996.

[10]. B. Rajapandian, V. Harini, D. Raksha, and V. Sangeetha, "A novel approach as an AID for blind, deaf and dumb people," 2017 Third International Conference on Sensing, Signal Processing and Security (ICSSS), Chennai, 2017, pp. 403-408, doi: 10.1109/SSPS.2017.8071628.

[11]. M. Nassih, I. Cherradi, Y. Maghous, B. Ouriaghli and Y. Salih-Alj, "Obstacles Recognition System for the Blind People Using RFID," 2012 Sixth International Conference on Next Generation Mobile Applications, Services and Technologies, Paris, 2012, pp. 60-63, doi: 10.1109/NGMAST.2012.28.

[12]. B. J. Koskowich, M. Rahnemoonfai, and M. Starek, "Virtual — A Framework Enabling Real-Time Coordinate Transformation & Occlusion Sensitive Tracking Using UAS Products, Deep Learning Object Detection & Traditional Object Tracking Techniques," IGARSS 2018 - 2018 IEEE International Geoscience and Remote Sensing Symposium, Valencia, 2018, pp. 6416-6419, doi: 10.1109/IGARSS.2018.8518124.

[13]. A. A. Nada, M. A. Fakhr, and A. F. Seddik, "Assistive infrared sensor-based smart stick for blind people," 2015 Science and Information Conference (SAI), London, 2015, pp. 1149-1154, doi: 10.1109/SAI.2015.7237289.

[14]. V. Kunta, C. Tuniki and U. Sairam, "Multi-Functional Blind Stick for Visually Impaired People," 2020 5th International Conference on Communication and Electronics Systems (ICCES), COIMBATORE, India, 2020, pp. 895-899, doi: 10.1109/ICCES48766.2020.9137870.

[15]. S. Kumar KN, R. Sathish, S. Vinayak, and T. Parasad Pandit, "Braille Assistance System for Visually Impaired, Blind & Deaf-Mute people in Indoor & Outdoor Application," 2019 4th International Conference on Recent Trends on Electronics, Information, Communication & Technology (RTEICT), Bangalore, India, 2019, doi: 10.1109/RTEICT46194.2019.9016765.

[16]. David A. Forsyth, Jean Ponce, "Computer Vision: A Modern Approach", Publisher: Prentice Hall Professional Technical Reference, ISBN:978-0-13-085198-7, March 2002

Cite this article as :

Ruchir Shah, Dhaval Tamboli, Ajay Makwana, Ravindra Baria, Kishori Shekokar, Dr. Sheshang Degadwala, "A Survey on Visionary Eye for Visual Impairment", International Journal of Scientific Research in Science, Engineering and Technology (IJSRSET), Online ISSN : 2394-4099, Print ISSN : 2395-1990, Volume 7 Issue 6, pp. 33-39, November-December 2020. Available at doi : https://doi.org/10.32628/IJSRSET20765
Journal URL : http://ijsrset.com/IJSRSET20765