IJCSE International Journal of Computer Sciences and Engineering Open Access
Research Paper Vol.-8, Issue-2, Feb 2020 E-ISSN: 2347-2693

An Overview on Object Detection and Recognition

N. Raviteja[1*], M. Lavanya[2], S. Sangeetha[3]

[1,2,3]Information Science & Engineering, Global Academy of Technology, Visvesvaraya Technological University, Bangalore, Karnataka, India

[*]*Corresponding Author: ravie.teja.80@gmail.com Tel.: +91 7022074577*

DOI: https://doi.org/10.26438/ijcse/v8i2.4245 | Available online at: www.ijcseonline.org

Accepted: 12/Feb/2020, Published: 28/Feb/2020

***Abstract*—** The main intent of enhancing the present system of object detection is to enhance its productivity and improve the detection accuracy. The method initializes a model in a way to store the background pixels and to compare each frame pixels with the same model background pixels. The intention is to separate the background pixels with the foreground moving pixels to improve the ability to detect the object in the foreground. An improved object detection algorithm based on "Gaussian mixture model" and "three-frame difference method". This three-frame difference method uses a dynamic segmentation threshold and edge detection technology to solve the problem of illumination mutation and discontinuity of target edge.

***Keywords*—** Gaussian mixture model, illumination mutation, object detection, pixels, three-frame difference method, dynamic segmentation, edge detection technology.

I. INTRODUCTION

Throughout the world, there are about 314 million people, who are visually impaired. Out of these, approximately 45 million are blind. Most of the people who have vision impairment live in developing countries.

It is also estimated that more than 2.2 billion people have a physical impairment, out of whom more than 1 billion are visually impaired that could be prevented or has yet to be addressed.

These 1 billion people are those suffering from far-sightedness or blindness due to un-addressed refractive error of about 123.7 million, cataract of about 65.2 million, glaucoma of about 6.9 million, corneal opacities of about 4.2 million diabetic retinopathy of about 3 million, and trachoma of about 2 million, as well as short-sightedness caused by un-addressed presbyopia of about 826 million.

In terms of regional differences, the prevalence of distance vision impairment in high income regions is estimated to be four times more than low- and middle-income regions. With regards to near vision, rates of un-addressed near short-sightedness are estimated to be greater than 80% in western, eastern and central sub-Saharan Africa, while comparative rates in high income regions of North America, Australasia, Western Europe, and of Asia-Pacific have also reported to be lower than 10%.
Population growth and aging will increase the risk that more people acquire vision impairment [2].

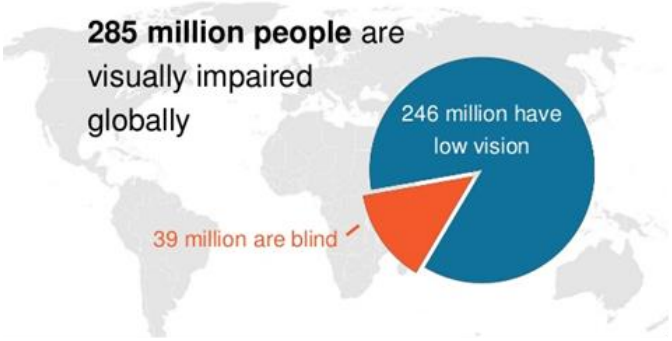


Figure 1. Global blind statistics

II. RELATED WORK

The main strategy implemented here is to differentiate still background and the moving foreground system. Consider an example of a still staircase and a moving escalator along with is direction. That is, when a person walks, he or she concerns about moving people or moving objects towards him or her rather than detecting and identifying what is situated around. An obstacle-free path can change anytime on the arrival of any object. Therefore, the information about the moving objects and their directions with respect to the blind persons perspective is more important than information about all the objects in the entire scene [3].

Proposed by Aniqua Nusrat Zereen et.al., [4], The approach of detection of moving objects (Object moving direction identification using depth image for blind) is a simple and

realistic blind navigation support system. RGB camera and depth sensor of Microsoft Kinect is used for image acquisition. To detect any moving object irrespective of the direction, the entire front scene image is divided into four vertical line numbered one to four from right to left then capturing 30 frames per second, the video clip from a fixed position using Kinect three consecutives depth frames (0th,15th and 30th frame) are extracted.

With the introduction of YOLO (You Only Look Once) and SSD, CNN based object detection as attracted attention because it achieves high accuracy. CNN is a basic structure in which a convolution operation of performing tasks on multiple layers. Most of the CNN calculation are product sum operations. For detecting small objects an image split method is used. The split-image is acquired from the camera to obtain multiple images and performing CNN inference on each image. The splitting mainly depends on the throughput. As the Throughput decreases, the splitting increases [4].

Described by Israh Akbar, et.al., [5] the research methodology used was an ethnographic study design. Ethnographic research is a qualitative method where researchers observe or interact with study's participants in real life environment. The aim of this study is to get to know the reality of design problem which was identified in the survey to find out the usage of GPS technology among the blind community. This study was suitable for this research where the researcher is fully involved with participants and facilitates the researchers for documentation of the findings. As described by the authors of "Intelligent eye: A mobile application for assisting blind people" [6].

Light detection: Embedded light sensors allows the blind people to capture light accuracy in mobile and convert those values in spoken description with sound beep which vary depending on the light intensity. Color detection approach mainly works based on the images that are taken by the back camera of smart phones. Mobile-phones have become invaluable devices frequently used to meet the user's needs according to their requirements. Currently, Android (Google) and iPhone (Apple) operating systems embed Color Vision Deficiency (CVD) functions for supporting color-blind people. However, usability problems of complexity and flexibility still exist in the User Interfaces (UIs) of many Apps. Among the others, Color-Find from Golden-Shores technologies LLC provides an interface characterized by the only camera view together with the quit button and a simple line showing as output the color name of the Object at the center of the camera view. The ideology as described by Saransh Sharma, et.al., [7] HAAR Cascade classifiers and LBPH recognizer has been used for face detection and recognition respectively. Various hand gestures and human faces can be detected and identified using this system.
Incrementally, the proposed architecture deals with hand gesture recognition and facial recognition.

Facial identification LBP, abbreviation for Local Binary Pattern, is the measure of fixed grey scale texture obtained by thresholding the nearby area of each pixel and the result obtained is considered as a binary number. The LBP operator is the ratios of the pixel [7].

As implemented by Marcin Iwanowski at.al.[11], a mobile application was introduced which had the following characteristics. Speech-Recognizer class ensures access to the Google Speech Recognition module [13] and is also responsible for data transfer between the application and the server. In turn, the Recognition-Listener allows tracking processes running in the Speech-Recognizer and getting the response vector or the error number. Over a dozen error types have been included, Eg., the lack or disruption of the internet connection, lack of understanding of commands, no answer from the server, or no results. Through the Text-To-Speech class, the sequence of characters qualified for reading is forwarded to the google speech synthesizer [7].

Furthermore, this class makes it possible to monitor the synthesis as well as suspending the launching of other functions until the current content has resounded [9]. The appropriate queuing of announcements has turned out necessary because autonomous processes continually wait for notifications about new events (Eg., about a new message), and immediate voice statements could disrupt other processes, particularly speech recognition. Under such circumstances, the announcements accompanying currently executed functions get always a higher priority while the others go to a special buffer. The program checks processes cyclically and at the right moment it reads the content gathered in the buffer [8].

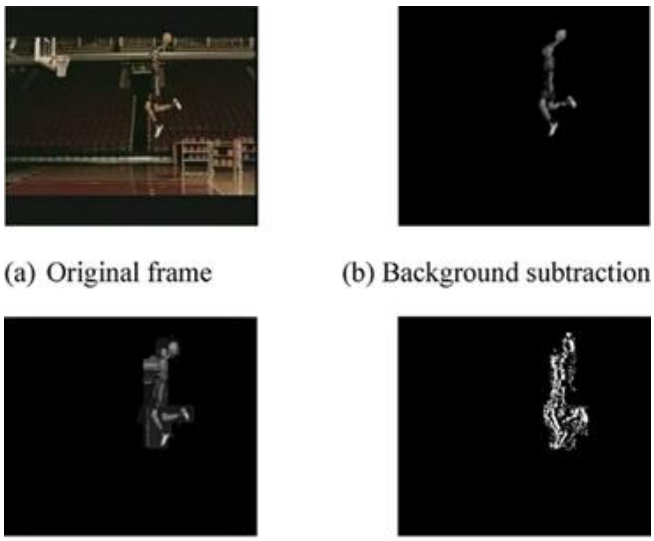
III. THEORETICAL MODEL

Background subtraction methods typically require a scene model to compare and segment the regions of the frames as a background or foreground. In the meantime, every model needs a size and initialization procedure which significantly has an effect on speed and accuracy of the method. Higher amount of sample makes the method more accurate, but it also causes the method to be slower and take longer to start segmentation. Therefore, if the system runs on a real-time the size of model should be chosen based on the hardware [12].

The proposed method for full system initialization requires N (N=20 in our experiments) number of frames to produce the complete model. However, the system starts segmentation when it passes Min-Frame (Min-Frame=3) which is the minimum number of required samples to identify foreground.

Most of other methods require a high number of frame and time for initialization. However, in comparison to the other methods, our approach can complete initialization at a faster

time. This will be carried out by blindly storing the first N number of depth and colour frames in to the model and then with the aid of the updated stage the model will gradually improve. This process will help the system to start rapid detection of moving object.



(a) Original frame (b) Background subtraction

(c) Frame difference (d) Optical flow

Figure 2. Different Frame Specification

IV. RESULTS AND DISCUSSIONS

As discussed here, we obtain the key details of a moving object detection using adaptive blind updating is proposed. The main contribution of this paper is adding adaptive blind update method, more complex segmentation, proposed bootstrapping detection and segmentation, Face Recognition System for Blind People," 2019 6th International Conference on Signal Processing and Integrated Networks (SPIN), Noida, India, 2019, pp. 534-539. doi: 10.1109/SPIN.2019.8711706

[8] proposed shadow detection method based on L*a*b* colour space and complete evaluation of the proposed method [9]. By tracking the moving object, the frequently of blind update will be changed according to the speed of moving object. This strategy will help the scene model to adapt to the complex scenarios such as environment changes, illumination changes and shadows as well as detecting the fast, slow and stationary objects while reducing the ghost phenomenon. A simple tracking method has been used in this work to reduce the computational costs. It also produces details or a brief overview on the different novel methods on object detection and classification. Light is imparted on moving objects and how accurately their details could be detected [13].

V. CONCLUSION AND FUTURE SCOPE

The proposed methods here [3][4][5], can significantly improve the accuracy of the algorithm when stationary objects are existing. On the other hand, on fast moving sequences, the ideology achieved a slightly improved or equal performance to the old methods. In general, the main advantages of the proposed methods are to improve segmentation accuracy in stationary moving objects, bootstrapping, shadow and depth camouflage scenarios. Overall the methodologies are more consistent in all the scenarios. The main disadvantage is that if the system is implemented, it might have a higher computational cost compared to the original method (static-distinguish approach). However, this method can be applied in live applications. The best performance of this method is only achievable when one moving object exists. In the future, more complex decision-making system can be developed to deal with more than one moving objects. Implementing this with a lower cost might also improvise the economical approach will be a major future enhancement.

REFERENCES

[1] Chen, Liang-Bi & Su, Jian-Ping & Chen, Ming-Che & Chang, Wan-Jung & Yang, Ching-Hsiang & Sie, Cheng-You, "An Implementation of an Intelligent Assistance System for Visually Impaired/Blind People", IEEE Access, **12-Apr-2019.**

[2] Mahajan J.R., C.S. Rawat, "Object Detection and Tracking using Cognitive Approach", International Journal of Scientific Research in Network Security and Communication, ISSN**: 2321-3256,** Volume:**5,** Issue.**3,** pp.**136-140, Jun-217**

[3] A. N. Zereen and S. Corraya, "Detecting real time object along with the moving direction for visually impaired people", International Conference on Electrical, Computer & Telecommunication Engineering (ICECTE), pp**.1-4**, **Dec-2016.**

[4] A. Jinguji, Y. Sada and H. Nakahara, "Real-Time Multi-Pedestrian Detection in Surveillance Camera using FPGA", International Conference on Field Programmable Logic and Applications (FPL), Barcelona, Spain, pp**.424-425, 2019**.

[5] I. Akbar and A. F. Misman, "Research on Semantics Used in GPS Based Mobile Phone Applications for Blind Pedestrian Navigation in an Outdoor Environment", International Conference on Information and Communication Technology for the Muslim World (ICT4M), pp**. 196-2018.**

[6] M. Awad, J. E. Haddad, E. Khneisser, T. Mahmoud, E. Yaacoub and M. Malli, "Intelligent eye: A mobile application for assisting blind people", IEEE Middle East and North Africa Communications Conference (MENACOMM), pp.**1-6**, **2018.**

[7] S. Sharma, S. Jain and Khushboo, and Kulyukin, Vladimir & Nicholson, John & Ross, David & Marston, James & Gaunet, Florence, "The Blind Leading the Blind: Toward Collaborative Online Route Information Management by Individuals with Visual Impairments", . International conference on Social Information Processing, pp.**54-59, March 26-28, 2008.**

[9] Yu, Tianming & Yang, Jianhua & Lu, Wei, "Combining Background Subtraction and Convolutional Neural Network for Anomaly Detection in Pumping-Unit Surveillance".

[10] Baygin, Mehmet & Karakose, Mehmet & Sarimaden, Alisan & Akin, Erhan, "An Image Processing based Object Counting Approach for Machine Vision Application", International

Conference on Advances and Innovations in Engineering (ICAIE),**16-Feb-2018.**

[11] S. L. Joseph, X. Zhang, I. Dryanovski, J. Xiao, C. Yi and Y. Tian, "Semantic Indoor Navigation with a Blind-User Oriented Augmented Reality", IEEE International Conference on Systems, Man, and Cybernetics, Manchester,**pp.3585-3591**, **2013** .

[12] C. K. Lakde and P. S. Prasad, "Navigation system for visually impaired people", International Conference on Computation of Power, Energy, Information and Communication (ICCPEIC), pp.**0093-0098, 2015 .**

[13] Md.T. . Akhtar, S.T. Razi, K.N. Jaman, A. Azimusshan, Md.A. Sohel, "Fast and Real Life Object Detection System Using Simple Webcam", International Journal of Scientific Research in Computer Sciences and Engineering, ISSN: **2320-7639,** Volume:**6**, Issue.**4,** pp.**18-23, Aug-2018.**

Authors Profile

Mr.Ravi Teja N pursuing BE in Global Academy of Technology in the Department of Information Science and Engineering Bangalore. His areas of interest are Image Processing, Deep Learning and Natural Language Processing.

Ms.Lavanya M pursuing BE in Global Academy of Technology in the Department of Information Science and Engineering Bangalore. Her areas of interest are Big Data, Machine Learning and Natural Language Processing.

Ms.Sangeetha S pursuing BE in Global Academy of Technology in the Department of Information Science and Engineering Bangalore. Her areas of interest are Image Processing, Machine Learning and Information Security.

