Real-Time Object Detection and Counting

Virag Patel and Priteshkumar Prajapati
Chandubhai S. Patel Institute of Technology (CSPIT),
Charotar University of Science and Technology (CHARUSAT),
Changa, Gujarat, India
19it114@charusat.edu.in, priteshprajapati.it@charusat.ac.in

***Abstract*—Humans can detect and recognize the objects present in an image. The human system is fast and accurate for performing a wide range of complex tasks like identifying multiple objects, detecting obstacles, with conscious thought. Due to the availability of large data sets, faster processing, and better approaches, computers can be trained easily to detect and classify multiple objects within an image or a video with a high percentage of accuracy.**

Object Detection has always been a more challenging task as it involves the combination of the two tasks, Image Classification, and Object Localisation, and drawing a bounding box around each object of interest in the image or the video and assigning them a label. Altogether, this problem is stated as Object Recognition.

Object detection systems always construct a model for an object class from a set of training examples. In the case of a fixed rigid object in an image, only one example may be needed, but more generally multiple training examples are necessary to capture certain aspects of class variability

Moving object detection has recently become a focus area of research in the field of computer vision. This is mainly because of its application in tracking objects, video surveillance, pedestrian detection, people counting, and self-driving cars.

***Index Terms*—Object Detection,Computer Vision**

I. Introduction

Digital image processing is one of the most researched fields nowadays. The ever-increasing need of surveillance systems has further on made this field the point of emphasis. Surveillance systems are now a days used for security reasons, intelligence gathering and many individual needs. Object tracking and detection is one of the main steps in these systems. Different techniques are used for this task and research is vastly done to make this system automated and to make it reliable. In the proposed system the background subtraction is done using the BackgroundSubtractor MOG2 algorithm which is a Gaussian Mixture-based mixture segmentation algorithm. One important feature of this algorithm is that it selects appropriate number of gaussian distribution for the pixels which makes it more adaptable for varying scenes.

II. Study of previous System

A. Count Things from Photo

Count Things from Photos app helps businesses automate counting. Open the app, select the right Counting Template for your items, take a photo, and count. If we don't have a Counting Template for your needs, let us know, and we can create one.

B. Dot dot Goose

DotDotGoose is a free, open-source tool to assist with manually counting objects in images. DotDotGoose was purpose-built since most conservation researchers and practitioners working on counting objects in images were using popular software which are not ideally suited for many conservation applications.

III. Literature Review

1) Object Detection

One of the simplest techniques for the object detection is the background subtraction. In this process the observed frame or image is compared with the same scene but with the exclusion of any objects in the scene. The subtraction from the original scene results in the difference of the two images. The difference highlights the areas with significant change and hence identifies the areas of interest [1] [2] 2.Object Tracking The technique for the background modeling can broadly be classified in two categories i) Non-Recursive and ii) Recursive.

A. Non-Recursive Technique

This technique involves the use of storing the previous L video frames. The estimation of the background is done with the temporal variation of each pixel within the buffer [3]. A few non-recursive techniques used commonly are described below.

B. Frame differences

This technique arguably is one of the simplest techniques for background modeling. The video frame used as a background for the frame at time t is the frame at time t – 1 is used [4]. b. Median filtering This technique defines the background estimation as a median of all the frames in the buffer at each pixel location processed. It is also assumed that for half of the frames in the buffer the pixel stays in background [4] [5]

C. Linear predictive filter

In this technique the current background is computed by applying linear predictive filter to the pixels of the frames in the buffer. It also estimates the filter coefficients [4].

D. Non-parametric models

This model has the model structure determined from the data and not specified from a priori. So, it is different from parametric model. This however does not imply that it does not have parameters but rather the parameters it has are flexible and not fixed [6].

E. Recursive Technique

This technique updates a single background model recursively based on the input frames. So, the current model can be significantly affected from the frames of the distant past. Some of the representative recursive techniques are described below:

F. Approximated median Filter

This technique has the scheme of increasing by one the running estimate of the median if the input pixel is larger than the estimate and so on decremented by one if the input pixel is smaller than the estimate [4].

G. Kalman Filter

It is a widely-used recursive technique for tracking linear dynamical systems under Gaussian noise [7].

H. Mixture of Gaussian (MoG)

This technique uses the method of tracking multiple Gaussian distributions simultaneously [4,2 ,5].

2) Object Tracking

Object tracking is one of the leading areas of research and quite important one in the field of computer vision [7]. The availability of higher-grade processors, higher quality cameras and the need of automated video analysis has further on spiked the interest in the object tracking algorithms [3]. The main tracking categories of tracking are [2] i) Point Tracking, ii) Kernel Tracking and iii) Silhouette Tracking

I. Point Tracking

In this technique the object is represented by points. These points are detected in consecutive frames to detect the object. These points are associated on the previous knowledge of the points which shows the object motion and positions in previous frames. An external mechanism is required to detect the objects from each frame [6].

J. Kernel Tracking

In the technique the appearance and shape of object is referred to by the kernel. The kernel can be any shape as an elliptical shape or round or a square shape. The motion of the kernel is tracked in various consecutive frames so as to track the motion of the object. This motion of the kernel can be translation, affine or rotational i.e. parametric transformation [6].

K. Silhouette Tracking

In this technique the region of the object in each frame is estimated. This method uses the encoded information inside the region of the object estimated. The encoded information of the object region can be of the form of shape density or appearance density. These are usually of the edge maps form [6].

IV. Proposed System

In the first part of this project, object detection is carried out using the pretrained HAAR-CASCADE classifier provided by the open cv-community.

For the next, where object detection and tracking are carried out in a video frame, a background Subtraction algorithm is used. The background Subtraction is based on pixel color distortion and brightness distortion is used to determine the foreground.

First of all, the video frames are recognized one by one and the mask is applied to it after applying the Background-SubtractorMOG2 method which returns the background-ratio parameter of the algorithm.

Then from the mask, where the white portion represents the necessary objects and the black represents the unnecessary objects, contours (lines drawn to surround the objects) are drawn. After this step, the recognition is improved by removing the noise from the frames. This is achieved by eliminating the objects with a specific area.

Here the function cv2.createBackgroundSubtractorMOG2 was added including two optional parameters, history is the first parameter, in this case, it is set to 100 because the camera is fixed. varthreshold instead is 40 because the lower the value the greater the possibility of making false positives. In this case, we are only interested in the larger objects. As the objects need not be detected nor tracked from the entire frame, so we proceed by defining a region of interest ROI by extracting a part of the frame. The region of interest can be extracted from anywhere on the frame.

Till now our objects are completely recognized by the system but a problem being there that the shadow of the object is also detected along with objects. So as to remove it the threshold function becomes handy. Starting from the mask we tell it to showcase only the white pixels and omit all the other shades of black by writing “254, 255” i.e., only the pixels having values between 254 and 255 will be shown and others will be discarded.

(0-black, 255-white, and the number between 0 and 255 represent the various shades starting from black to white.)

After this is done, the coordinates of the detected object are obtained and using those coordinates a rectangle is drawn around the object after replacing it with the already drawn contours. Once the object has been created, the position of the bounding box is taken and inserted in a single array.

From this project, all the objects that pass through the ROI are identified and their positions inserted in a specific array. Obviously, the more motorcycles identified the larger our array will be. After this, the array is passed to the tracker file and we will again get an array with the potions but in addition, a unique id will be assigned for each object.

At this point, we just have to draw the rectangle and show the vehicle ID.

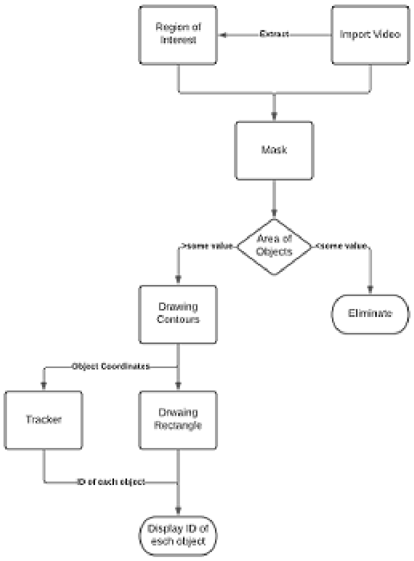

Fig. 1. Flow chart of the algorithm.

V. Results and Discussion

A. *Results of object counting using HAAR-CASCADE classifier*

Fig. 2. Counting the total faces.

B. *Results of object detection and tracking using the BackgroundSubtractorMOG2 algorithm*

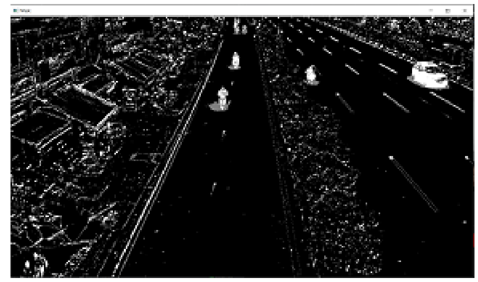

Fig. 3. Mask applied to the input video frame

Fig. 4. Displaying Contours on the interested Objects.

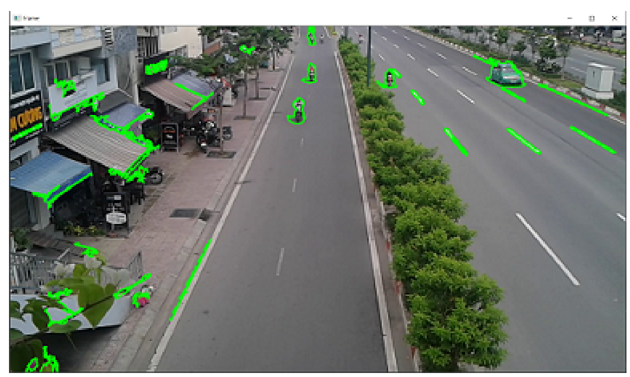

Fig. 5. Removing the noise from the object detection

VI. Discussion

This algorithm was tested on different video frames having different resolution. The detection of the perfectly distinguishable objects was perfectly returned by the algorithm, but it produced errors when there were a large number of objects in the ROI which degrades it accuracy. Specifically for this project, as far as the object detection is concerned, the coordinates of the detected object were sent the tracker which assigned a unique ID to each of the object. But as the frames were rolled, due to a drastic change in the pixel intensity, same object (here specifically bikes, cars, trucks etc.)

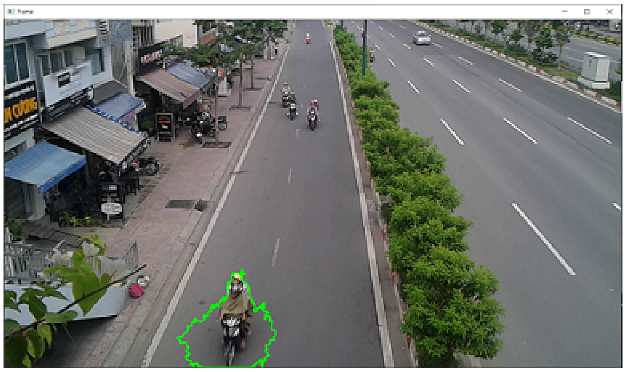

Fig. 6. Restricting our detection in Region Of Interest.

Fig. 7. Object detected along with its shadow.

Fig. 8. Object detected without its shadow.

Fig. 9. Rectangle drawn and displayed the id..

were counted twice which resulted in wrong numbering of the detected objects.

VII. Conclusion

This paper discussed about one of the algorithms for object detection and tracking of moving objects. This approach is based on a new technique for identifying the foreground pixels and color distortion. It detects the objects by eliminating the shadow from the objects. Experiments have been performed and it is found that this particular algorithm achieves the detection of objects more precisely for simple scenarios. For better detection more precise algorithms needs to be used. From this project, it can be inferred that the accuracy provided by this algorithm is nearly 75%. As there are issues with the tracking of the objects it adds on the decrement of its accuracy.

References

[1] C. I. Patel and R. Patel, "Illumination invariant moving object detection," *International Journal of Computer and Electrical Engineering*, vol. 5, no. 1, p. 73, 2013.

[2] S. Gundimada, L. Tao, and V. Asari, "Face detection technique based on intensity and skin color distribution," in *2004 International Conference on Image Processing, 2004. ICIP'04.*, vol. 2. IEEE, 2004, pp. 1413–1416.

[3] S. Y. Elhabian, K. M. El-Sayed, and S. H. Ahmed, "Moving object detection in spatial domain using background removal techniques-state-of-art," *Recent patents on computer science*, vol. 1, no. 1, pp. 32–54, 2008.

[4] U. Chandrasekhar and T. Das, "A survey of techniques for background subtraction and traffic analysis on surveillance video," *Universal Journal of Applied Computer Science and Technology*, vol. 1, no. 3, pp. 107–113, 2011.

[5] K. P. Seng, A. Suwandy, and L. Ang, "Improved automatic face detection technique in color images," in *2004 IEEE Region 10 Conference TENCON 2004.* IEEE, 2004, pp. 459–462.

[6] B. Yang, C. Huang, and R. Nevatia, "Segmentation of objects in a detection window by nonparametric inhomogeneous crfs," *Computer Vision and Image Understanding*, vol. 115, no. 11, pp. 1473–1482, 2011.

[7] W. Khong, W. Kow, H. Tan, H. Yoong, and K. Teo, "Kalman filtering based object tracking in surveillance video system," in *Proceedings of the 3rd CUTSE International Conference*, 2011, pp. 669–675.