MOVING OBJECT DETECTION AND TRACKING USING BACKGROUND SUBTRACTION

Ms. Angila Tshering Bhutia*

Mr. Sourav Paul**

Abstract—

Moving object detection is a fundamental task required for many image and video processing applications like video surveillance, medical imaging, satellite imaging, biometrics identification, automated inspection in industry products etc. An important approach is to perform background subtraction, which identifies moving objects from the portion of a video frame that differs significantly from a background model. By detecting the object in the scene and tracking it many parameters can be determined. One of the output parameter of Object detection and tracking is to estimate the velocity of the object in the scene by calculating the distance travelled. This estimate can then be used for Intelligence video surveillance, Sports analysis etc.

Keywords—Moving Object Detection,Background subtraction,Background model.

* M.Tech (Dept. of Computer Science and Engg.), Sikkim Manipal Institute of Technology, Majhitar, Sikkim

** Assistant Professor-II (Dept. of Computer Science & Engineering), Sikkim Manipal Institute of Technology, Majhitar, Sikkim

I. Introduction

The main aim of object tracking and detection is to establish a correspondence between object parts in consecutive frames and to extract information about objects such as trajectory, posture, speed and direction. Tracking detected objects frame by frame in video is a significant and difficult task [1]. It is a crucial part of smart surveillance systems since without object tracking, the system could not extract cohesive temporal information about objects and higher level behavior analysis steps would not be possible. Moving object detection is the first step in video analysis.

Some of the applications are as follows [2]:

(i) Visual surveillance: A human action recognition system, process image sequences captured by video cameras, monitoring sensitive areas such as bank, departmental stores, parking lots and country border to determine whether one or more humans engaged are suspicious or under criminal activity.

(ii) Content based video retrieval: A human behavior understanding system scan an input video, and an action or event specified in high-level language as output. This application will be very much useful for sportscasters to retrieve quickly important events in particular games.

(iii) Sports analysis: Precise analysis of athletic performance for e.g athlete action is becoming an important tool for sports training, since it has no intervention to the athletic.

In all these applications fixed cameras are used with respect to static background and a common approach of background subtraction is used to obtain an initial estimate of moving objects.

II. Literature Survey

In [3] a background subtraction algorithm is used for detecting moving vehicles and pedestrians in urban traffic video sequences. Approaches varying from simple techniques such as frame differencing and adaptive median filtering, to more sophisticated probabilistic modeling techniques is used. While complicated techniques seems often to produce superior performance, their experiments show that simple techniques such as adaptive median filtering can produce good results with much lower computational complexity. In [4] the proposed approach exploits gray color information for both background subtraction to improve object segmentation. The approach proves fast, flexible and precise in terms of pixel accuracy. The implementation of the background subtraction algorithm is written in Matlab. Improving segmentation results as well as being able to extract additional information such as frame deference, background subtraction allows for improved object detection and thus tracking. The subtraction of the two images gives the good results of the moving object in the surveillance area. The resultant subtracted frame contains the information or data from both the input frames. In [5] for detection of the object, background subtraction method, Kalman filter is used to track the object. The Gaussian Mixture model though experimented to be easy to implement and give complete operation but takes long time for computation and it also gives more noise. Therefore morphological filtering has been used for removing noise. In [6] sports video analysis has been carried out to detect the movement of player from the background image in video sequence and for the player tracking. Motion of a moving player and tracking in a video stream is studied and its velocity is detected. The experimental results show that the proposed method using GMM (Gaussian Median Model) runs quickly, accurately and fits for the real-time detection. A benefit of this method is that it is time efficient, and it works well for small numbers of moving objects.

III. PROBLEM DEFINITION

The main objective is to develop an algorithm that can detect and track moving objects in a video and then estimate the velocity of the object within the frame. Motion detection in consequent images is nothing but the detection of the moving object in the scene. Object tracking is the process of locating and following the moving object in sequence of video frames. The process of locating the moving object in sequence of

frames is known as tracking. This tracking can be performed by using the feature extraction of objects and

detecting the objects in sequence of frames. Velocity estimation is to calculate the position and velocity of the moving object by using the position values of object in every frame. Few algorithms are tested to improve the image quality, to detect moving object, calculation of distance and velocity of the moving object. A good background removal algorithm would handle the relocation of background objects, non-stationary background objects e.g. waving trees, and image changes due to camera motion which is common in outdoor applications e.g. wind load. A background removal system should adapt to illumination changes whether gradual changes (time of day) or sudden changes (light switch), whether global or local changes such as shadows and inter reflections. A foreground object might have similar characteristics as the background, it become difficult to distinguish between them (camouflage). A foreground object that becomes motionless cannot be distinguished from a background object that moves and then becomes motionless (sleeping person). A common problem faced in the background initialization phase is the existence of foreground objects in the training period, which include the actual background, and on the other hand often it is impossible to clear an area to get a clear view of the background, this puts serious limitations on system to be used in high traffic areas. Some of these problems can be handled by very computationally expensive methods, but in many applications, a short processing time is required.

IV. PROPOSED SOLUTION STRATERGY

It has been proposed to use Background Subtraction Technique in order to detect the motion of the moving object. Since the background will remain static, background subtraction method will give the best and complete result in this case. Even though there exist numerous of background removal algorithms in the literature, most of them follow a simple flow diagram, defined by Cheung and Kamath [3],passing through four major steps, which are (1) pre-processing(simple image processing tasks that change the raw input video into a format that can be processed by subsequent steps), (2) background modeling (also known as background maintenance), (3) foreground detection (also known as background subtraction) and (4) data validation(also referred to as post-processing, used to eliminate those pixels that do not correspond to actual moving objects). To obtain background subtraction, the background has to be modeled first. Then, the incoming frame is obtained, and subtracted out from the background model [7]. With the background model, a moving object can be detected. This algorithm is called as "Background Subtraction‟ [8]. Background modeling [9], is the backbone of the Background Subtraction algorithm. Background model defines the type of model selected to represent the background, and the model representation can simply be a frame at time t. Data validation is involved with the collection of techniques to reduce the misclassification of pixels. It is the process of improving the candidate foreground mask based on

information obtained from outside the background model. The background models have three main limitations: first, they ignore any correlation between neighboring pixels; second, the rate of adaption may not match; the moving speed of the foreground objects; and third, non-stationary pixels from moving leaves or shadow cast by moving objects are easily mistaken as true foreground objects. The first problem typically results in small false-positive or false-negative regions distributed randomly across the candidate mask. The most common approach is to combine morphological filtering and connected component grouping to eliminate these regions [10]. Applying morphological filtering on foreground masks eliminates isolated foreground pixels and merges nearby disconnected foreground regions. There are several problems that a good background subtraction algorithm must resolve. Therefore the most commonly used, Gaussian mixture model (GMM) based background subtraction algorithms can be used in this case proposed for the background subtraction in Friedman and Russell, [11] and efficient update equations are given in Stauffer and Grimson,[12]. In Power and Schoonees, [13] the GMM is extended with a hysteresis threshold. This method uses a Gaussian probability density function to evaluate the pixel intensity value. It finds the difference of the current pixels intensity value and cumulative average of the previous values. So it keeps a cumulative average (μ) of the recent pixel values. If the difference of the current images pixel value and the cumulative pixel value is greater than the product of a constant value and standard deviation then it is classified as foreground [14]. That is, at each t frame time, the I pixels value can then be classified as foreground pixel if the inequality:

$|I_t - \mu_t| > k\sigma$ holds; otherwise, it can be considered as background

where ,k is a constant

σ is standard deviation.

Here background is updated as the running average:

$$\mu_{t+1} = \mu_t * I_t + (1-\alpha) * \mu_t$$

$$\sigma^2_{t+1} = \alpha (I_t - \mu_t)^2 + (1-\alpha)\sigma^2 t$$

where, σ, the learning rate, is typically 0.05

I_t is the pixels current value

μ_t is the previous average.

The process of locating the moving object in sequence of frames is known as tracking. This tracking can be performed by using the feature extraction of objects and detecting the objects in sequence of frames. By using the position values of object in every frame, we can calculate the position and velocity of the moving object. By using a Rectangular Bounding Box, a bounding box is plotted around the foreground objects produced from GMM based Background subtraction. By using the dimensions of rectangular bounding box, a centroid is plotted. The position of the centroid is stored in the array and the distance is calculated using

Euclidean distance formula. The velocity of the object movement from frame to frame is calculated by using

the distance and frame rate of the recorded video. The distance travelled by the object is determined by using the centroid. It is calculated by using the Euclidean distance formula. The variables for this are the pixel positions of the moving object at initial stage to the final stage [14].

$$\text{Distance} = \sqrt{(X2 - X1)^2 + (Y2 - Y1)^2}$$

Where X1=previous pixel position

X2=present pixel position in width

Y1=previous pixel position

Y2=present pixel position in height

The velocity of moving object is calculated by the distance it travelled with respect to the time. Euclidean distance formula is used to calculate the distance between the sequences of frames [14]. By using the values of distance with respect to frame rate, the velocity of the object is defined. The defined velocity is of 2-dimension (since camera is static).Velocity of moving object is determined by using the distance travelled by the centroid to the frame rate of the video.

Velocity =Distance travelled / Frame rate

The velocity of moving object in the sequence frames is defined in pixels / second.

The fowchart of the proposed algorithm is given below in fig1.

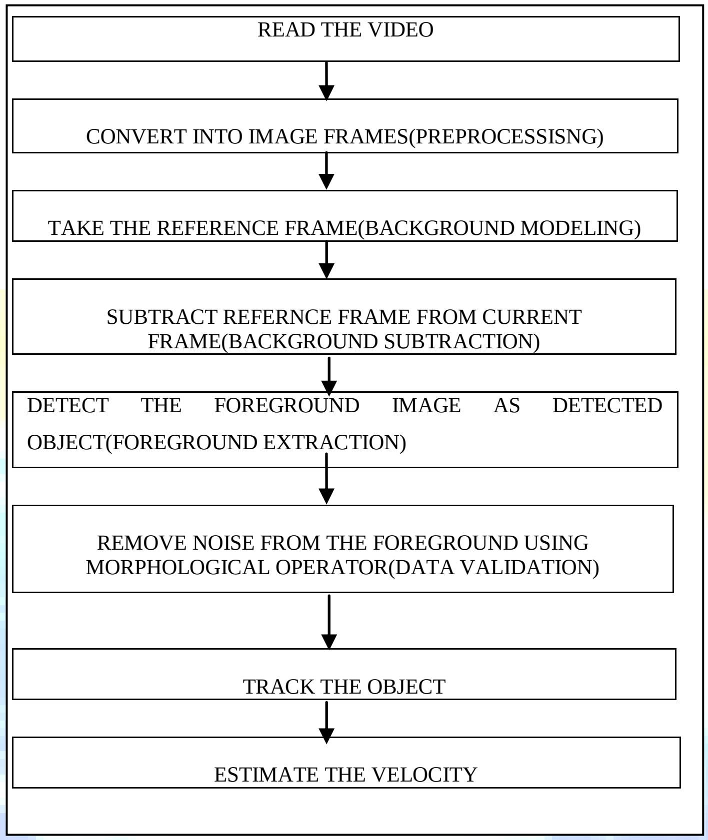

Fig1: Block Diagram for the Proposed Algorithm

REFERENCES

[1] A. Yilmaz, O. Javed and M. Shah,"Object Tracking: A Survey" in ACM Journal of Computing Surveys," Vol. 38, No. 4, 2006.

[2] Priti P.Kuralkar, Prof. V.T.Gaikwad," Human Object Tracking using Background Subtraction and Shadow Removal Techniques," in International Journal of Advanced Research in Computer Science and Software Engineering, Vol. 2, No. 3, 2012.

[3] Cheung S-C, Kamath C. Robust techniques for background subtraction in urban traffic video. In S. Panchanathan and B. Vasudev,editors, Proc Elect Imaging:Visual Comm Image Proce 2004 (Part One) SPIE, 2004; 5308: 881-892

[4] Mahamuni P., R. P. Patil, H.S. Thakar:MOVING OBJECT DETECTION USING BACKGROUND SUBTRACTION ALGORITHM USING SIMULINK IJRET: International Journal of Research in Engineering and Technology ,Volume: 03 Issue: 06 | Jun-2014.

[5] Modi Nirav D, Dr. Sachin R. Gengaje:Intelligence Video Surveillance Added and Removed Object Detection, International Journal of Advanced Research in Computer Science and Software Engineering, Volume 4, Issue 3, March 2014

[6] R.Manikandan, R.Ramakrishnan:Human Object Detection and Tracking using Background Subtraction for Sports Applications , International Journal of Advanced Research in Computer and Communication Engineering Vol. 2, Issue 10, October 2013

[7] A.McIvor, "Background subtraction techniques," in Proceedings of Image and Vision Computing , Auckland, New Zealand,2000.

[8] M. Hedayati, Wan Mimi Diyana Wan Zaki , Aini Hussain, "A Qualitative and Quantitative Comparison of Real-time Background Subtraction Algorithms for Video Surveillance Applications" Journal of Computational Information Systems ", pp 493 – 505, 2012.

[9] M. Hedayati, Wan Mimi Diyana Wan Zaki , Aini Hussain, "A Qualitative and Quantitative Comparison of Real-time Background Subtraction Algorithms for Video Surveillance Applications" Journal of Computational Information Systems ", pp 493 – 505, 2012.

[10]C.Stauer and W. Grimson, Learning patterns of activity using real-time tracking," in IEEE Trans. on Pattern Analysis and Machine Intelligence, 22, pp. 47-57, Aug 2000.

[11] Friedman N., Russell, S, "Image segmentation in video sequences: a probabilistic approach", In: Proc. 13th Conf. on Uncertainty in Artificial Intelligence, 1997.

[12] C.Stauffer, E.Grimson, "Adaptive background mixture models for real-time tracking", IEEE International Conference on Computer Vision and Pattern Recognition (CVPR), 2:246-252, 1999.

[13] Power, P.W., Schoonees, J.A., "Understanding background mixture models for foreground segmentation", In: Proc. of the Image and Vision Computing New Zealand, 2002.

[14] Z.Zivkovic,"Improved adaptive Gausian mixture model for background subtraction", IEEE International Conference on Pattern Recognition (ICPR), pp 28-31, 2004.