

# Moving Object Detection and Validation in Video Surveillance System

Chetan M. Turkane<sup>1</sup>, Prof. Ms. V.D.Nagrle<sup>2</sup>

*Department of Electronics, AISSMS COE Pune*

*University of Pune*

*Pune, India*

chetanturkane@gmail.com<sup>1</sup>, vismita.nagrle@gmail.com<sup>2</sup>,

**Abstract**—The analysis of human body motion is an important method in which computer vision combines with bio-mechanics. This method is widely used in motion detection, motion analysis, intelligent control and many other fields. In the analysis of human body motion; the moving human body detection is important part. The moving human body is detected from the background image in video sequences. Here the new method for the moving object detection based on background subtraction is defined by establishing a reliable background updating model which uses a dynamic optimization threshold method to obtain a more complete moving object. After getting moving object to remove the noise morphological filtering is done. The noise is in form of disturbances which present in the background. For removing the effect of shadow contour projection analysis is combined with the shape analysis, so that moving human bodies detection is done more accurately and reliably. The Background Subtraction method is accurate, faster and fits in detecting real time environment.

**Keywords**—moving object detection, background subtraction, background model

## INTRODUCTION

There are three types of methods mainly used in moving object detection. These methods are the frame subtraction method, the background subtraction method and the optical flow method [1]. In the Frame subtraction method [2] the difference between two consecutive images is taken to determine the presence of moving objects. The calculation in this method is very simple and easy to develop. But in this method it is difficult to obtain a complete outline of moving object. Therefore the detection of moving object is not accurate. In the Optical flow method [1], calculation of the image optical flow field is done. The clustering processing is done according to the optical flow distribution characteristics of image. From this, the complete movement information of moving body is found and it detects the moving object from the background. But there are many disadvantages of this method such as sensitivity to noise, large quantity of calculation, poor anti-noise performance makes it unsuitable for real-time applications. The background subtraction method [2] is the method in which the difference between the current image and background image is taken for the detection moving objects by using simple algorithm. But it is very sensitive to the changes which occur in the external condition and it also has poor anti interference ability. One advantage of this method is, it can provide the most complete object information in the case of the background is known [3]. In the background subtraction method, in a single static camera condition, the dynamic background modeling is combined with dynamic

threshold selection method which depends on the background subtraction. The background is updated on the basis of accurate detection of object.

## FRAME SEPARATION

Frame processing [4] is the first step in the background subtraction algorithm, the purpose of this step is to prepare the modified video frames by removing noise and unwanted objects in the frame in order to increase the amount of information gained from the frame.

Preprocessing of a image is a process of collecting simple image processing tasks that change the raw input video info into a format. This can be processed by subsequent steps. Preprocessing of the video is necessary to improve the detection of moving object's, For example; by spatial and temporal smoothing, snow as moving leaves on a tree, can be removed by morphological processing of the frames after the identification of the moving object.

The numbers of frames are depending upon the frame rate of the video. We are selecting the video in .avi format. So that we can obtain the details of the video and we can select the number of frames for frame separation.

## BACKGROUND SUBTRACTION METHOD

Background subtraction is particularly a commonly used technique for motion segmentation in static scenes [2]. It attempts to detect moving regions by subtracting the current image pixel-by-pixel from a reference background image that is created by averaging images over time in an initialization period. The pixels are classified as foreground where the difference is above a threshold. After creating a foreground pixel map, some morphological post processing operations such as erosion, dilation and closing are performed to reduce the effects of noise and enhance the detected regions. The reference background is updated with new images over time to adapt to dynamic scene changes.

There are different approaches to the basic scheme of background subtraction in terms of foreground region detection, background maintenance and post processing.

In [5] Heikkila and Silven uses the simple version of this scheme where a pixel at location (x, y) in the current image, it is marked as foreground if is satisfied

$$|I(x, y) - B_t(x, y)| > \tau \quad (2.1) \quad (1)$$

Where,  $\tau$  is a pre-defined threshold. The background image  $B_t$  is up- dated by the use of an Infinite Impulse Response (IIR) filter as follows:

$$B_{t+1} = \alpha I_t + (1 - \alpha) B_t \quad (2.2) \quad (2)$$

The foreground pixel map creation is followed by morphological closing and the elimination of small-sized regions. Although background subtraction techniques perform well at extracting most of the relevant pixels of moving regions even they stop, they are usually sensitive to dynamic changes when, for instance, stationary objects uncover the back-ground (e.g. a parked car moves out of the parking lot) or sudden illumination changes occur.

#### Background Modeling

In the background modeling process, the reference background image and some parameters associated with normalization are computed over a number of static background frames [3]. The background is modeled statistically on a pixel by pixel basis. A pixel is modeled by a 4-finite sequence of pixels  $\langle E_i; s_i; a_i; b_i \rangle$  where  $E_i$  is the expected color value,  $s_i$  is the standard deviation of color value which is defined in  $a_i$  is the variation of the brightness distortion, and  $b_i$  is the variation of the chromaticity distortion of the  $i^{\text{th}}$  pixel. The expected color value of pixel  $i$  is given by

$$E_i = [\mu_R(i), \mu_G(i), \mu_B(i)] \quad (3)$$

Where  $\mu_R(i)$ ,  $\mu_G(i)$  and  $\mu_B(i)$  are the arithmetic means of the  $i^{\text{th}}$  pixel's red, green and blue values computed over  $N$  background frames. So far, we have defined  $E_i$  and  $s_i$ .

#### Background Update

For accurately extracting the moving object the background needs to be updated in real time and the background model can better adapt to light changes. In the moving object detection, the pixels judged as belonging to the moving object maintain the original background gray values, not be updated. We update the background model according to following rule for the pixels which are judged to be the background.

$$B_{k+1}(x,y) = \beta B_k(x,y) + (1-\beta) F_k(x,y) \quad (4)$$

Where  $B(x,y)$  is background image,  $F_k(x,y)$  is current image and  $F_3 E(0,1)$  is update coefficient, in this paper  $f_3 = 0.004$ .  $F_k(x,y)$  is the pixel gray value in the current frame.  $B_k(x,y)$  and  $B_{k+1}(x,y)$  are respectively the Background value of the current frame and the next frame. As the camera is fixed, the background model can remain relatively stable at one position for very long period of time. Using this method we can avoid the unexpected phenomenon of the Background, such as the sudden appearance of something in the background which is not included in the original background. Moreover, the impact brought by light, weather and other changes in the external environment can be effectively adapted by the updating of pixel gray value of the background

#### Moving Object Extraction

When the background image  $B(x, y)$  is obtained, subtract the background Image  $B(x,y)$  from the current frame  $F_k(x,$

$y)$ . Set threshold as  $T$ . If the pixel difference is greater than threshold  $T$ , then determines that the pixels appear in the moving object, otherwise, as the background pixels. The moving object can be detected after threshold operation. Its expression is as follows:

$$D_k(x, y) = \begin{cases} 1 & |F_k(x,y) - B_{k-1}(x,y)| > T \\ 0 & \text{others} \end{cases} \quad (5)$$

Where,  $D_k(x, y)$  is the binary image of differential results.  $T$  is gray-scale threshold. Its size determines the accuracy of object identification.

As in the algorithm  $T$  is a fixed value, only for an ideal situation, is not suitable for complex environment with lighting changes. Therefore, this paper proposes the dynamic threshold method, we dynamically changes the threshold value according to the lighting changes of the two images obtained. On this basis, add a dynamic threshold  $T$  to the above algorithm. Its mathematical expression is as follows:

$$\Delta T = \lambda \cdot \frac{1}{M \times N} \sum_{t=0}^{N-1} \sum_{j=0}^{M-1} |F(i, j) - B(i, j)|$$

$$D_k(x, y) = \begin{cases} 1 & |F_k(x, y) - B_{k-1}(x, y)| > T + \Delta T \\ 0 & \text{others} \end{cases} \quad (6)$$

Where,  $\lambda$  is the inhibitory coefficient. According to the requirements of practical applications set the value of the  $\lambda$ , and the reference values is 2.  $M \times N$  is the size of each image to deal with.  $M \times N$  numerical results indicate the number of pixels in detection of motion region.  $T$  shows the overall changes in the environment. If small changes in image illumination, dynamic threshold  $T$  takes a very small value. Under the premise of enough pixels in the detection region,  $T$  will tend to 0. If the image illumination changes significantly, then the dynamic threshold  $T$  will increase significantly. This method can effectively suppress the impact of light changes.

#### A. Extraction of Moving Human Body

Some accurate edge regions are got after doing the median filtering and morphological operations. But the moving human body regions could not be determined. By observation, we can find out that when moving object appears, shadow will appear in some regions of the scene. The presence of shadow [3] it is difficult to extract the moving object accurately. By analyzing the characteristics of motion detection, we combine the projection operator with the previous methods. Based on the results of the methods above, adopting the method of combining vertical with horizontal projection to detect the height of the motion region. This can remove the impact of the shadow to a certain degree. Then we analyze the vertical projection value and set the threshold value (determined by experience) to remove the pseudo-local maximum value and the pseudo-local minimum value of the vertical projection to determine the number and width of the body in the motion region, we will get the moving human body with precise edge[5]. We

are assuming that people in the scene are all in upright-walking state.

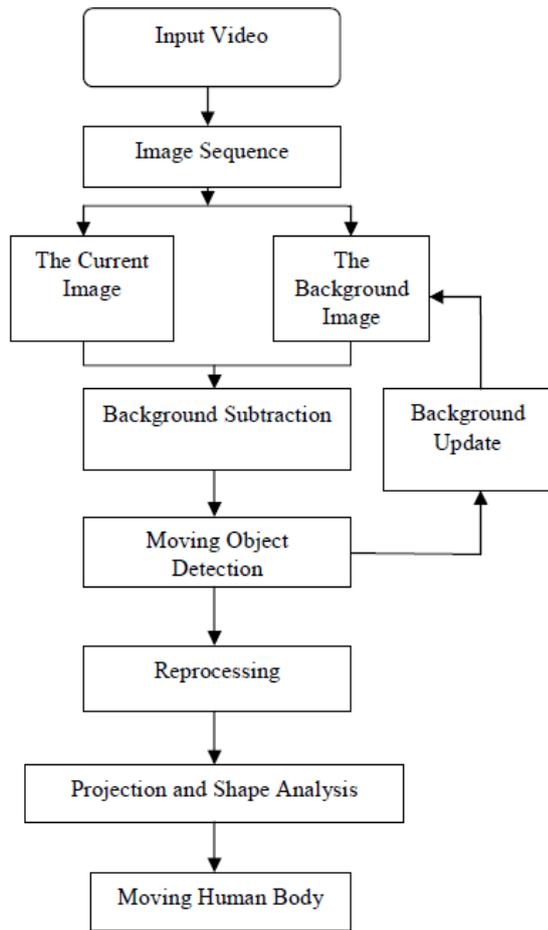


Figure 1: The Flowchart of Moving human Body Extraction

VALIDATION

In the Validation Process, the result of the proposed method is compared to the ground truth image of that frame. The ground truth image is created by using the Photoshop software. It is the accurate method for the detection of the moving object. By comparing those two images we can find out various parameters. Such as mean square error, signal to noise ratio, entropy, correlation coefficient, precision, recall, figure of merit and similarity.

Recall, [7] also known as detection rate, gives the percentage of detected true positives as compared to the total number of true positives in the ground truth.

$$Recall = tp / (tp + fn) \tag{7}$$

Where *tp* is the total number of true positives, *Fn* is the total number of false negatives, and  $tp / (tp + fn)$  indicates the total number of items present in the ground truth. It should be observed that, if the ground truth is given as a binary

detection mask. Recall acts as a pixel-based measure. The ground truth is given as information on the number and features of moving objects in the sequence.

Recall alone is not enough to compare different methods, and is generally used in conjunction with Precision [7] also known as positive prediction, that gives the percentage of detected true positives as compared to the total number of items detected by the method.

$$Precision = tp / (tp + fp) \tag{8}$$

Here, *fp* is the total number of false positives, (*tp+fp*) indicates the total number of detected items, and considerations analogous to those reported for Recall should be made to clarify the pixel-based or frame-based approach of the adopted metric.

F1 also known as Figure of Merit [7] or F-measure, define as the weighted harmonic mean of Precision and Recall

$$F1 = 2 * recall * precision / (recall + precision) \tag{9}$$

And Finally the Similarity [7] can be measured by the following formula

$$Similarity = tp / (tp + fn + fp) \tag{10}$$

By Using these formulas various parameters are find out. Hence we can know the similarity between the proposed method and the actual method for the detection of the moving object.

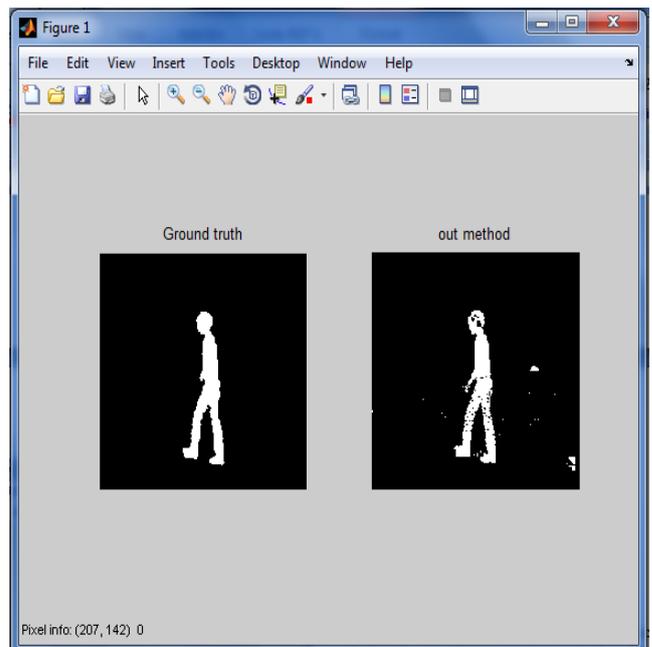


Figure 2: Comparison of Ground Truth image and proposed method image

## RESULT AND DISCUSSION

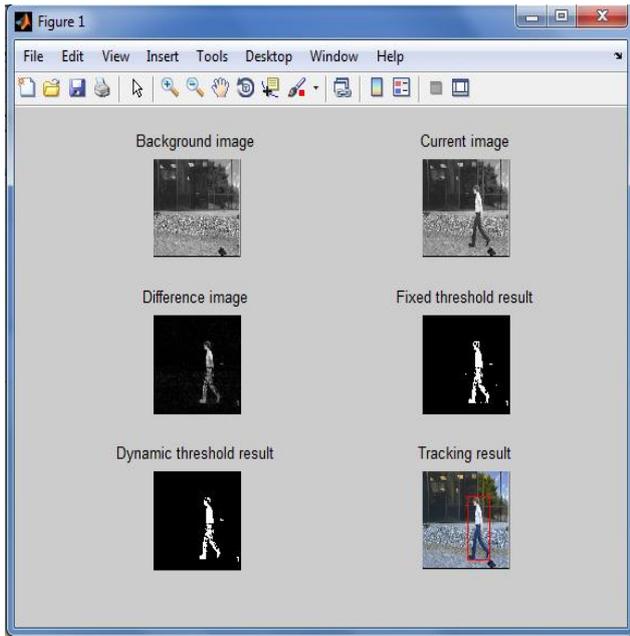


Figure 3: Moving Object Detection

By Comparing the Ground Truth image and proposed method Image some parameters are obtained. Those parameters are Mean square error, Signal to noise ratio and Entropy. They are as follows-

TABLE I- Experimental parameters

S. No	Parameters	Values
1	MSE	0.0134
2	PSNR	66.8458
3	Entropy	0.2938
4	Correlation Coefficient	0.8626
5	Precision	0.9931
6	Recall	0.9934
7	Figure of Merit	0.9932
8	Similarity	0.9866

## CONCLUSION

In this proposed method, a real-time and accurate new method for detecting moving human body is developed. This method is based on background subtraction. To overcome the imperfection in the traditional method of object detection, we establish reliable background model, use dynamic threshold method to detect moving object and update the background in real time. to For removing the shadow effect we combine contour projection analysis with

shape analysis. Experiments show that the algorithm is fast and simple, able to detect moving human body better and it has a broad applicability.

## REFERENCES

- [1] Cheng-Ming Huang, Yi-Ru Chenand, Li-Chen Fu "Real- Time Object Detection and Tracking on a Moving Camera Platform" ICROS-SICE International Joint Conference 2009 Fukuoka International Congress Center, Japan August 18-21, 2009, pp. 717-722
- [2] Lianqiang Niu, Nan Jiang "A Moving Objects Detection Algorithm Based on Improved Background Subtraction" DOI 10.1109/ISDA.2008.337 978-0-7695-3382-7/08 \$25.00 © 2008 IEEE
- [3] Lijing Zhang, Yingli Liang "Motion Human Detection Based on Background Subtraction" 2010 Second International workshop on Education Technology and computer science,2010
- [4] Du-Ming Tsai and Shia-Chih Lai "Independent Component Analysis-Based Background Subtraction for Indoor Surveillance" IEEE TRANSACTIONS ON IMAGE PROCESSING, VOL. 18, NO. 1, JANUARY 2009
- [5] J.Heikkila and O.Silven. "A real-time system for monitoring of cyclists and pedestrians" In Proc. of Second IEEE Workshop on Visual Surveillance, pages 74–81, Fort Collins, Colorado, June 1999.
- [6] I. Haritaoglu, D. Harwood, and L.S. Davis. W4: A real time system for detecting and tracking people. In Computer Vision and Pattern Recognition, pages 962–967, 1998.
- [7] Lucia Maddalena and Alfredo Petrosino. "A self-organizing approach on background subtraction for visual surveillance application" IEEE Transaction on Image Processing, Kerkyra, Greece, pages 1168–1177, 2008.
- [8] S.J.McKenna, S. Jabri, Z. Duric, and H. Wechsler "Tracking interacting people" In Proc. of International Conference on Automatic Face and Gesture Recognition, pages 348–353, 2000.
- [9] H.T.Chen, H.H. Lin and T.L. Liu. "Multi-object tracking using dynamical graph matching" In Proc. of IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pages 210–217, 2001.