Moving Object Detection using Background Subtraction Shadow Removal and Post Processing

Adesh Hardas
Vidyalankar Institute of Technology
Mumbai

Dattatray Bade
Vidyalankar Institute of Technology
Mumbai

Vibha wali
Vidyalankar Institute of Technology
Mumbai

ABSTRACT
In many vision based application identifying moving objects is important and critical task. For different computer vision application Background subtraction is fast way to detect moving object. Background subtraction separates the foreground from background. However, background subtraction is unable to remove shadow from foreground. Moving cast shadow associated with moving object also gets detected making it challenge for video surveillance. The shadow makes it difficult to detect the exact shape of object and to recognize the object. Now days many methods are available for background subtraction. The core of background subtraction is background modeling. Gaussian Mixture model is good balance between accuracy and complexity. For better result post processing is done to output of Gaussian Mixture model. The experimental results give good performance for the proposed method.

Keywords
Background subtraction, Gaussian mixture model, Shadow removal, RGB color space

1. INTRODUCTION
Moving object detection applications are normally car detection, person identification or wild life monitoring. Therefore for real time application background subtraction method should be simple and fast. Background subtraction methods are divided into parametric and non parametric background subtraction. Background model is also classified as static or dynamic. Dynamic background model consist of moving background mostly in outdoor environment.

Background subtraction techniques are mostly used for motion detection in many real-time vision surveillance applications. Difference between the coming frame and the background image is performed to detect foreground objects. Background subtraction gives the most complete feature data, but it is sensitive to dynamic scene changes due to illumination changes and extraneous events. Most researchers are now devoted to developing robust background models to prevent falseness in motion detection caused by scene changes. For example, the algorithm proposed by Stauffer and Grimson [7] uses a mixture of Gaussian distributions to model a multimodal background image sequence, and a technique to update the background model. Any good background algorithm must have the following characteristics (1) should adapt to various levels of illuminations at different times of the day (2) It must handle different weather condition such as fog or snow (3) It must handle the moving objects that first merge into background and then become foreground at a later time. Frame difference is a simple method where current frame is subtracted from reference image. If | frame i – background i| > threshold, then the pixel i is foreground [6]. These techniques are very fast, but the segmentation performance can be quite poor, especially when there is fluctuating in illumination. In order to manage the change in background several complex background methods have been developed.In the moving object detection process, one of the main challenges is to differentiate moving objects from their cast shadows. Since the most common techniques for foreground object detection are inter-frame difference or background suppression, all the moving points of both objects and shadows are detected at the same time. The shadow makes it difficult to detect the exact shape of object and to recognize the object. The accurate detection of a moving object and the acquisition of its exact shape by removing shadow will have great influence on the performance of subsequent steps such as tracking, recognition, classification, and activity analysis.

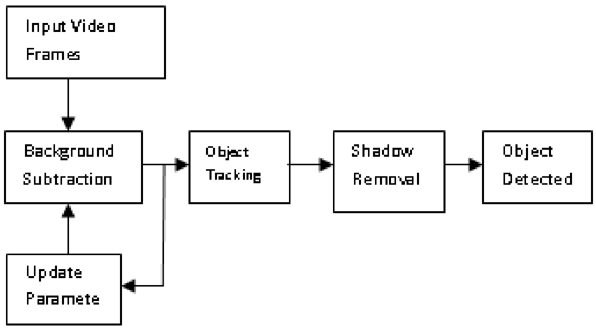

Fig 1. Flow chart of proposed method.

Fig 1 shows the flow of proposed method which is based on RGB colour space for background subtraction and shadow removal. The rest of the paper is organised as follows. In section 2 Background subtraction method is presented. In section 3 characteristics of shadow and its detection in RGB color space is analyzed. Section 4 is experimental results and discussions. Section 5 Future scope of project is discussed. Section 6 is the conclusion.

2. BACKGROUND SUBTRACTION METHODS
Background subtraction techniques will classify each pixel as foreground object or background. If there is a too much variation in pixel value then it is considered as moving object. In [11] Cheung and Kamath have categorized Background Subtraction techniques as: 1) non-recursive and 2) recursive. Non recursive techniques are adaptive to scene changes depending on the buffer size. But while detecting slow moving objects or sudden stop large memory is required for

these techniques. For the recursive technique buffer is not required. They try to update the background model recursively using either a single or multiple model(s) as each input frame is observed. Fig 2 shows general block diagram of Background Subtraction method where current image is subtracted from reference background image. The subtracted result is compared with predefined threshold to detect foreground object.

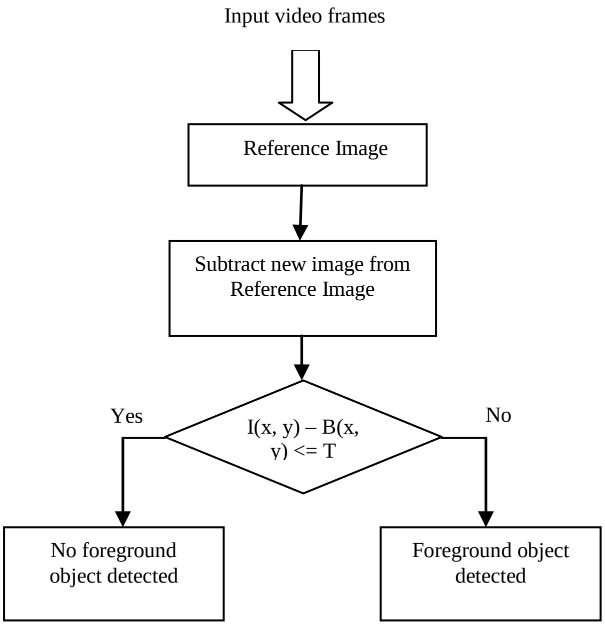

Fig. 2 General Block diagram of Background Subtraction Method

2.1 Frame Difference Algorithm:

Frame differencing is the simplest method in Background subtraction. A background image with no moving objects of interest is taken as the reference image. Pixel value for each co-ordinate (x, y) for each color channel of the input image is subtracted from the corresponding pixel value of the background image. After subtraction If the resulting value is greater than a selected threshold value, then that is a foreground pixel otherwise background. If

$$|\,Frame\,i - Background\,i\,| \geq Threshold \qquad (1)$$

then the pixel i is foreground.

This technique is very fast but the result is very poor when there is illumination change.

2.2 Single Gaussian:

Many authors have modeled background pixel with probability density function learned over sequence of frames. Here with PDF thresholding a pixel with low probability is considered as foreground moving object. In this algorithm t no of frames are required to estimate the mean μ and the standard deviation σ in each color component separately. After computing the parameters, a pixel is considered as a part of the foreground object based on the following formula:

$$|\,\mu(x,y,t) - p(x,y,t)\,| \geq c \times \sigma(x,y,t) \qquad (2)$$

Where c is a constant. This method is suitable for indoor environments where there is gradual illumination changes, this method fails where moving background objects are like trees, flags, etc.

2.3 Mixture of Gaussian model:

The Mixture of Gaussians technique was first introduced by Stauffer and Grimson in [7]. This method represents each pixel of the scene by using a mixture of normal distributions to handle multimodal background images from video. It finds the difference of the current pixel's intensity value and cumulative average of the previous values. So it keeps a cumulative average (µt) of the recent pixel values. If the difference of the current image's pixel values and the cumulative pixel value is greater than the product of a constant value and standard deviation then it is classified as foreground. In this method each pixel is modeled as mixture of k normal distribution. The probability that a certain pixel has a value $\boldsymbol{X_t}$ at time can be written as

$$P(X_t) = \sum_{i=1}^{k} \omega_{i,t} * \eta\left(X_t, \mu_{i,t}, \Sigma_{i,t}\right) \qquad (3)$$

Where $\boldsymbol{k}$ is the number of distributions (currently, 3 to 5 is used), $\boldsymbol{\omega_{i,t}}$ is the weight of the $\boldsymbol{k}$**th** Gaussian in the mixture at time $\boldsymbol{t}$ and $\boldsymbol{\eta}$ **(**$\boldsymbol{X_t}$, $\boldsymbol{\mu_{i,t}}$, $\boldsymbol{\Sigma_{i,t}}$**)** the Gaussian probability density function. For computational reasons, the covariance matrix is assumed to be of the form

$$\Sigma_{k,i} = \sigma_k^2 I \qquad (4)$$

Where σ is the standard deviation The distribution of recently observed values of each pixel in the scene is characterized by a mixture of Gaussians. A new pixel in the current frame is represented by one of the K components of the mixture model That is, at each t frame time, the I_t pixel's value can then be classified as foreground pixel if the inequality:

$$|\,I_t - \mu_t\,| \geq k.\sigma \quad holds \qquad (5)$$

otherwise, the pixel can be considered as background, where k is a constant and σ is standard deviation [8].

The GMM method models the intensity of each pixel with a mixture of $\boldsymbol{k}$ Gaussian distributions. To update the background model, each new pixel in the current frame, is checked against the existing K Gaussian distributions, until a "match" is found. A match is defined as a pixel value within 2.5 standard deviations of a distribution (out of the K components), and this matched component can be a background component or a foreground component which will be verified later. Here the background model is updated using following formulas:

Weights of k distributions at time t are adjusted as:

$$\omega_{k,t} = (1-\alpha)\omega_{k,t-1} + \alpha\left(M_{k,t}\right) \qquad (6)$$

The parameter of the distribution which matches the new observations are updated as

$$\mu_t = (1-\rho)\mu_{t-1} + \rho X_t \qquad (7)$$

$$\sigma_t^2 = (1-\rho)\sigma_{t-1}^2 + \rho(X_t - \mu_t)^T (X_t - \mu_t) \qquad (8)$$

Gaussians are ordered by the value of ***ω/σ***

Then the first B distributions are chosen as the background model, where

$$B = \arg b \min \left(\sum_{k=1}^{b} \omega j > T \right) \quad (9)$$

(a) Original Frame (f_181)

(b) After Background Subtraction

Fig. 3 Extracted moving region by Gaussian mixture method

Fig 3 shows result of background subtraction using Gaussian mixture method. Fig 3 (a) is the original frame from input video and fig 3 (b) is the GMM result where we can see that background is eliminated (black color)and foreground is detected with some shadow points.

3. OBJECT TRACKING

The next step in the video analysis is object tracking. High powered computers with high quality video cameras have increased the need for automated video analysis with a great deal of interest in object tracking algorithms. Kalman filter recursively estimates the state of target object; hence in tracking it is a useful technique which predicts the states of the moving objects. The Kalman filter is a set of mathematical equations that provides an efficient computational (recursive) means to estimate the past, present, and even future states, and even without knowing the precise nature of the modeled system [6]. The Kalman filter estimates a process by using feedback control. At some time the filter estimates the process state and then obtains feedback. There are two sets of equations for Kalman filters: time update equations and measurement update equations. The current state for projecting forward is calculated by time update equations and error covariance estimates to obtain the a priori estimate for the next time step. The measurement update equations are responsible for the feedback. The time update equations are also called as predictor equations and the measurement update equations can also be called as corrector equations.

Time update equations:

Predicted State estimate $X_{k|k-1} = F_k\, X_{k-1|k-1} + B_k\, u_k$ (10)

Predicted Covariance $P_{k|k-1} = F_k\, P_{k-1|k-1}\, F_k^T + Q_k$ (11)

Measurement update equations:

Optimal Kalman gain $K_k = P_{k|k-1}\, H_k^T\, S_k^{-1}$ (12)

Updated state estimate $X_{k|k} = X_{k|k-1} + K_k\, y_k$ (13)

Updated covariance $P_{k|k} = (I - K_k\, H_k)\, P_{k|k-1}$ (14)

In eq. (12) K_k is Kalman gain. H is matrix converting state space into measurement space and R is measurement noise covariance. X_k is a process actual state. The final step in Kalman filter is to update the error covariance $P_{k|k-1}$ into P_k as given in eq. (14). The process is repeated after each time and measurement update pair, with previous posteriori estimates used to project or predict the new priori estimate [6].

4. SHADOW DETECTION AND REMOVAL

From above result as shown in Figure 3 we conclude that the major disadvantage of background subtraction is shadow which is classified as a part of foreground. This is because shadow share same motion patterns as that of foreground. This result in false detection of foreground objects which decreases the tracking performance of the object interested. Example scenarios where detection and tracking performance are affected are: **(i)** Because of shadow, close objects gets merged together, **(ii)** since the shadow pixels are detected they decreases the reliability of the appearance model for each person, increasing the likelihood of tracking loss. There are two types of shadows In particular, that part of the object not illuminated is called self-shadow, while the area projected on the scene by the object is called cast shadow [10]. This last one is more properly called moving cast shadow if the object is moving.

In [3] Horprasert design a color model that separates brightness from chromaticity components. In three dimensional RGB colour space model consider a pixel, i, in the image. Let Ei represent the pixel's expected RGB colour in the reference or background image. The line passing through the origin toward the point Ei is called expected chromaticity line. Let Ii denote the pixel's RGB colour value in a current image. Basically shadow has similar chromaticity but lower brightness. Therefore colour distortion and brightness distortion can be calculated by

$$\alpha i = \arg\min\left(Ii - \alpha i\, Ei\right) \quad (15)$$

$$CD_i = \| I_i - \alpha_i\, E_I \| \quad (16)$$

Where α_i is brightness distortion and CD_i is colour distortion In the proposed method to discriminate the shadow pixel and the object pixel, the R, G, B components are normalised and then multiplied with a fixed value matrix. Output is compared with threshold and the value greater than threshold (YT) is used to calculate L as follows:

$$L = YT \times (116 \times Y3 - 16) + (\sim YT) \times (903.3 \times Y) \quad (17)$$

If L is less than threshold the output is set to 0 as shadow region and if the value is greater than threshold then the output is set to 1 as foreground region. This way shadow mask and foreground mask is calculated.After previous estimations, moving pixels are divided into two sets that are shadow mask and the moving object mask. After this estimation two types of errors may occur, first is shadow detection failure and the other is object detection failure. To improve the accuracy of detection of moving object, a post processing method is used described as follows: In order to break the weak connection between foreground regions, an opening morphological operator is applied to the detected results. Then a flood-fill operation is used to fill the holes in foreground regions.

5. EXPERIMENTAL RESULTS

Experiment is performed on various videos. The performance of proposed method is analyzed by calculating Recall and Precision as follows:

$$\mathrm{Re}\mathit{call} = \frac{TP}{TP+FN} \tag{18}$$

$$\mathrm{Pr}\mathit{ecision} = \frac{TP}{TP+FP} \tag{19}$$

where *TP* is the true positives, *FN* is the false negatives and FP is false positive.

Table 1 Recall and Precision for different Video Sequences

Videos	Recall	Precision
Indoor_1 (f-193)	78.36	72.92
Indoor_2 (f-82)	88.15	82.74
Outdoor person (f-188)	90.14	92.32
Traffic(f-61)	83.49	89.07

Experimental results for two videos are shown in fig 4. First video is outdoor traffic and other is outdoor persons. Original frames are shown in figure (a). Background subtraction using Gaussian Mixture method is performed on these two videos and results are shown in fig 4 (b). After Background subtraction as shown in fig 4 (b) shadows are also detected as moving object which causes false detection of foreground object. Hence shadow is removed by proposed algorithm and after post processing the result is shown in fig 4(c). The parameters of the method are set such that method is useful for both traffic surveillance and outdoor persons detection without any change.

Traffic Frame_61 outdoor person Frame_188

(a)Original Frames

(b) After Background Subtraction

(c) Results after shadow removal

Fig. 4 Final results of proposed method

6. FUTURE WORK

In terms of improving the current system, there are many opportunities for future work in the object segmentation and the object tracking. By studying the motion of the object, different pattern of activity in the video can be learned and then can then be used to detect, classify, and analyze future activities in the area under surveillance.

Current system is based on RGB color space. Other color spaces can also be used for testing different videos like Indoor/Outdoor person detection and traffic videos and compare the systems for proper results such as which system is suitable for indoor or outdoor purpose.

7. CONCLUSION

A robust method is presented in this paper for moving object detection from video with static background scene which has shadow regions also. Experimental result shows that the method is accurate, reliable and efficient. Background subtraction method used in the system is adaptive, thus it updates the changes in the background such as illumination change or if any new object is inserted in the background then it will become a part of background model.

8. REFERENCES

[1] P. Kumar, and A. Mittal,―Study of robust and intelligent surveillance in visible and multimodal framework,‖ informatica, Vol. 31, 2007 pp. 447-461.

[2] M. Piccardi, T. Jan, "Efficient mean-shift backgonmd subtraction", to appear in Proc. of IEEE 2004 KIP, Singapore, Oct. 2004.

[3] T. Horprasert, D. Harwood, and L.S. Davis, "A statistical approach for real time robust background subtraction and shadow detection." In the proceedings of IEEE ICCV'99 FRAME-RATE workshop,1999.

[4] R. Cucchiara, C. Grana, M. Piccardi, A. Prati, and S. Sirotti, "Improving shadow suppression in moving object detection with HSV color information," in Proceedings of IEEE Int'l Conference on Intelligent Transportation Systems, Aug. 2001, pp. 334–339.

[5] A. Prati, Ivana Miki, Mohan M. Trivedi, Rita Cucchiara, “Detecting Moving Shadows: Algorithms and Evaluation” IEEE Trans.Pattern Analysis and Machine Intelligence, vol. 25, no. 7, pp. 918-923, July 2003.

[6] http://en.wikipedia.org/wiki/Kalman_filter.

[7] C. Stauffer, WEL. Grimson, “Adaptive Background Mixture Models for Real- Time Tracking,” IEEE Computer Society Conf. on Computer Vision and Pattern Recognition CVPR,; vol.2, pp. 246-252, 1999.

[8] A. Sanin, C. Sanderson, B.C. Lovell. Shadow Detection: A Survey and Comparative Evaluation of Recent Methods. Pattern Recognition, Vol. 45, No. 4, pp. 1684–1695, 2012.

[9] Zhan Chaohui Duan Xiaohui, Xu Shuoyu Song Zheng Luo Min, 2007. An improved Moving Object Detection Algorithm Based on Frame Difference and Edge Detection”, Fourth International Conference on Image and Graphics 0-7695-2929-1/07 $25.00, 2007 IEEE.

[10] E. Salvador, A. Cavallaro, and T. Ebrahimi. Cast shadow segmentation using invariant color features. Computer Vision and Image Understanding, 95(2):238–259, 2004.

[11] S.C. Cheung and C. Kamath, "Robust techniques for background subtraction in urban traffic video," Video Communications and Image Processing, SPIE Electronic Imaging, UCRL Conf. San Jose, vol.200706, Jan 2004.

[12] C.-T. Chen, C.-Y. Su, and W.-C. Kao. An enhanced segmentation on vision-based shadow removal for vehicle detection. In International Conference on Green Circuits and systems, pages 679-682, 2010.