Modern Education
and Computer Science
PRESS

Comprehensive Study and Comparative Analysis of Different Types of Background Subtraction Algorithms

Priyank Shah
Charotar University of Science and technology, Gujarat, India
e-mail: priyanksshah1988@gmail.com

Hardik Modi
Charotar University of Science and technology, Gujarat, India
e-mail: modi8584@yahoo.com

Abstract—There are many methods proposed for Background Subtraction algorithm in past years. Background subtraction algorithm is widely used for real time moving object detection in video surveillance system. In this paper we have studied and implemented different types of methods used for segmentation in Background subtraction algorithm with static camera. This paper gives good understanding about procedure to obtain foreground using existing common methods of Background Subtraction, their complexity, utility and also provide basics which will useful to improve performance in the future . First, we have explained the basic steps and procedure used in vision based moving object detection. Then, we have debriefed the common methods of background subtraction like Simple method, statistical methods like Mean and Median filter, Frame Differencing and W4 System method, Running Gaussian Average and Gaussian Mixture Model and last is Eigenbackground Model. After that we have implemented all the above techniques on MATLAB software and show some experimental results for the same and compare them in terms of speed and complexity criteria.

Index Terms—Background Subtraction, Moving Object Detection, Video Surveillance, Mean Filtering, Median Filtering, W4 System, Frame Differencing, Running Gaussian Average, Gaussian Mixture Model, Eigen Background.

I. INTRODUCTION

To perform high level tasks like Video Surveillance, traffic monitoring and automated event detection, it is necessary to model the background. The performance of these systems is depended on the accuracy and speed of object detection algorithm.

The basic block diagram of object detection procedure is explained in the figure 1. The real time image is captured through camera and then using background subtraction algorithm we can achieve noisy foreground. This noise can be reduced by applying filtering operation; generally Morphological Operation is carried out in this phase. After obtain connecting region, we can able to find foreground mask and ultimately by applying this mask we can detect and track our object. For real time processing Background subtraction algorithm is well suited. There are many challenges to detect object such as illumination changes like clouds moving in the sky, Motion changes like swaying of trees, Secondary illumination effects like static shadows and moving shadows and camouflage. Many methods are suggested to reduce these problems.
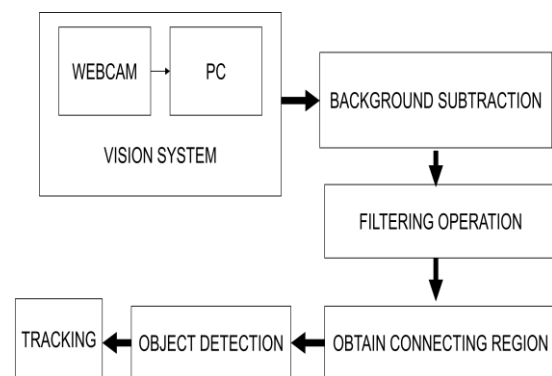

Fig.1. Block Diagram of Object Detection

In this paper we have review some of the Background Subtraction methods which are used for segmentation, which will be explain in following sections.

- Simple method
- Mean Filter
- Median Filter
- W4 system
- Frame differencing
- Running Gaussian Average
- Mixture of Gaussian Model(MGM)
- Eigen background

II. SEGMENTATION METHODS

In basic method for Background subtraction, the static background image without object is taken first as a reference image. After that the current image of the video is subtracted pixel by pixel from the background image and resultant image is converted into binary image which is worked as a foreground mask. For conversion in binary image threshold is required.

$$|I^t(x,y) - B(x,y)| > T \qquad (1)$$

Where, $I^t(x,y)$ is pixel intensity of frame at time t, B(x, y) is mean intensity on background pixel and T is threshold. When difference reaches beyond threshold the pixel categorize as a foreground pixel. So the effectiveness of the object detection is depends on the threshold value. Although this method is very fast, it is very sensitive to illumination changes and noise.

The next method is mean filtering. In this method the background is calculated using the mean of the last n frames as per (2) given below.

$$B(x,y,t) = \frac{1}{n} \sum_1^{n-1} I(x,y,t-i) \qquad (2)$$

Where, B(x, y,t) is reference background calculated at time t and I(x,y) is the pixels intensity. So, the foreground can be found using (3), where T is a threshold.

$$| I(x,y,t) - B(x,y,t) | > T \qquad (3)$$

Another method is median filter, in which the background is calculated as per (3), in which background B(x, y, t) is given by (4).

$$B(x,y,t) = median\ of$$
$$n\ frames\{\ I(x,y,t-i)\} \qquad (4)$$

The advantages of these two above methods are they are fast, easy to implement and adaptive background calculation. The disadvantages are accuracy is depends on object speed, also memory requirement is very high and global threshold i.e. same threshold for every pixel.

The next method is based on method used in W4 system. As mentioned in [1] and [2], we can say, during training sequence maximum intensity, minimum intensity and maximum intensity difference (Dmax) in successive frame of the pixel is calculated and then foreground pixel is calculated based on following equation.

$$I^t(x,y) > Imax(x,y)\ Or$$
$$I^t(x,y) < Imin(x,y)\ Or \qquad (5)$$
$$|I^t(x,y) - I^{t-1}(x,y)| > Dmax(x,y)$$

Another method is based on frame differing also known as Temporal Differencing. As explain in [3] and [4], this method moving object is found by taking difference between two or more consecutive frames using following equation.

$$|I_n(x,y) - I_{n-1}(x,y)| > T \qquad (6)$$

The disadvantage is that if the object moves very slowly it might not detect. So, it is only works when object satisfy certain speed criterion. Also, accuracy is very sensitive with the selection of threshold voltage. This method is based on non-adapting background subtraction.

The next approach is probabilistic known as Running Gaussian average. We can summaries from [2] and [5] that this approach is generally used with (R, G, B) and (Y, U, V) color spaces. In this model for each pixel location a Gaussian probability density function (pdf) of last n pixel is calculated and then it is compared with current frame. For this mean or average and standard deviation are calculated for each pixel. At time t the running average is given by

$$\mu_t = (1-\alpha)\mu_{i-1} + \alpha I_t \qquad (7)$$

Where, I_t is current value of pixel, μ_t is last average and α is the empirical weight, which is useful to give higher weight for current frames and lesser for old frames. Thus, we can update the reference background if changes are not so fast. After computing last average we can distinguish pixels whether it is foreground pixel or background pixel using μ_t and σ_t parameters.

$$|I_t - \mu_t| > k\sigma_t \qquad (8)$$

Where, σ_t is a standard deviation at time t. The pixel classified as foreground if it satisfy above equation otherwise it is classify as background. So, using (7), average can be computed without storing to much memory so the it will increases the speed of operation is well.

As by using running average formula it will also consider the foreground region for calculation of average. To eliminate this other formula is suggested.

$$\mu_t = M\mu_{i-1} + (1-M)\big((1-\alpha)\mu_{i-1} + \alpha I_t\big) \qquad (9)$$

Where M is 1 for foreground pixel and 0 for background pixel. This is also known as selective background update method.

In practice mixture of Gaussian in used. From [5-9] we can write about the basic idea of this method is to define k Gaussian distribution for a pixel to represent its state. The pixel is separated as foreground pixel if it does not match with background Gaussians and then they can be grouped using connected component analysis. Generally value of k is defined between 3-5. If the pixel value is represented as X_t, we can write the probability of pixel in terms of k Gaussian equation.

$$f(X_t = x) = \sum_{i=1}^k \omega_{i,t}\ \eta(x/\mu_{i,t}, \Sigma_{i,t}) \qquad (10)$$

Where,

$$\eta(X_t/\mu_{i,t}, \Sigma_{i,t}) = \frac{1}{(2\pi)^{\frac{n}{2}} |\Sigma_{i,t}|^{\frac{1}{2}}}\ e^{-\frac{1}{2}(X_t-\mu_{i,t})^T \Sigma_{i,t}(X_t-\mu_{i,t})}$$

$\eta(X_t/\mu_{i,t}, \sum_{i,t})$ is the ith Gaussian distribution at t time, $\mu_{i,t}$ is a mean, $\omega_{i,t}$ is covariance matrix and is mean of weight of i th Gaussian distribution and it will also satisfy equation:

$$\sum_{i=1}^{k} \omega_{i,t} = 1 \qquad (11)$$

For simplicity we can assume that each channel of color image is independent. So, we can simply write.

$$\sum_{i,t} = \delta_{i,t}^2 I \qquad (12)$$

If a new pixel value X_{t+1} can be matched with the exiting Gaussians using the (16) than weight of all Gaussians mean and variance can be updated using (13),(14) & (15).

$$\omega_{i,t} = (1 - \beta)\omega_{i,t-1} + \beta \qquad (13)$$

$$\mu_{i,t} = (1 - \rho)\mu_{i,t-1} + \rho\mu_{i,t} \qquad (14)$$

$$\delta_{i,t}^2 = (1 - \rho)\delta_{i,t-1}^2 + \rho(I_t - \mu_{i,t})^2 \qquad (15)$$

Where,

$$\rho = \beta\eta(I_t/\mu_{i,t-1}, \delta_{i,t-1})$$

If X_{t+1} do not match with any of the existing k Gaussian than the probably distribution with lowest weight is replaced by new distribution with $\mu_{i,t}= X_{t+1}$, a high variance and low weight. Then the first B distributions are chosen as background model,

$$B = srg_{b}^{min}\left(\sum_{k=1}^{b} \omega_k > T\right) \qquad (16)$$

Where, T is predefined threshold which is a minimum weight for representing a background. If a current pixel background is matches with any one of this B distribution it consider as a background pixel.
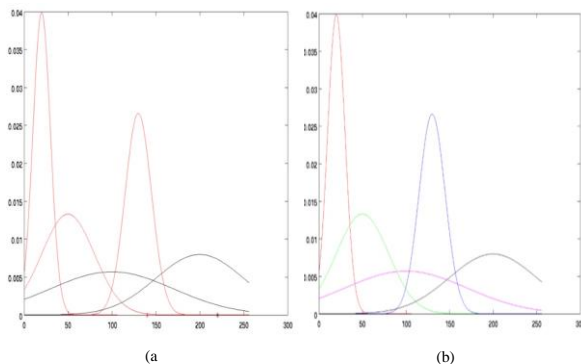

Fig.2. Representation of pixel history in k=5 modes [21]

For example if we consider gray image and set k=5, we will observe the history of a pixel something like as shown in figure 2-(a). If we choose B is three than the distribution with largest three weights becomes back-

ground model which is shown as red color in figure-2-(b).

The advantages of GMM are that we can decide threshold for every pixel and threshold is automatically updated. Secondly, this method allows object to become the part of background without destroying existing background model and GMM does a fast recovery compare to other methods explain above. The main disadvantage of GMM is that it is very sensitive to sudden light changes.

The next method is Eigen Background method. This approach is based on Eigen value decomposition. Principle component analysis (PCA) is the method which is commonly used to mold the background by significantly reduces the dimension of data. PCA can be applied on m frames and it is much faster than Gaussian Mixture Model. The method can be summarized from [10-20] as follow.

- Training set of m images of size NxN are represented by vectors of size N2. So, Now, this two dimensional vector is changed to one dimensional vector Γi.
- Each image is represented by the vector Γi.
- Average image is calculated by

$$\Psi = (\Gamma1 + \Gamma2 + \Gamma3 + \Gamma4+..... \Gamma M)/M \qquad (17)$$

- Each face differs from the average by Φi= Γi - Ψ which is called mean centered Image.
- A covariance matrix is constructed as:
  $C = AA^T$, where A = [Φ1, Φ2, ..., ΦM] size of N2 x N2.
- Eigen vectors corresponding to this covariance matrix is needed to be calculated. AXi is the Eigen vector and λi is the Eigen value.
- Only best M eigenvectors stored in an eigenvector matrix thus eigenvector matrix size will reduce and Eigen vectors of the covariance matrix AA^T are AXi which is denoted by Ui.
- For every new image I can be projected on Eigen space I' $= U^T(I - \Psi)$; where (I-Ψ) is the mean centered image.
- I' is converted back to the image space as I" = U I' + Ψ. Because Eigenspace is suited for static part of seen. So, I" will not contain any moving object.
- So, we can found foreground object using equation

$$I - I"| > T \qquad (18)$$

Where, T is user define threshold.

By, using this step we can able to obtain foreground pixels and by connecting them we can locate foreground object.

III. Experiment Results

We have presented here experimental results on two different videos [22] performed .All experiments are performed over MATLAB 2012 version. The processor is 1.7 GHz and RAM is of 4 GB RAM. The Operating System is Windows 7. Figure 3 shows the original sequence of image on which the operation takes place via different methods. This session explains the results obtain by above methods.

When we apply simple Background Subtraction algorithm, we got results shown in figure 4 & 5. We used two different level of threshold of value 10 and 20. We can conclude that output is very much sensitive to threshold value. Also due to fix background model, output is distorted due to lighting changes.

The next results shown in figure 6-9 is mean and median filtering. We have put result by taking consideration of mean and median of last 5 and 15 frames. By comparing results of two different training signals, we can say that if training signal is too long than memory requirement is very high and if to short than updating rate gets very high. So, slow moving object is not detected property.

Fig.3. Original Video Sequence

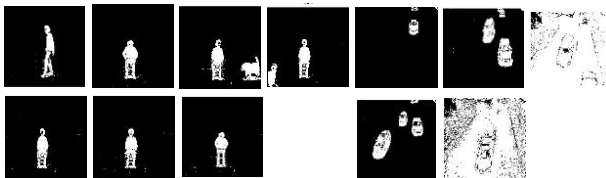Fig.4. Simple Background Subtraction with Threshold 10

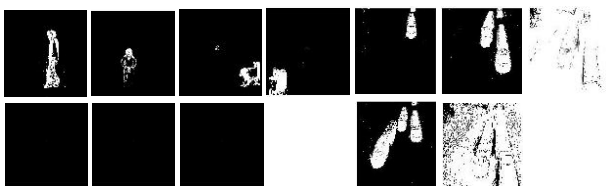Fig.5. Simple Background Subtraction with Threshold 20

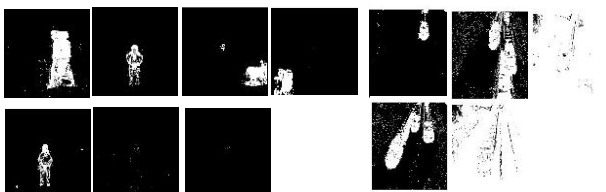Fig.6. Mean Filtering With Threshold 5
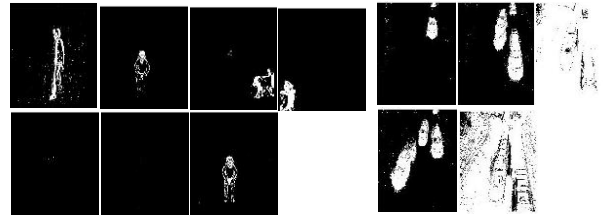

Fig.7. Mean Filtering With Threshold 15

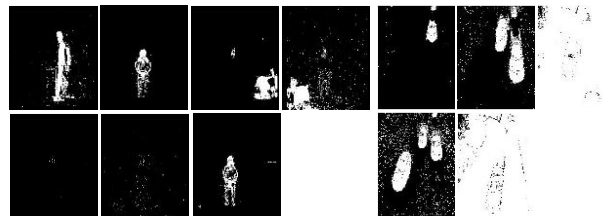Fig.8. Median Filtering With Threshold 5

Fig.9. Median Filtering with Threshold 15

The following figures 10-11 shows foreground detected by w4 system and frame differencing. As output depends on neighboring frames in both techniques, output is not so much sensitive to lighting changes and for the same reason the slow moving object is not detected accurately. In this experiment we used 10 as threshold.
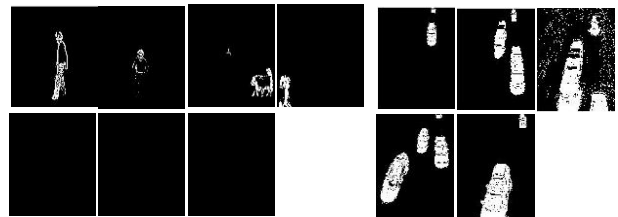

Figure 10 W4 System

Fig.11. Frame Differencing

GMM uses maximum computation to calculate foreground shown in figure 12 than other explained method. Though it can able to detect slightest object movement. It depends on many parameters and have different threshold for each pixel. But we can see that it is very much sensitive to lightening changes.

Output of Eigen background method is shown in figure 13 with no adaption. Though, it is not much affected by lighting changes but it is only useful for small & medium object. Also it uses Principle Component Analysis, so method is faster than GMM.
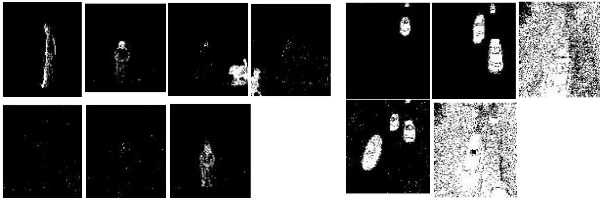
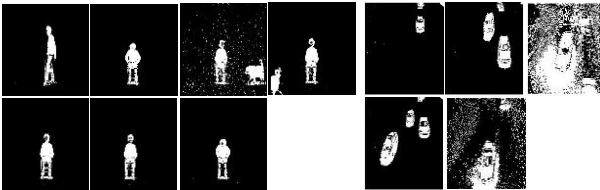Figure 12 GMM With Alpha 0.05, SD 9 and Threshold 0.3

Fig.13. Eigenbackground

By using this output we can compare them as per memory requirement and speed of algorithm which is shown in table 1. Here, values of speed are the observation that we have experience during our output evaluation for each frame. It may not be same for different conditions. Memory requirement specify here is for calculation of foreground for one pixel.

TABLE 1 Memory requirement per pixel and execution speed observed for particular video series per frame

Method	Memory	Speed
Simple	1	.05T
Mean	N	1.5T (5 Training Signal)
		2.5T (15 Training Signal)
Median	N	2.3T (5 Training Signal)
		3.18T (15 Training Signal)
W4	N+3	0.1T
Frame differencing	1	.05T
GMM	M	8T
Eigenbackground	N	.15T

Where, N is no of training signal required to build background model. M is number of Gaussian in GMM and T is one unit time.

IV. CONCLUSION

In this paper, we have studied different types of background subtractions algorithms. We have also done comparative analysis of all the algorithms. This paper gives very good idea about the techniques used for moving object detection using Background Subtraction method. We have experience that Simple method is the fastest among others but it is very sensitive to noise. From all the methods, W4 method is only useful for gray images. Accuracy of Frame differencing to detect object is totally

depends on the speed of the moving object, so we cannot use it if speed of the object is less. Furthermore Running Gaussian average method performs faster than GMM and Eigen background. Also, running Gaussian average requires less memory than GMM and Eigen backgrounds because it uses single threshold for a pixel to decide it is foreground or background, also it is adaptive and fast because it have to update only two-three parameters to update background. If we compare the accuracy, GMM and Eigen backgrounds have good accuracy compare to other methods. But the GMM is more complex because it uses different threshold for each pixel and have many parameters to update, which reduces its speed and increases memory consumption. On the other hand, Eigen background which uses PCA so it will convert n dimension to m dimension (m<n), thus the complexity is reduce by (n - m). Thus, Eigen background is faster than GMM.

Thus, we can select any method from above based on the environment, speed, memory requirement and accuracy of our system.

REFERENCES

[1] A. M. McIvor," Background subtraction techniques, Proc. of Image and Vision Computing,pp.155-163",2000.
[2] J Nascimento and J Marques, "Performance evaluation of object detection algorithms for video surveillance", IEEE Transactions on Multimedia , Vol- 8 , Issue-4,pp-761-774,Aug 2006.
[3] K Gupta1, A Kulkarni,"Implementation of an Automated Single Camera Object Tracking System Using Frame Differencing, and Dynamic Template Matching", CISSE 07 Co-Sponsored by IEEE,pp 245-250, Dec 2007.
[4] K Jadav, M .Lokhandwala, A Gharge," Vision based moving object detection and tracking",National Conference on Recent Trends in Engineering & Technology,pp.13-14, May 2011.
[5] M Piccardi,"Background subtraction techniques: a review", IEEE International Conference on Systems, Man and Cybernetics, Vol-4, pp 3099-3104, 2004.
[6] X Song, J Chen, X Zhou,"A Robust Moving Objects Detection Based on Improved Gaussian Mixture Model", IEEE International Conference on Artificial Intelligence and Computational Intelligence, Vol.2, pp 54-58, 2010.
[7] Z Bian and X Dong,"Moving Object Detection Based on Improved Gaussian Mixture Model", 5th International Congress on Image and Signal Processing, pp 109-112,2012.
[8] C. Stauffer and W Grimson, "Adaptive background mixture models for real-time tracking,", IEEE CVPR, Vol.2, June 1999.
[9] C. Wren, A. Azarhayejani, T. Darrell, and A.P. Pentland, "Pfinder: real-time tracking of the human body", IEEE Trans. on Patfern Anal. and Machine Infell., vol. 19, Issue. 7, pp. 780-785, 1997.
[10] N Oliver, B Rosario, and A Pentland, "A Bayesian computer vision system for modeling human interactions", IEEE Trans. on Patfern Anal. and Machine Zntell., vol. 22, Issue. 8, pp. 831-843, 2000.

[11] C Zhang ,A Pan, S Zheng, X Cao,"Motion Object Detection Of Video Based On Principal Component Analysis", IEEE International Conference on Machine Learning and Cybernetics, Vol-5,pp 3938-3943, 2008.

[12] K Joshi, D Thakore ,"A Survey on Moving Object Detection and Tracking in Video Surveillance System", , International Journal of Soft Computing and Engineering, Vol-2, Issue-3, pp 44-48, July 2012.

[13] I Haritaoglu, D Harwood, and L Davis, "W4: real-time surveillance of people and their activities", IEEE Trans. on Pattern Analysis and Machine Intelligence, vol. 22, no. 8, pp. 809-830, August 2000.

[14] D Maniciu, P Meer "Mean shift: A robust approach toward feature space analysis", IEEE Trans. Patt. Analy. Mach. Intell. Vol-24, Issue- 5, pp 603–619, ,August 2002.

[15] B Moore," Principal component analysis in linear systems: Controllability, observability, and model reduction",IEEE Transactions on Automatic Control, Volume:26 , Issue: 1, PP-17-32, 06 January 2003.

[16] H Hotelling , "Simplified calculation of principal components", Psychometrica, vol. 1, pp.27 -35,1936.

[17] A. Elgammal, D Hanvood, L Davis, "Non- parametric model for background subtraction", Proc. ECCV 2000, pp. 751-767, June 2000.

[18] C Wang, L Lan, Y Zhang, M Gu, "Face Recognition Based on Principle ComponentAnalysis and Support Vector Machine", 3rd IEEE International workshop on Intel. System and Application, pp 1-4,2011.

[19] H Zhipeng, Y Wang, Y Tian, T Huang"Selective eigenbackgrounds method for background subtraction in crowed scenes",18th IEEE International Conference on Image Processing,Issue-1522-4880,pp3277-3280,2011.

[20] M. Seki, T Wada, H Fujiwara, K Sumi, "Background subtraction based on cooccurrence of image variations", Proc. CVPR 2003, Vol. 2, pp. 65-72, 2003.

[21] http://www.cs.utexas.edu/~grauman/courses/fall2009/slides/lecture9_background.pdf.

[22] http://imagelab.ing.unimore.it/visor/video_details.asp?idvideo=45.

Priyank Shah has received the B.E. degree in Electronics and Communication Engineering from Sardar Vallabhbhai Patel Institute Of Engineering & Technology,Vasad, Gujarat and presently pursuing M.Tech. in Communication System engineering from Charotar University of Science & Technology, Changa, Gujarat. His research interests include computer vision, image and video processing and Software Development.

Hardik Modi has received the B.E. degree in Electronics and Communication Engineering from Sardar Patel University and M.E. degree in Electronics and Communication System engineering from Dharmsinh Desai University. Presently, he is Research Scholar of Charotar University of Science & Technology. He is also working as Assistant Professor in Department of Electronics and Communication Engineering of Charotar University of Science & Technology, Changa, Gujarat. His research interests are in Microcontroller's applications, Image & Video Processing, Medical Imaging, Computer Vision and Embedded System.