International Journal of Engineering and Advanced Technology (IJEAT)
ISSN: 2249 – 8958, Volume-3, Issue-1, October 2013

Background Subtraction Techniques for Moving Object Detection in Video Frames

Vrunda A. Mahamuni, Madhuri Khambete

Abstract— Identifying moving objects is a critical task for many computer vision applications. Background subtraction approach is used to separate the moving objects from the background. Many different methods have been proposed over the recent years. This paper provides implementation of five background subtraction techniques these are , Frame differencing, Mean, Median, Single Gaussian distribution and codebook. Implemented techniques are compared based on different parameters e.g. TP rate FP rate Precision and computation time, Such a comparison can effectively guide the designer to select the most suitable technique for a given application in a principled way.

Index Terms—Background modeling, BGS, BG, Foreground.

I. INTRODUCTION

Background subtraction (BGS) is a widely used approach for detecting moving objects in videos. This approach is obtained by taking the difference between the current frame and a "background image", or "background model". For modeling a background each pixel in the image is analyzed using intensity values to determine whether the pixel belongs to the background(BG) or the foreground. The background image must be a part of the scene with no moving objects, and foreground image is the part of scene with moving object.

Several methods for performing background subtraction have been proposed in the recent literature [1,2]. All of these methods try to effectively estimate background model from the sequence of the frames. 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 adverse weather condition such as fog or snow that modified the background. (3) It must handle moving objects during background modeling.

In this work, five background subtraction techniques are discussed and implemented. In section II we discussed the five background subtraction techniques such as frame differencing (FD) , mean method, Median method, single Gaussian distribution method, and codebook method. Section III gives experimental results and performance evaluation, and we conclude in section IV.

II. BACKGROUND SUBTRACTION TECHNIQUES

The methods of background subtraction for moving object detection are as follows:

Manuscript received October, 2013.

Vrunda A. Mahamuni, Department of Electronics and Telecommunication, Cummins College of Engineering for Women, Pune, India.

Dr. Madhuri Khambete, Principal, Cummins College of Engineering for Women, Pune, India.

A. Frame differencing

Frame differencing makes use of the pixel-wise differences between two or three consecutive frames in an image sequence to extract moving regions or moving object as proposed in [3].

Pixels belongs to foreground if

$$B(x,y,t) = I(x,y,t-1) \qquad (1)$$

$$|I(x,y,t) - B(x,y,t)| > Th \qquad (2)$$

Where, *B(x, y, t):* is background image at time t , *I(x, y, t):* is the current image at time t, for n number of frames & Th: is threshold value

If the difference of a pixel intensity of current frame and pixel intensity of Previous frame is greater than threshold then the pixel is classified as foreground pixel otherwise,it becomes background. Depending on the object structure, speed, frame rate and global threshold, this approach may or may not be useful.

B. Mean Method

In this case the background is the mean or average of the previous n frames mathematically is given by the formula.

$$B(x,y,t) = \frac{1}{n}\sum_{i=0}^{n-1} I(x,y,t-1) \qquad (3)$$

$$|I(x,y,t) - B(x,y,t)| > Th \qquad (4)$$

.

Background modeled is the mean of previous n frames on pixel by pixel basis , this modeled background is subtracted from current frame or image . if this difference is greater than threshold value then that pixel is belongs to foreground otherwise it becomes a background pixel.

C. Median Method

This method models background as median of the previous n frames on pixel by pixel basis. Mathematically this method is represented as ,

$$B(x,y,t) = median\{\, I(x,y,t-i)\} \qquad (5)$$

$$|I(x,y,t) - \ median\{\, I(x,y,t-i)\}| > Th$$
$$where\ i \in \{0, \ldots \ldots . . n-1\} \qquad (6)$$

Background modeled is the mean of previous n frames on pixel by pixel basis. Pixel belongs to foreground If difference between current frame and background frame is greater than threshold else it becomes a background pixel.

D. *Single Gaussian distribution*

The Single Gaussian distribution (unimodal distribution) model used in [5,6] models each pixel as single Gaussian distribution by calculating mean and standard deviation on pixel by pixel basis.

The Algorithm for background subtraction:

i. Model the values of a each pixel (i,j) as a single Gaussians for n-1 frames
ii. For modeling the pixel value find mean and standard deviation for n-1 frames with condition of the area covered under the pdf curve that is bounded by the distribution of $[\mu \pm 2.5\sigma]$
iii. Check current pixel value with existing Gaussian of n-1 frames
iv. If pixel value fits in the background distribution model is declared as background, else it becomes foreground.

E. *Codebook Method*

Codebook method proposed in [7] construct a background model by using quantization / clustering technique. A background is modeled as a codebook for each (i,j) pixel position for n-1 frames on pixel by pixel basis.
The method proposed in [7] is summarized as follows:
Let C=(c1,c2,……cL) is the codebook for a single pixel consisting of L codeword's ,each pixel has a different codebook size based on its sample variation .Each codeword ci, i=1,2,….L consist of an RGB vector vi=(Ri,Gi,Bi) & 6 tupel auxi as auxi=(Ǐi ,Îi ,*f*i ,λi ,pi,qi), Where Ǐ ,Î is min and max brightness, *f* is frequency with which the codeword has occurred, λ is maximum negative run length(MNRL) i.e. longest interval during training ,that the codeword has not activated & p,q is first and last access time of the codeword has occurred.
During background generation the value of incoming pixels in the current frame is compared to the codebook to find a matching codeword . this is done by calculating a color distortion and brightness range between current pixel and matching codeword, as follows.

$$colordist(x_t, v_m) \le \epsilon \quad (7)$$

$$brightness\{I, (\check{I}m, \hat{I}m)\} = true \quad (8)$$

Where x_t is current pixel, v_m matching codeword, ϵ is threshold, I is intensity of x_t , $\check{I}m, \hat{I}m$ is min and max brightness of matching codeword.
Finally for background subtraction determine the current pixel value x_t and its intensity I, then find the matching codeword c_m to x_t based on its color distance and brightness range same as in equation 7& 8. if there is no any matching codeword to x_t ,then the current pixel is classified as foreground ,otherwise it becomes background pixel.

III. EXPERIMENTAL RESULTS AND PERFORMANCE EVALUATION

In this section we apply various background subtraction methods on two test video sequences in wallflower dataset. The methods implemented are Frame differencing in section A, Mean method in Section B, Median method in section C, Single Gaussians Distribution in Section D & codebook method in section E.

Test were run on two video sequences in wallflower dataset, one is waving trees sequence and second is foreground aperture sequence, The camera is stationary in both of the video sequences. The waving trees sequence consists of 286 frames at 120×160 resolution , frame 247 is provided as ground truth frame, in this sequence trees are moving in background, experimental results are tested on 50 frames as shown in fig1. Foreground aperture sequence consist of frames at 120×160 resolution, frame 489 is a ground truth image. This sequence shows one person is sleeping and wakes up after some time, test is done on 50 frames as shown in figure 2.
Performance evaluation:
The four measurement parameters are used to evaluate the ability of each method to correctly detect motion by using a ground truth data , such as true positive rate (TP) is also called as recall rate or true foreground detection rate , falls positive rate(FP) or true background detection rate ,precision and computation time for background training and background subtraction are defined as follows.

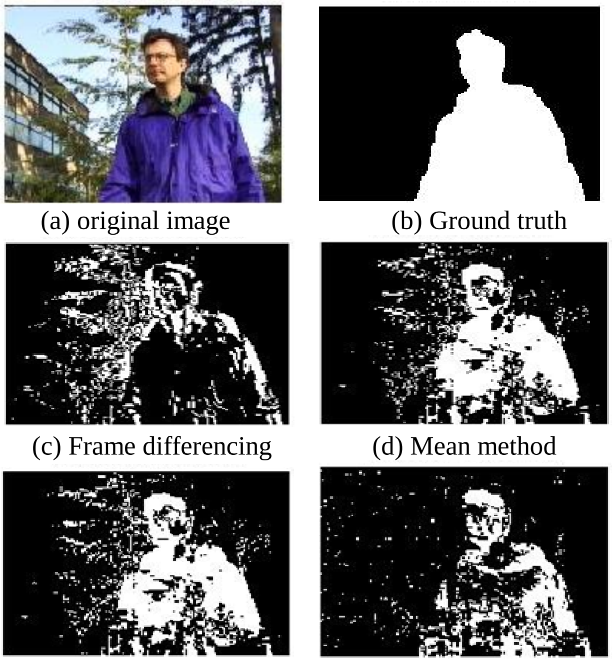
$$Tp\ rate = \frac{tp}{tp + fn} \quad (9)$$

$$Fp\ rate = \frac{fp}{fp + fn} \quad (10)$$

$$Precision = \frac{tp}{tp + fp} \quad (11)$$

Where,tp is true positives pixels, fp is false positive pixels, fn is false negative pixels. Precision denotes positive predictive values which are defined as the proportion of the true positives against all the positive results. The performance of each method is evaluated on two video sequences in wallflower dataset by calculating TP, FP and Precision parameters as shown in table 1and table 2.

Results on waving trees sequence:



(a) original image (b) Ground truth
(c) Frame differencing (d) Mean method
(e) Median method (f) Single gaussian model

(g) Codebook

Fig:1 Foreground Detection results on Waving trees sequence

Evaluation:

BGS techniques	Performance parameters				
	TP rate	FP rate	Precision	BG training time (sec)	BGS time (sec)
Frame diff	0.19	0.14	0.36	---	---
Mean Method	0.64	0.064	0.81	0.50	0.53
Median Method	0.65	0.067	0.80	0.55	0.51
Single Gaussian Distribution	0.47	0.02	0.91	0.50	0.49
Codebook	0.97	0.018	0.95	1.48	0.61

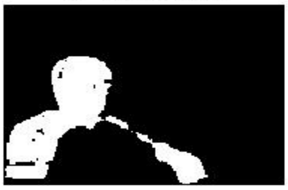
Table 1: Performance parameters on waving trees sequence

Results on Foreground aperture sequence:

(a) original image

(b) Ground truth

(c) Frame differencing



(d) Mean method

(e) Median method

(f) Single Gaussian model

(g) Codebook

Fig:2. Detection results on Foreground aperture sequence

Evaluation:

BGS techniques	Performance parameters				
	TP rate	FP rate	Precision	BG training time (sec)	BGS time (sec)
Frame diff	0.17	0.028	0.67	---	---
Mean Method	0.48	0.034	0.83	0.50	0.60
Median Method	0.48	0.034	0.83	0.58	0.59
Single Gaussian Distribution	0.78	0.84	0.24	0.54	0.69
Codebook	0.96	0.007	0.97	9.13	0.58

Table 1: Performance parameters on Foreground aperture sequence

IV. CONCLUSION

In this paper we have presented most widely used methods for background subtraction. Goal of this work is, effectively guide the user to select the most suitable method for a given application in a principled way.

By observing the experimental results and performance evaluation factors, the codebook method worked best in detecting moving objects. This method offers high TP rate , with good accuracy and limited memory, only the problem that it takes more computation time for background modeling so this is a quit slow method as compared with other methods of background subtraction.

The mean and median methods worked second best since it detects the most details in the video sequences after the codebook method. These methods are Extremely easy to implement and use. The limitation of Mean and median background methods have relatively high memory requirements with acceptable accuracy.

Amongst the methods studded, frame differencing is the simple and very fast method of background subtraction. The accuracy of frame differencing depends on object speed and frame rate. The single Gaussian method is computationally complex as compared with basic methods, This method is not sufficient for dynamic background modeling.

REFERENCES

[1] Massimo Piccardi.”Background subtraction techniques: a review”, IEEE International Conference on Systems, Man and Cybernetics 2004.

[2] Marek CHOVANEC, ”Computer vision vehicle tracking using Background subtraction” , journal of Information, Control and Management Systems, Vol.1, (2005).

[3] R.C. Jain and H.H. Nagel, “On the analysis of accumulative difference pictures from image sequences of real world scenes,” IEEE Trans. on Pattern Analysis and Machine Intelligence, vol. 1, no. 2, pp. 206–213, 1979.

[4] B. Tamersoy ,"Background Subtraction – Lecture Notes".University of Texas at Austin. (September 29, 2009).

[5] C.R. Wren, A. Azarbayejani, T. Darrell, and A. Pentland, “Pfinder: Real-time tracking of the human body” , IEEE Transactions on PAMI, Vol. 19, no. 7, pp. 780-785, 1997.

[6] Horprasert T, Harwood D, Davis LS. A statistical approach for real-time robust background subtraction and shadow detection. IEEE Frame-Rate Applications Workshop, Kerkyra, Greece; 1999.

[7] Kim, K. , Chalidabhongse T. H., Harwood D, Background modeling and subtraction by codebookConstruction” Proc. Int. Conf. Image Processing, Singapore, 3061-3064, (2004).

Vrunda A Mahamuni, received B.E. degree in Electronics and Telecommunication from TPCT’s College of Engineering, Osmanabad, Maharashtra, India, in 2008. She is presently pursuing M.E in Signal Processing from Cummins College of Engineering, Pune, Maharashtra, India.

Dr. Madhuri Khambete, Principal Cummins College of Engineering for Women ,Pune, INDIA.