Performance Enhancement through Handling of False Classification for Smart Video Surveillance

Lavanya Sharma[1,*] and Nirvikar Lohan[2]
[1,*]Department of Computer Science, Uttarakhand Technical University, Dehradun, India.
[2]Department of Computer Science, College of Engineering Roorkee, Roorkee, India.

Over the last decennium, Visual surveillance has become an active research domain for academicians, researchers or industry due to its rapid day by day growing importance in terms of realistic environment. The solutions of the video surveillance are security tools that help us to monitor various things in terms of moving object (i.e. locations, monuments, building, people, etc. In this paper we proposed an efficient method for detection of object using background subtraction technique by enhancing the exiting method. To preserve the shape and removal of noisy pixels some post processing tools were also used. Comparative analysis of our method with considered state-of-the-art method reveals that proposed method shows better outcomes both in terms of qualitative and Quantitative analysis.

Keywords: Video Surveillance, Background Subtraction, Object Detection and Tracking, Pedestrian, Morphology.

1. INTRODUCTION

The world has onlooked several safety issues in current times. An undeviating impact is day to day rapid growth of requirement in visual surveillance system or computer vision system that lead to large scaled research work in this active area of research. Over last few decades it becomes an ideal choice for academicians, researchers and industrials point of view. The automated surveillance system is a very rigorous and efficacious area of research due to rapid growth of several suspicious social events and sturdy terrorism activities [1,2]. In several cases it becomes very difficult to classify the background or foreground pixel's which lead to miss classifications of foreground pixels as background pixels [3-6]. So, it becomes very important to classify pixels clearly as foreground or background because rest of the sub sequential steps are highly dependent on object detection. In above scenario, non moving camera with static background technique is used to achieve correct pixels classification in nature. To overwhelm this problematic issues some further modification are to be added to enhance the pixels quality. The detection of a moving object has to deal with the problems encountered during the process of object detection using background subtraction technique such as how to develop an effective method that is computationally less expensive, selection of a threshold value for pixels classifications (background or foreground) and most important aspect is how to improve the detection quality by avoiding various kind of issues available in background scene [7-9]. In this paper, we proposed an efficient method for object detection using background subtraction for video surveillance that results in performance improvement as comparison to existing work. The main motive behind this work is to detect meaningful moving information from each consecutive video frame.

In literature various Background Subtraction (BGS) techniques for detection of moving object are present. Stauffer *et. al.* [1] presents a model where each pixel was represented as a gaussian and to update the background model an online approximation was used. Lee [3] presents a spatial similarity based technique with low complexity and faces shadowing issue that is resolved by structural similarity. Jung *et. al.* [4] proposed an efficient method for gray scale video sequences using BGS and shadow removal. In this work a robust statistical descriptors are applied and then pixels are classified using pixel ratio. Haque *et. al.* [10] has improved Stauffer's method [1] and proposed a Gaussian Mixture Model (GMM) based non static background generation method in order to get better results in terms of both stability and detection quality against Stauffer's [1] model. Barnich *et. al.* [6] employed particle swarm optimization that automatically tuned the parameters of GMM algorithms parameters. A combined approach of GMM and region based algorithm with color histogram and texture information is proposed. This model performs well but fails in case of highly dynamic background like swaying leafy movement or illumination variations causes several false alarms. Ng *et. al.* [7] presented a non parametric algorithm named Motion Detection (MODE) that is independent of several background challenges like dynamic background, noise, and bootstrapping. Yadav *et. al.* [11] proposed a Quasi-Euclidian distance function that concludes variability in terms of distance between corresponding pixels of background model and test frame. To improve the detection quality, enhancement step is carried out where a threshold connected components and blob labeling is also applied. This method also detected shadow along with moving object.

This paper is categorised into four parts, first section deals with the introductory part of the article. In the next section 2, proposed method for detection of foreground object in video surveillance is discussed. Section 3 deals with experimental analysis of both methods. In the last section 4, conclusion and future work is discussed.

2. PROPOSED METHOD

In our proposed work we enhanced the work of Haque *et. al* [10] by modifying the background (BG) mode and then applying some post processing tools, where every pixel within a intense foreground domain is predicted by the bulk of voting when the corr. pixel's (neighborhood) weight $\{W_{ti}\,(x, y> 0.5)\}$. Firstly, Model Learning is performed then object detection is done on the basis of four criteria which are listed below:

(i) Probabilistic background subtraction.

(ii) Basic background subtraction with multi-background.

(iii) Detection of foreground object.

(iv) Post processing process.

2.1. Probabilistic Background Subtraction

In this work, Gaussian mixture (GM) is used to model a background adaptively, where each and every pixel is modeled independently by a Mixture of Gaussian (MoG) distribution. Here, each Gaussian represents the intensity distribution at various levels. [11-15].

2.2. Basic Background Subtraction with Multi-background

The probabilistic technique is experimented with the basic background subtraction and then this process classified a pixel within a dense foreground domain ($W_{t1}(x, y) > 0.5$) and ($W_{t2}(x, y) > 0.5$)), the corr. pixel can be classified as a foreground (FG) regardless of its classification because this results in improvement of the detection quality of a pixel inside that particular object domain.

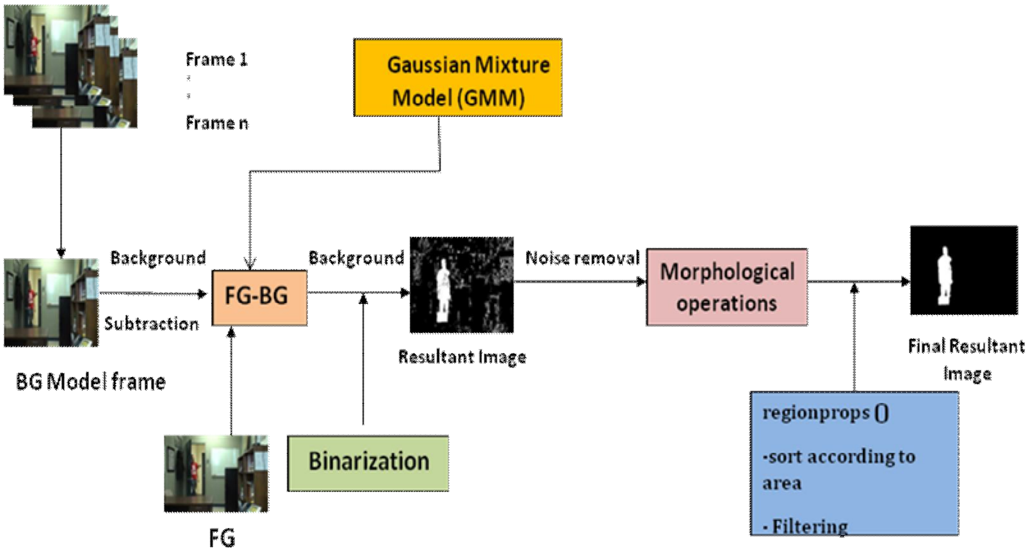


Fig. 1: Block diagram of proposed Method.

2.3. Detection of Foreground Object

To detect the object two decisions $D_i(x, y)$ $(1 \leq i \leq 2)$ were taken into consideration for every pixel (x, y). If pixel belongs to foreground (FG) then its value is 1 and if pixel belongs to background (BG) then its value is 0, respectively. Final detection $D_i(x, y)$ for a pixel (x, y) can be determined by using the object detection method. For each pixel-level detection decision $D_i(x, y)$, a neighbourhood weight of a pixel is computed that represents the part of foreground (FG) within the neighbourhood including that pixel:

$$W_i(x, y) = \forall\ (m, n) \text{€} N\ D_i\,(m, n)/N_n \qquad (1)$$

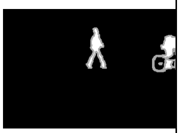
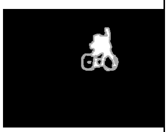
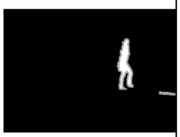
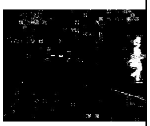
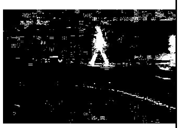
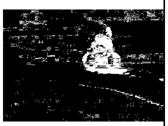
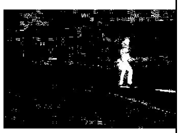
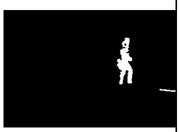
Where, N denotes the neighbourhood of a pixel (x, y) and m, n represents the matrix.

3. EXPERIMENTAL SETUP

In this section proposed enhanced method's results are compared with Haque *et. al.* [10]. For experimental analysis, analysis is carried out on gray-scale video frames. Here, we use 2 datasets publicly available on internet: Change Detection [2] and Pedestrian [5] dataset. Both datasets contains video sequences from realistic scenario. All experiments were carried out using Windows-XP environment. All the experimental analysis have been performed and implemented on MATLAB 2011*b*.

3.1. Qualitative Analysis

The results of Haque *et. al.* [10] and proposed method are evaluated at a learning rate, α = 0.01 for both methods. The proposed method shows outstanding detection results which clearly reflect the elimination of slow leafy movement, illumination variations, and dynamic background.

Pedestrian Dataset [5]	Frame no 412	Frame no 499	Frame no 544	Frame no 485	Frame no 400
Original frame					
Ground truth					
Haque *et. al.* [10]					
Proposed					

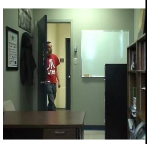
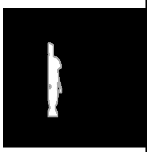
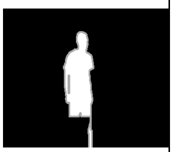
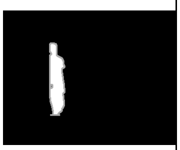
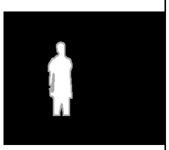
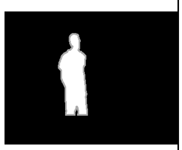
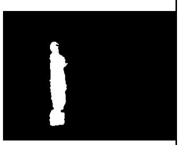
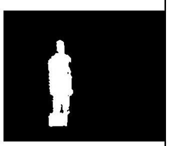
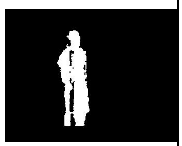
Office Dataset [2]	Frame no 498	Frame no 646	Frame no 598	Frame no 612	Frame no 639
Original frame					
Ground truth					
Haque et. al [10]					
Proposed					

Fig. 2: Qualitative results: row wise (i) Dataset and their frame sequences (ii) Original frame (iii) Ground truth mask (iv) Haque *et. al.* [10] (v) Proposed work.

3.2. Quantitative Analysis

In this section, we have evaluated that in case of pedestrian dataset [8]. The proposed method takes less time to execute all video sequence whereas in case of office dataset [2], the considered method takes less time but the time difference is not much. But terms of time constraint, our method works in an optimized manner. This work has experimented over office and pedestrian frame sequences of change detection dataset. The Jaccard index is used to measure the similarity of proposed result with the available ground truth. Another parameter, Jaccard distance is used to compute the dissimilarity of the proposed result with the available ground truth. Similarity, this work computed these parameters for Haque *et. al.* [10] as shown in Table 1.

3.3. Performance Metrics

For performance analysis of two metrics are taken into considerations, precision and recall where, TP: True Positive; FN: False Positive; FP: False Negative.

$$\text{Precision (P)} = TP/(TP+FP) \tag{2}$$

$$\text{Recall (R)} = TP/(TP+FN) \tag{3}$$

The Table 2, depicts the computation time under which the proposed method and considered method has executed. According to the given time analysis, the total average time is less than considered state of the art method i.e. Haque *et. al.* [10]. In proposed method we have improved the Haque *et. al.* [10] and minimizes the unwanted blobs, pixels or noise. Apart from these, the proposed method generate better performance (refer Table 1 - Table 2) as compare to considered peer method.

Table 1: Jaccard Index and Jaccard Distance of Results over both sequences.

Dataset	**Jaccard Index** (Similarity Measure)		**Jaccard Distance** (Dissimilarity Measure)	
	Haque *et. al*. [10]	**Proposed**	**Haque *et. al.* [10]**	**Proposed**
Office	0.3564	0.7787	0.6436	0.2213
Padestrian	0.3944	0.6909	0. 6056	0.3091

Table 2: Time taken by algorithms to generate results on different dataset.

Dataset	**Haque *et. al.* [10]**	**Proposed Method**
Office	1588.804 sec	1590.240 sec
Pedestrian	1892.342 sec	1882.600 sec
Average Time	1740.573 sec	1736.420 sec

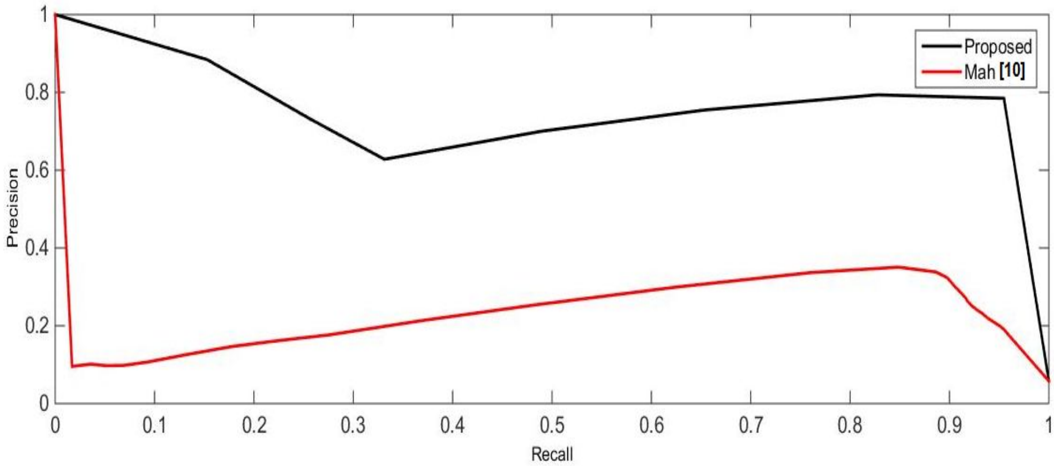


Fig. 3: Precision - Recall curve for both methods.

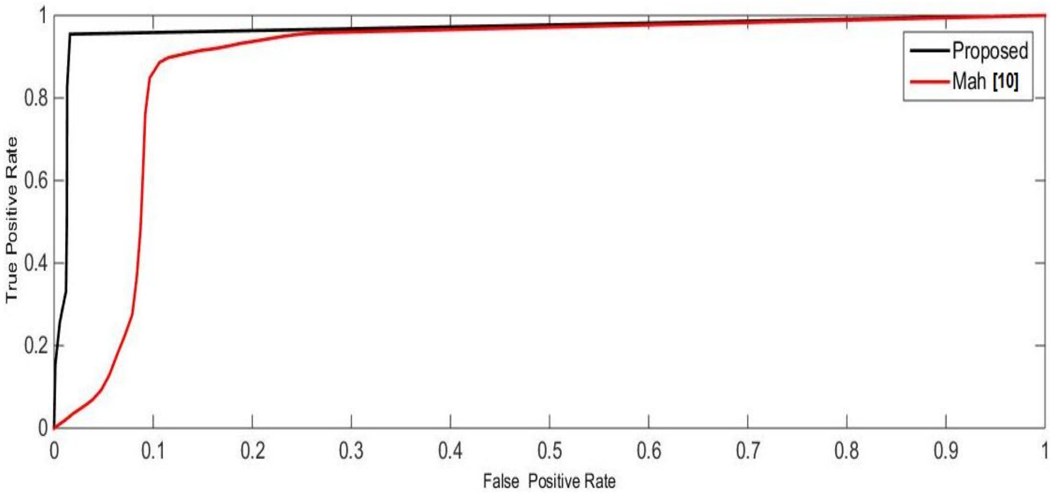


Fig. 4: ROC curve for both methods.

4. OBSERVATIONS

The results obtain from both qualitative (Fig. 2) and quantitative analysis (Table 1, Table 2) reflects:

4.1. Quantitave analysis (Fig. 3, Fig. 4) shows that the proposed method is applicable to real time video surveillances systems with stand still cameras.

4.2. Table 1, reflects the Jaccard Index and Jaccard Distance of Results over both sequences which also shows better results of proposed work.

4.3. From Table 2, it clearly depicts that the proposed work takes less time in case of pedestrian data set and total average time is also very less than compared method but proposed method take very little more time in case of office dataset.

4.4. The quantitative analysis of the proposed methods shows more accurate outcomes in terms of performance metrics such as Receiving operating characteristics (ROC) curve, precision and recall.

Finally, the proposed work achieves better performance as comparable to the considered state-of-the-art method Haque *et. al.* [10].

5. CONCLUSION & FUTURE WORK

This work has proposed an improved method for detection moving object using background subtraction technique which can be applicable to many real time applications with non moving camera. The outcomes show that our method performs very well in terms of quantitative and qualitative analysis. The qualitative results of proposed method clearly depict better detection quality than considered peer method. The quantitative results of proposed work achieve better performance than state-of-the-art method. The proposed method takes less time in case of pedestrian data set and total average time is also very less than compared method. But proposed method takes very little more time in case of office dataset. Proposed method also represents more accurate results in terms of various performance metrics such as precision, recall, and ROC curve. In future we will work on data set with moving camera and dynamic background, and on cloud environment.

ACKNOWLEDGEMENT

The first author is very thankful to her Daadu Ji *(Late. Shri. Ram Krishan Choudhary Ji)* for being the one who always inspired her to achieve the goal.

REFERENCES

[1] C. Stauffer and W.E.L. Grimson; “Adaptive Background Mixture Models for Real-Time Tracking”, IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Vol. 2, pp. 252-258, 1999. DOI:10.1109/CVPR.1999.784637

[2] Change Detection Dataset: Office Dataset: http://www.changedetection.net/ .

[3] D.S. Lee; “Effective Gaussian Mixture Learning for Video Background Subtraction”, IEEE Trans. On Pattern Analysis and Machine Intelligence, Vol. 27(5), pp. 827-832, May 2005.

[4] C.R. Jung; “Efficient Background Subtraction and Shadow Removal for Monochromatic Video Sequences", IEEE Transactions on Multimedia, Vol.11(3), pp. 571-577, April 2009.

[5] OTCBVS Benchmark Dataset: OSU-Pedestrian: http://vcipl-okstate.org/pbvs/bench/.

[6] O. Barnich and M.V. Droogenbroeck; "Vibe: A Universal Background Subtraction Algorithm for Video Sequences", IEEE Transactions on Image Processing, Vol. 20(6), pp. 1709-1724, June 2011.

[7] K.K. Ng and E.J. Delp; "Background Subtraction using a Pixel-Wise Adaptive Learning Rate for Object Tracking Initialization", Visual Information Processing and Communication II, Proceedings of SPIE, Vol. 7882, Jan 2011. DOI:10.1117/12.872610

[8] M. Piccardi; "Background Subtraction Techniques: A Review", International Conference on Systems, Man and Cybernetics, IEEE, Vol. 4, pp. 3099-3104, Oct 2004.

[9] D.K. Yadav, L. Sharma and S.K. Bharti; "Moving Object Detection in Real-Time Visual Surveillance using Background Subtraction Technique", 14th International conference on Hybrid Intelligent Systems (HIS), IEEE, pp. 79-84, Dec 2014.

[10] M. Haque, M. Murshed and M. Paul; "On Stable Dynamic Background Generation Technique using Gaussian Mixture Models for Robust Object Detection", 5th Int. Conference on Advanced Video and Signal Based Surveillance, IEEE, pp. 41-48, Sept 2008.

[11] D.K. Yadav and K. Singh; "Moving Object Detection for Visual Surveillance Using Quasi-Euclidian Distance", IC3T-2015, LNCS, Advances in Intelligent Systems and Computing Series, Springer, Vol. 381, pp. 225-233, Sept 2015.

[12] R. Zhang, W. Gong, V. Grzeda, A. Yaworski and M. Greenspan; "An Adaptive Learning Rate Method for Improving Adaptability of Background Models", IEEE Signal Processing Letters, Vol. 20(12), pp. 1266-1269, Dec 2013.

[13] X. Zhou, C. Yang and W. Yu; "Moving Object Detection by Detecting Contiguous Outliers in the Low-Rank Representation", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 35(3), pp. 597-610, Mar 2013.

[14] T. Bouwmans; "Recent Advanced Statistical Background Modeling For Foreground Detection: A Systematic Survey", Recent Patents Computer Science, Vol. 4(3), pp. 147-171, 2011.

[15] L.S. Sharma, D K. Yadav and S.K. Bharti; "An improved method for visual surveillance using background subtraction technique", 2nd Int. Conf. on Signal Processing and Integrated Networks, IEEE, pp. 421-426, Feb 2015.