An improved Gaussian Mixture Model with post-processing for multiple object detection in surveillance video analytics

Original Scientific Paper

Fancy Joy
Department of Computer Science
Sri Ramakrishna College of Arts and Science, Coimbatore
fancyjoy1@gmail.com

V. Vijayakumar
Department of Computer Science
Sri Ramakrishna College of Arts and Science, Coimbatore
Veluvijay20@gmail.com

Abstract – *Gaussian Mixture Model (GMM) is an effective method for extracting foreground objects from video sequences. However, GMM fails to detect the object in challenging scenarios like the presence of shadow, occlusion, complex backgrounds, etc. To handle these challenges, intrinsic and extrinsic enhancement is required in traditional GMM. This paper presents a novel framework that combines improved GMM with postprocessing for multiple object detection. In the proposed system, GMM with parameter initialization is considered an intrinsic improvement. Video preprocessing and postprocessing are considered extrinsic improvements. Integration of morphological operation with GMM helps for better segmentation than traditional GMM, and it also helps to increase detection performance by reducing false positives. Video preprocessing is the process of noise removal that prepares input video ready for further processing. In the final step gradient of morphological operations is used for postprocessing. The proposed approach was tested on challenging surveillance video sequences from benchmark datasets such as PETS 2009 and CD 2014(Change Detection). The experimental results are compared using ground truth and performance evaluation metrics. The results show that the proposed approach performs better than GMM, and the method can detect the object effectively even in illumination variation and partial occlusion.*

Keywords*: gaussian mixture model, background subtraction, multiple object detection, postprocessing*

1. INTRODUCTION

In this era, video surveillance systems have a prime role in implementing public safety and security. Even though intelligent surveillance analytics systems exist, peoples still rely on humans for monitoring surveillance systems. The object detection process in surveillance video plays a significant part in video analytics. In this real-world, detection of objects has a wide variety of applications, such as theft detection, street monitoring, intruder detection, traffic monitoring etc. [1]. The significant challenges involved in object detection are the sudden illumination variation [2], presence of shadow, long-term occlusion [3], and dynamic background due to natural phenomena (wind, rain, moving trees etc.) [4]. An intelligent surveillance system requires an efficient object detection approach to segment the foreground from the background. The main objective of the proposed method is to integrate the proposed approach with the surveillance system for object detection without human intervention. This paper presents a robust object detection method that can apply in real-world scenarios. The technique uses the Gaussian Mixture Method (GMM) [5] with updated parameters and gradient filter technique as advanced filtering. This approach can detect the object in surveillance videos even with challenging scenarios.

A simplified block diagram of the proposed method shows in Fig.1. The proposed approach introduces a hybrid object detection framework that consists of video preprocessing, background subtraction and postprocessing. GMM has been chosen here as the backbone for detection. Background subtraction uses the gaussian mixture method (GMM) to determine the background from video frames, which helps to segment the foreground object from the background model. The approach modifies the GMM by integrating both intrinsic and extrinsic enhancement.

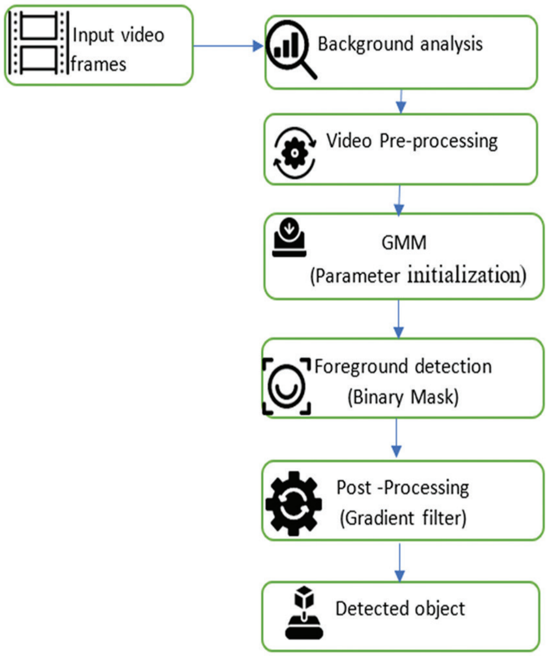


Fig. 1. Block diagram of proposed gradient GMM for object detection

Here GMM with parameter initialization technique as intrinsic improvements. Both preprocessing and postprocessing parts consider extrinsic modifications. GMM initializes its parameters by performing the histogram analysis of training frames. Preprocessing video frames is an essential task that makes the video frames noise-free and ready for advanced processing like detection and tracking. The proposed method uses a bilateral edge-preserving method for preprocessing. Advanced filtering is one of the relevant tasks in proposed GMM-based object detection. A gradient-based morphological filter introduces here, which helps effectively segment objects from the background. The postprocessing phase applies an advanced filter in segmented results from the gaussian mixture model (GMM) and boosts the detection of multiple objects more accurately. In most challenging scenarios, background subtraction and segmentation play a vital role in GMM-based object detection.

The significant contribution of this paper is a fast and improved method for object detection in surveillance monitoring with an advanced filtering process. Advanced filtering introduces a gradient filter mechanism for better segmentation. Hence, the proposed method can detect foreground objects efficiently in various challenging scenarios. The rest of the paper organizes as follows: Section 2 examines the literature regarding object detection and the proposed method discussed in section 3. Section 4 presents the experimental results, and the conclusion is in section 5.

2. REVIEW OF LITERATURE

Object detection is one of the major research areas in computer vision. The simplest method for object detection is background subtraction. Background subtraction includes deriving a background reference image for a video, subtracting each new video frame from the reference image and outputting the result [6]. Many researchers have focused on background subtraction methods to deal with different challenges in object detection. Background estimation, kernel density estimation, subspace learning, background clustering and statistical models are some background modelling methods. Statistical models maintain an excellent balance between performance and computational cost. Gaussian Mixture model [6], kernel density estimation [7] and subspace learning [4] are the commonly used statistical models. Subspace learning methods work with the idea of dimensionality reduction [4] and perform well in varying illumination. Kernel density estimation [7] uses the kernel function to estimate the density of pixels. Gaussian Mixture Model is the most used background subtraction approach. Some researchers improved the GMM to handle the challenges faced by the video surveillance system. The following part reviews various background subtraction approach based on Gaussian Mixture Model and its variations.

Stauffer and Grimson [5] introduced adaptive background mixture models for real-time tracking. This method models each pixel as a mixture of Gaussians and updates the model with an online approximation. The authors also discussed learning patterns of activity used in real-time tracking [8]. The main drawback of their method was that approach learns very slowly in the beginning. To overcome the limitation of the above technique, KaewTraKulPong and Bowden [6] improved the adaptive background mixture model with shadow detection. They applied the expectation-maximization method to update the parameters of the existing GMM. But this method also suffers a slow learning rate.

Lee[9] presented adaptive learning rate-based GMM for background subtraction. A dynamic learning rate helps the GMM to achieve convergence faster. The method shows better performance, but it takes more computational time. Zivkovic and Heizden [10] suggested adaptive Gaussians to each pixel to improve GMM. This approach achieves fast execution than GMM. Its single learning nature causes an increased number of false positives and false negatives. Shah et al. [11] also built an adaptive local learning rate for GMM. They proved that this method is efficient in both memory and time, but it is computationally expensive due fixed number of Gaussians. Shah et al. [12] again modify the GMM with an adaptive learning rate. The method derived the learning rate from the dynamic nature of the image. The algorithm uses YUV color space and SURF features for better segmentation. R. Chavan et al. [13] modified GMM by parameter tuning. The authors used EIR with the expectation-maximization algorithm to choose the value for parameters such as threshold and learning rate. The method performs better detection in the dynamic background with sudden illumination variation.

Del Blanco et al. [14] used GMM with postprocessing to improve detection accuracy. They used parametric background subtraction for detecting moving regions in video data. Estimating and fitting ellipses apply here as postprocessing to refine the foreground. Pathan et al. [15] proposed a Gaussian mixture model with a median filter and morphological operation for moving object detection. The method assures complete detection of moving objects. Chen and Ellis [16] improved the gaussian mixture model with varying learning rates and the number of Gaussians. The authors proved that the approach performs well in illumination variation and dynamic background.

Fradi and Dugelay [17] proposed a new approach based on incorporating on uniform motion model into GMM background subtraction. The authors combined improved GMM background subtraction with a uniform motion model and proposed a single framework. The approach has improved accuracy even in complex backgrounds. Video object detection based on GMM was discussed by Fu and Wang [18]. They introduced the combination of foreground and background gaussian distribution for efficient segmentation.

Xiaofeng Lu and Caidi Xu[19] built an improved gaussian mixture model based on mean block images. The approach improves accuracy using a wavelet-based denoising method with a semi-soft threshold used in the foreground detection stage. The author's used mean images of image blocks as background modelling and wavelet-based denoising for foreground detection to improve the speed of GMM.

Chen et al. [20] developed a spatiotemporal gaussian mixture model based on pixel hierarchy. The method used optical flow, spanning trees and super-pixel segmentation. The approach worked well even with sudden or frequent changes in pixel values. Ghedia and Vithalani [21] modified the GMM with parameter initialization and adaptive thresholding. They used intrinsic and extrinsic improvements to GMM for efficient object detection.

From the literature reviews, it is clear that the Gaussian Mixture model requires both intrinsic and extrinsic improvements. Intrinsic improvements aim the parameter optimization and tuning. Extrinsic enhancement such as preprocessing and postprocessing reduces the noise and improves the performance of detection. Hence both intrinsic and extrinsic modification to the GMM is required.

3. PROPOSED METHOD

The proposed object detection method has a hybrid object detection method that uses modified GMM. The gaussian mixture model is the flexible foreground detection approach introduced by Stauffer et al. [8]. To handle the multimodality of the background, the GMM represents each frame pixel using a mixture of normal distributions. GMM first generate the background model and then subtracts the background model from the current frame pixel-by-pixel. Surveillance video background may vary due to various constraints such as illumination, presence of shadow, dynamic background etc. The intrinsic and extrinsic enhancement makes the GMM that handle different conditions. Modified GMM depicts in Fig 2. The proposed object detection method has a hybrid framework that consists of preprocessing, object detection and postprocessing. Preprocessing and postprocessing are extrinsic improvements. Here parameter initialization of the GMM is the intrinsic modification.

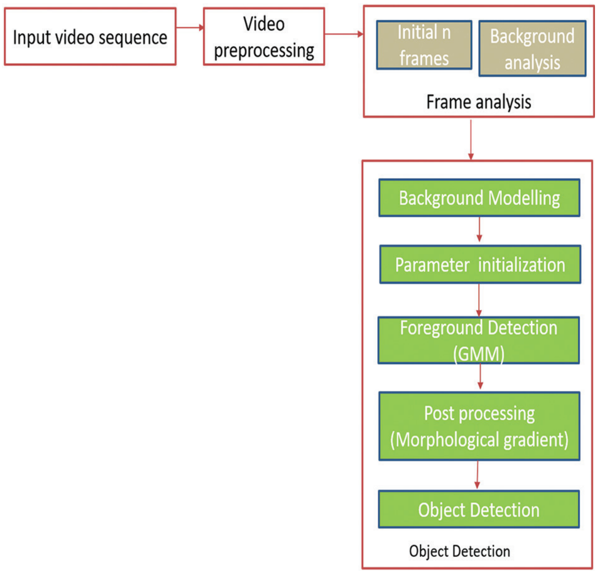


Fig. 2. Proposed Gradient GMM approach

3.1. PRE-PROCESSING

Pre-processing is one of the significant tasks in video processing. Video pre-processing improves the quality of video frames, thereby reducing or eliminating noise. Hence it can increase the performance of further video processing applications such as object detection, tracking and monitoring etc. In the proposed system bilateral filter is used for preprocessing. A bilateral filter converts the video frames into a smoother version by removing noise and fine details[22]. The main advantage of this filter is that it can preserve the edges of objects without blurring.

Let 'I' be the input image, then the bilateral filter output image $I^{filtered}$ is defined as

$$I^{filtered}(q) = \frac{1}{W_s} \sum_{q_i \in \rho} I(q_i) r_k(\| I(q_i) - I(q) \|)\, s_k(\| q_i - q \|) \quad (1)$$

Here the intensity values of each pixel in the input image are replaced with a weighted average of intensity values of the adjacent pixel. The bilateral filter uses spatial and range kernel intensity values to preserve the sharp edges. In the equation w_s denotes weight normalization and can be defined as

$$w_s = \sum_{q_i \in \rho} r_k(\| I(q_i) - I(q) \|) s_k(\| q_i - q \|) \quad (2)$$

Here I and $I^{filtered}$ indicate input image and filtered image respectively and 'q' denotes the coordinates of the current pixel to be filtered. ρ indicates the window that is centred in q i.e. $q_i \in \rho$ refers to another pixel in the image. r_k is the range kernel, which performs smoothing differences in intensities of pixels. s_k is the spatial kernel which performs smoothing differences in spatial coordinates of pixels[22].

3.2. OBJECT DETECTION

Object detection in the proposed system divides into two tasks: frame analysis with parameter initialization and background subtraction. The flowchart of frame analysis with parameter initialization depicts in figure 3. This step analyzes the initial training frames of the video dataset. Analyze various frame intensities using histogram analysis to understand the nature of the background and declare the set of parameter vectors according to it. Here the approach considers two significant parameters:α(learning rate) and T(threshold). Here α is the learning constant for background subtraction, and T is the threshold, which means the minimum proportion of the frame that regards the background. Initial training using GMM started using default parameters and then compares its binary mask results with ground truth. If the parameter gives a better result, initialize the GMM with the same α and T or check the next set of parameters. Repeat the steps until the final set of parameter values. Histogram analysis and background model comparison use here to choose the parameters for a particular dataset.

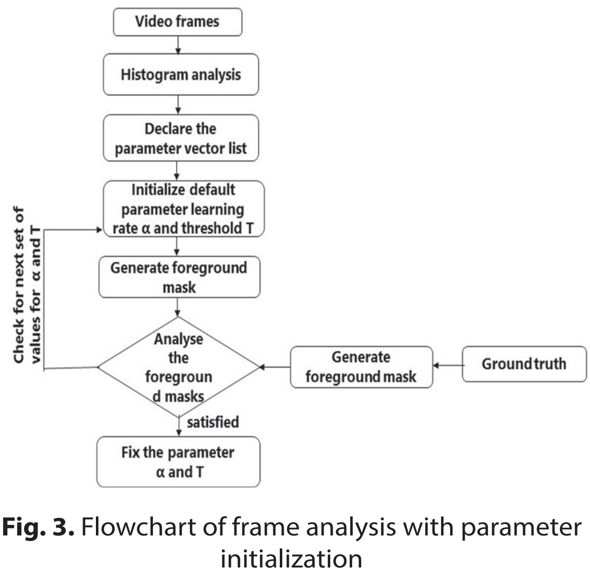


Fig. 3. Flowchart of frame analysis with parameter initialization

The significant task in this object detection step is background subtraction. For background subtraction, the proposed method uses a statistical probability-based approach. Gaussian Mixture Model (GMM) is the widely used background subtraction method. Here probability density function of the gaussian mixture model is used to represent each pixel of video frames. GMM used a mixture of gaussian distribution to model the {Q1, Q2,.........Qs} of each pixel. To model the history of each pixel, GMM uses the mixture of K Gaussian distribution.

The probability of the current pixel is represented as

$$P(Q_s) = \sum_{i=1}^{K} w_{i,s} * \eta(Q_s, \mu_{i.s}, \Sigma_{i,s}) \quad (3)$$

Where

K - number of gaussian distribution

$w_{i,s}$ - weight of i[th] gaussian in the mixture at time s

$\mu_{i.s}$ -mean of pixel intensities

$\Sigma_{i,s}$ -covariance matrix of pixel intensities

η refers the gaussian probability density function which is defined by

$$\eta(Q_s, \mu_{i.s}, \Sigma\,) = \frac{1}{2\Pi^{\frac{m}{2}}|\Sigma|^{\frac{1}{2}}}\; e^{-1/2(Q_s-\mu_s)^{T-1}}\Sigma(Q_s - \mu_s) \quad (4)$$

Here $w_{i,s}$ can be defined as

$$w_{i,s} = (1-\alpha)w_{i,s} + \alpha(\mu_{i.s}) \quad (5)$$

Where α is the learning rate and $\mu_{i.s}$ is the mean value. Every new pixel Q_{s+1} is compared against the previous K gaussian distribution of the pixel until a match is found. Mean μ is set to 1 if the background model matches otherwise μ assigned to 0.

$$\mu_{i.s} = \begin{cases} 1\,, If\ the\ background\ model\ matches \\ 0, otherwise \end{cases} \quad (6)$$

If the distribution is found unmatched then the μ_s and σ_s^2 parameters remain the same. If the distribution finds a match, then the parameters μ_s and σ^s_2 are updated based on previous distribution values.

$$\mu_s = (1-\rho)\mu_{s-1} + \rho(Q_s) \quad (7)$$

$$\sigma_s^2 = (1-\rho)\sigma_{s-1}^2 + \rho(Q_s - \mu_s)^T\,(Q_s - \mu_s) \quad (8)$$

Where ρ is the second learning rate. It can be represented as

$$\rho = \alpha\eta\left(\frac{Q_s}{\mu_i}, \sigma_i\right) \quad (9)$$

The primary step in every background subtraction model is background modelling. This step compares the current frame to the background frame and then classifies the pixels as foreground and background. Here the first 'M' distribution is used for background modelling.'M' can be defined as

$$\mathrm{M} = \mathrm{argmin}_m\left(\sum_{i=1}^{m} w_i > T\right) \quad (10)$$

Where T is the threshold value, which is the measure of the minimum portion of data that should be considered as background.

3.3. POST-PROCESSING

The vital role in the proposed architecture lies in the postprocessing phase. Gradient morphological operations use here to process the video frames according to their shapes. There are mainly two kinds of morphological operations; dilation and erosion. The structuring element is the essential element of both dilation and erosion. The structuring element is a matrix that processes the pixel of the image and its neighbourhood.

Dilation operation adds the pixel to an image boundary, whereas erosion removes the pixel from the image boundary[21]. In this morphological gradient operation, the gradient is the difference between dilation and erosion of the image. The gradient operation of an image results in the contrast intensity in adjacent pixels, which will help for better segmentation and edge detection of an object. Hence the integration of gradient operation with the gaussian mixture model can improve the object detection performance. Let I(x,y) be the input image and S(x,y) be the structuring elements. Erosion of I by structuring element S produces $I \ominus S$. Dilation of the image I by structuring element S is given by $I \oplus S$. Hence the gradient G is defined as

$$\text{Gradient } G = (I \oplus S) - (I \ominus S) \tag{11}$$

4. EXPERIMENTAL RESULTS

4.1. EXPERIMENTAL SETUP

Implementation of proposed method Matlab 2017 with 8GB RAM have used. The datasets used for evaluation were the video sequences from the datasets VISOR [23], CD 2014(Change Detection) [24] and PETS 2009[25] and used several existing methods [8][21][26][27] for comparison were used. The terminologies used in the evaluation matrix were TP, TN, FP and FN.TP indicates the correct object detection, TN denotes the correct object rejection, false positive or false detection of an object represented by FP and FN measures the detection failure. Root Mean Square Error (RMSE) is the difference between the source and result image. Precision measures the ratio between true positives to overall positives. Recall calculates true positives to actual objects. The following evaluation metrics uses here to evaluate the performance of the proposed method.

$$RMSE = \sqrt{\frac{1}{mn}\sum_{i=1}^{m}\sum_{j=1}^{n}(x_{ij} - y_{ij})^2} \tag{12}$$

$$\text{Precision=TP/(TP+FP)} \tag{13}$$

$$\text{Recall=TP/(TP+FN)} \tag{14}$$

4.2. EXPERIMENTAL RESULTS

The results in table 1 indicate the comparison of existing GMM [5]and gradient GMM (proposed) in the aspect of foreground detection. Here benchmark GMM is used for comparison. Column 3 represents the ground truth frame. The results of GMM and gradient GMM has shown in column 4, column5 respectively. From the results, it becomes apparent that the gradient GMM performs better than GMM.RMSE(Root Mean Square Error) of the binary mask results of the GMM and the gradient GMM with the actual image has shown in Fig.4.The RMSE value of Gradient GMM is less when compared to GMM, which means that the gradient GMM results more similar to the actual image than other.

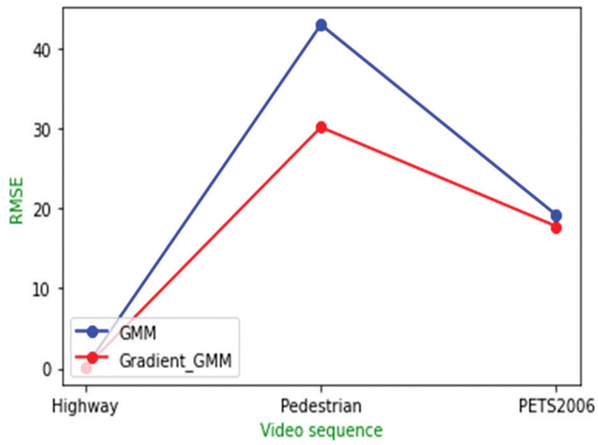


Fig.4. RMSE between the resultant frame of GMM [5] and the proposed method

Fig.5 and Fig.6 represent precision and recall graphs. The precision and recall of Nguyen [26], Poppe's method [27] and the proposed method have displayed in these graphs. The methods Nguyen [26] and Poppe's method [27] used the non-GMM method and used the same dataset for their experiment. The authors used change detection video sequences with various challenges for comparison. According to the precision graph Fig.5, the proposed method outperforms when compared to the other two approaches. But all methods result in almost the same precision value in PETS2006 video sequences due to high reflection in the video.

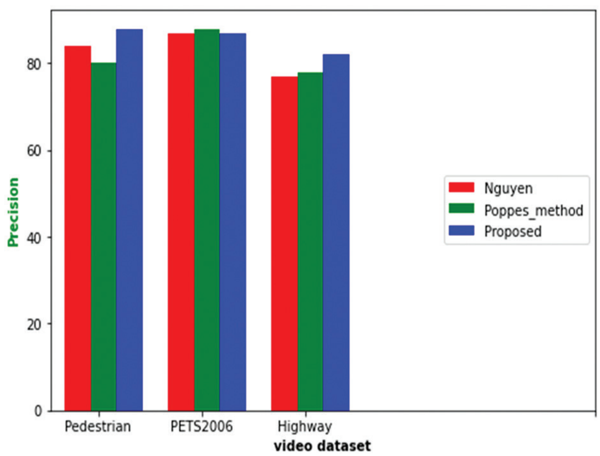


Fig. 5. Precision graph of Nguyen method [26], Poppe's method [27] and the proposed method

Table 1. Comparison of GMM with the proposed system

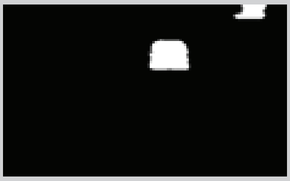
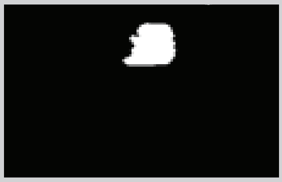
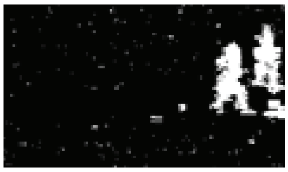
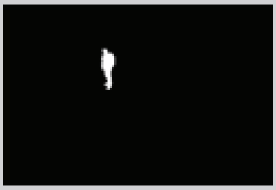
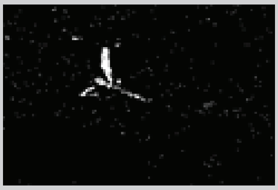
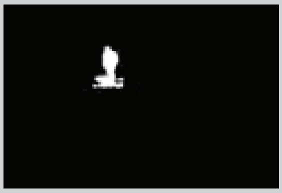
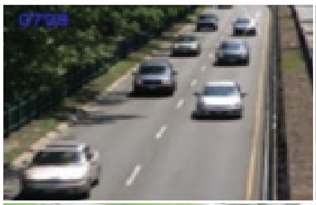
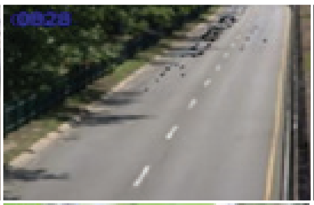
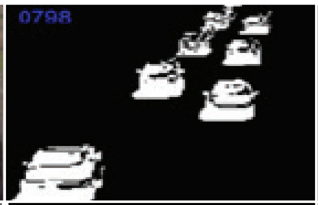
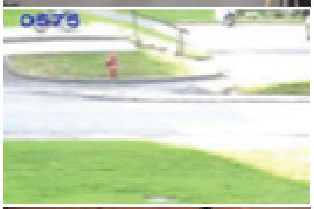
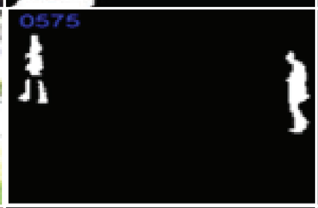
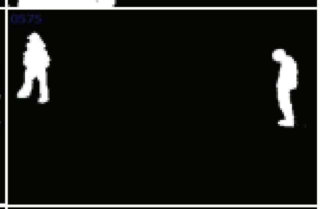
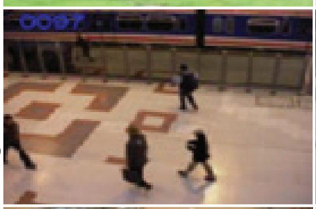
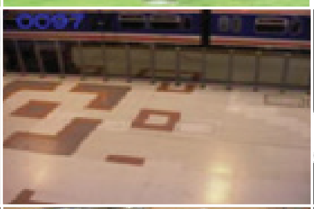
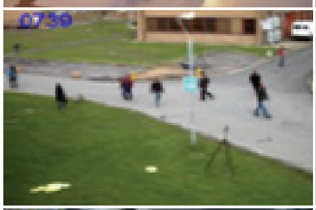
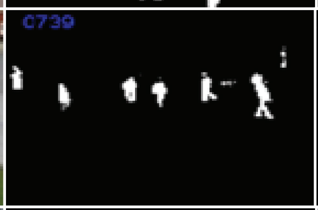
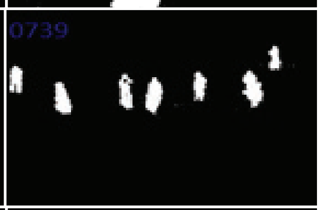
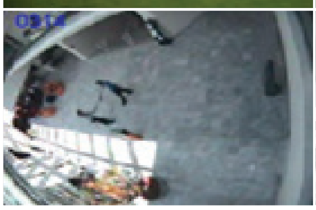
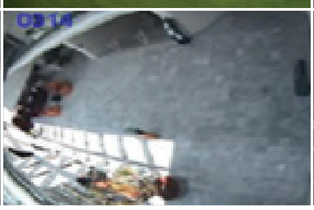
Dataset name	Frame number	Ground Truth	GMM	Gradient GMM (Proposed)
Highway	100			
Pedestrian	420			
PETS2006	300			

Table 2. Comparison of Foreground mask generated by self-adaptive GMM [21] and Gradient GMM(Proposed)

	Original frame and challenges	Background	Foreground Mask N S Ghedia [21]	Foreground Mask (Proposed)
clutter				
high illumination				
high reflection				
surveillance				
Far Field				

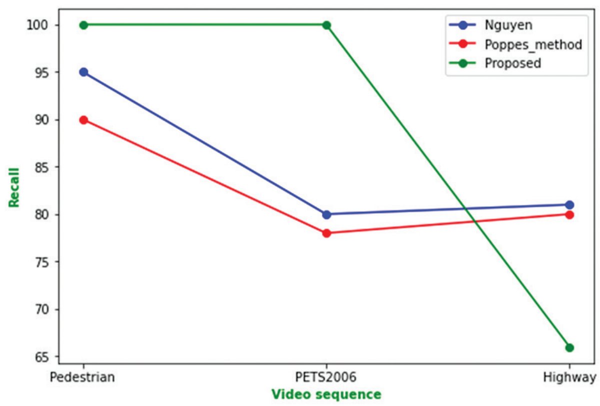


Fig. 6. Recall graph of Nguyen method [26], Poppe's method [27] and the proposed method

The recall is the measure of the true positives towards actual positives.Fig.6 shows the recall comparison of the proposed system performance with Nguyen [26] and Poppe's method [27].

The proposed method performs well in pedestrian and PETS 2006 video sequences compared to the other two methods [26][27]. But recall is very low in Highway sequences (dynamic background). The complex background motions may lower the recall value of the proposed method.

Table 2 represents the foreground mask comparison of various frames with the improved GMM method Ghedia [21] The method Ghedia[21] used a self-adaptive improved GMM method with the same experimental dataset. The table represents the frames from various challenging scenarios such as illumination variation, presence of shadow and clutter etc. In table 2, column1 represents the original frame, column2 shows the background frame, column3 displays the foreground mask of the self-adaptive GMM method Ghedia[21] and the last column shows the foreground mask of the proposed system. The table evidence that the foreground mask of the proposed method is better than the method of Ghedia[21]. From the above comparisons and results, it is clear that Gradient GMM performs more correctly than GMM. But Gradient GMM is less accurate in the case of high reflection.

5. CONCLUSIONS

This paper introduced a gradient GMM method for multiple object detection in surveillance videos. The Proposed gradient GMM for object detection performs better under various challenging circumstances. Gaussian Mixture Model is the backbone of the proposed system. Contributions to GMM classifies into intrinsic and extrinsic enhancement. Parameter initialization of gradient GMM is the intrinsic improvement and helps the proposed method for better background subtraction. The preprocessing and postprocessing are extrinsic enhancements to GMM. The morphological gradient filter plays a critical role in foreground detection and eliminating the noise in foreground objects. Integration of the gradient filter adds to the efficiency and robustness of the proposed GMM method. The experimental results proved that the proposed technique performs better than GMM. Gradient GMM can detect the object even in illumination variation, shadows, occlusion, and dynamic background to a range. But the method results in false positives in the case of moving backgrounds such as high reflection, high dynamic motion etc. A further modification is required to handle complex backgrounds and long-lasting occlusion.

6. ACKNOWLEDGEMENT

We thank Sri Ramakrishna College of Arts and Science for giving support for doing this research work.

7. REFERENCES:

[1] J. Fancy, V. Vijayakumar, "A Review on Multiple Object Detection and Tracking in Smart City Video Analytics", International Journal of Innovative Technology and Exploring Engineering, Vol. 8, No. 2S2, 2018, pp. 2278-3075.

[2] D. Zhou, V. Frémont, B. Quost, Y. Dai, H. Li, "Moving object detection and segmentation in urban environments from a moving platform ", Image and Vision Computing, Vol. 68, 2017, pp. 76-87.

[3] S. Li, C. Zeng, Y. Fu, S. Liu, "Optimizing multi-graph learning based salient object detection", Signal Processing: ImageCommunication, Vol. 55, 2017, pp. 93-105.

[4] T. Bowmans, "Subspace learning for background modelling: a survey", Recent Patents on Computer Science, Vol. 2, No. 3, 2009, pp.223-234.

[5] C. Stauffer, W. Eric L. Grimson, "Adaptive background mixture models for real-time tracking", Proceedings of the IEEE computer society conference on computer vision and pattern recognition, Vol. 2, Fort Collins, CO, USA, 23-25 June 1999.

[6] P. KaewTraKulPong, R Bowden, "An improved adaptive background mixture model for real-time tracking with shadow detection", Proceedings of the 2nd European workshop on advanced video-based surveillance systems: AVBS01, September 2001.

[7] A. Elgammal, D. Harwood, L. Davis, "Non-parametric model for background subtraction", Lecture Notes in Computer Science, Springer Berlin, 2000, pp.751-767.

[8] C. Stauffer, W. Eric L. Grimson, "Learning patterns of activity using real-time tracking", IEEE transactions on pattern analysis and machine intelligence, Vol. 22, No. 8, 2000.

[9] D. S. Lee, "Effective Gaussian mixture learning for video background subtraction", IEEE Transaction on Pattern Analysis Machine Intelligence, Vol. 27, Issue.5, 2005, pp.827-832.

[10] Z. Zivkovic Z, F. van der Heijden, "Efficient adaptive density estimation per image pixel for the task of background subtraction," Pattern Recognition Letters, Vol. 27, No. 7, 2006, pp. 773-780.

[11] M. Shah, J. Deng, B. Woodford B, "Localized adaptive learning of mixture of Gaussians models for background extraction", Proceedings of the Image and Vision Computing Conference, New Zealand, 2010, pp.1-8.

[12] M. Shah, J. Deng, B Woodford, "Video background modelling: recent approaches, issues and our proposed techniques", Machine Vision application, Vol. 25, No. 5, 2014, pp.1105-1119.

[13] R Chavan, S. R. Gengaje, S. Gaikwad, "Multi-Object Detection Using Modified GMM-Based Background Subtraction Technique", Proceedings of the International Conference on ISMAC in Computational Vision and Bio-Engineering ISMAC, 2019, pp.945-954.

[14] C. R del-Blanco, F Jaureguizar, N. García, "An efficient multiple object detection and tracking framework for automatic counting and video surveillance applications", IEEE Transaction on Consumer Electronics, Vol. 58, No. 3, 2012.

[15] I. Pathan, C. Chauhan, P. Kathiriya, "Real-time moving object detection and tracking using adaptive gaussian mixer model and Lucas -Kanade Method", International Journal of Engineering Science and Computing, Vol. 6, No. 5, 2016.

[16] Z. Chen, T. Ellis, "A self-adaptive gaussian mixture model", Computer Vision and Image Understanding, Vol. 122, 2014, pp. 35-46.

[17] H. Fradi, J. L. Dugelay, "Robust foreground segmentation using improved gaussian mixture model and optical flow", Proceedings of the International Conference on Informatics, Electronics &Vision, Dhaka, Bangladesh, 18-19 May 2012, pp. 248-253.

[18] Z. Fu, L. Wang, "Video object detection algorithm based on gaussian mixture models", Advanced Materials Research, Vol. 721, 2013, pp. 587-590.

[19] X. Lu, C. Xu, "Novel Gaussian mixture model background subtraction method for detecting moving objects", Proceedings of the IEEE International Conference of Safety Produce Informatization, Chongqing, China, 10-12 December 2018, pp. 6-10.

[20] M. Chen, X. Wei, Q. Yang, Q. Li, G. Wang, M.-H. Yang, "Spatiotemporal GMM for Background Subtraction with Superpixel hierarchy", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 40, No. 6, 2018.

[21] N. S. Ghedia, C. H. Vithalani, "Outdoor object detection for surveillance based on modified GMM and adaptive thresholding", International Journal of Information Technology, Vol. 13, 2020, pp.185-193.

[22] S. Paris, P. Kornprobst, J. Tumblin, F. Durand, "Bilateral filtering: Theory and applications", Foundations and trends in Computer graphics and Vision, Vol. 4, No. 1, 2008, pp. 1-73.

[23] Visordataset, https://aimagelab.ing.unimore.it/ (accessed:2022)

visor/video_categories.asp (accessed: 2022)

[24] Change detection dataset, https://www.kaggle.com/datasets/fc82cc044b7e90db502e947e-3a4d301a0ff2c498a38b75522543304a40c764f5 (accessed: 2022)

[25] PETS 2009 dataset, http://cs.binghamton.edu/~mrldata/ pets2009 (accessed: 2022)

[26] M. H Nguyen, T. L. Vuong, D. N. Nguyen, D. V. Nguyen, H. Le, T. Nguyen, "Moving object detection in compressed domain for high-resolution videos", Proceedings of the Eighth International Symposium on Information and Communication Technology, 2017.

[27] S Bruyne, C. Poppe, S Verstockt, P Lambert, R V D Walle, "Estimating motion reliability to improve moving object detection in the H.264/AVC domain", Proceedings of the IEEE International Conference on Multimedia and Expo, New York, NY, USA, 28 June - 3 July 2009, pp. 290-299.