

ABSTRACT

Robust Visual Monitoring system that detects and analyze activities based on human, vehicle or animal detection and tracking assures good surveillance. This paper focuses on two dimensional object detection and tracking using Modified Gaussian mixture model (GMM), Adaptive threshold and prediction filtering. Our proposed approach improves learning rate of model for the various constraint like clutter background, light variations, slow and fast moving objects, entering or removed objects, partial occlusion. Our aim is to design such an algorithm which works well on outdoor video sequences.

KEYWORDS: visual monitoring system, adaptive thresholding, background modelling, foreground detection, mixture of Gaussians.

I. INTRODUCTION

Video surveillance, monitoring and robot perception are the most researched area in Image processing that accept video sequences- Gray scale or Colour and detects and tracks objects in a series of sequences. Real time visual monitoring system helps in different application such as traffic monitoring, human computer applications, vehicle navigation, automated surveillance and helps in identifying object behaviour. So, with wide range of real time analysis moving detection and tracking in a video sequences is a most researched and growing area in computer vision. Video analysis generally focuses on the study of the behaviour of the objects, identification and location of the objects in video sequences [5].

The tracker task is to identify the object trajectory in the subsequent frames. Object tracker provides complete region occupied by the object. Robust and accurate tracking depend on robust background modeling and foreground detection. For accurate tracking major issues are to deal with objects that are non-rigid and fast moving, occlusions as well as abrupt changes of objects. For the accurate tracking we should also take certain constraint like constant Velocity, smooth motion or changes, size and number of objects and object silhouette. There are different methods of tracking are available like point, kernel or silhouette tracking. [3]

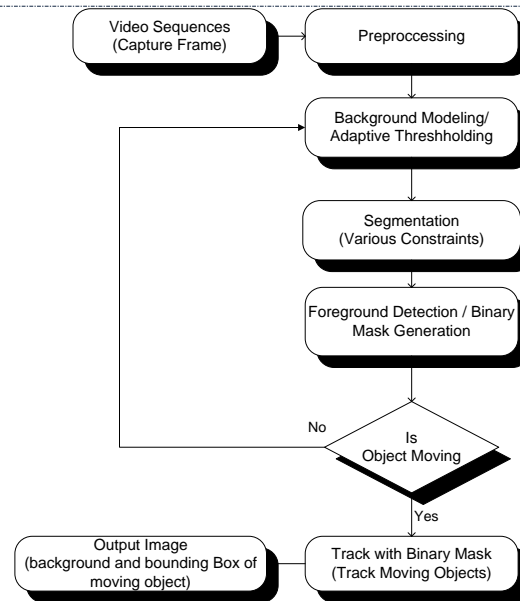


Figure 1 Block diagram of Visual Monitoring system

Object detection uses information from every frame and it is essential for the tracking purpose. Some of the common approaches for the object detection are point detectors, background subtraction, temporal frame differencing and optical flow [3,4]. Generally object detection methods also adopt certain preprocessing approaches like image averaging and morphology for the reduction of the resolution and noise in object space or region. Sometimes instead of color image gray scale image is used to adopt fast changes. It has also been noticed that due to clutter or non stationary background false positive increases and algorithm will fail to detect the target. So, low contrast or smoothing will reduce the error by reducing the scattering noise. Background subtraction gives good result for stationary camera and with static background. It cannot handle dynamic environments so, it fails to detect object in non stationary background.

II. RELATED WORK

Statistical probabilistic background modeling approach would be the best approach for robust surveillance system. It is able to handle static and dynamic scenes and easily detect objects in every frame under almost all constraint. Mixture of Gaussian will be the good approach for designing the background model. Gaussian mixture model handles color information of a pixel to classify background and foreground. Figure1 gives better understanding regarding object detection with almost no noise and clear segmentation of motion object. Stauffer and grimpson [1] explained a Gaussian mixture model as a statistical background analysis to detect objects in static as well as in dynamic environments. [2] Proposed a novel approach for object detection using RGB background modeling. Algorithm is able to detect and track multiple objects in dynamic environments. Morphological analysis can be used to take the challenge of non stationary background.[7] proposed an algorithm which handles non stationary background for real time object detection. Based on conditional random field, space and time constraints in the video sequences are included during a discriminative frame work. [9] Proposed a novel approach for object detection using improved Gaussian mixture model. It improves convergence speed and handle non stationary background by initialize and updating the learning model parameters.[10] proposed an adaptive GMM using spatio temporal dependency for real time object detection. It uses spatial dependency and RGB space for shadow removal. [11] Proposed a traditional approach for background subtraction and extract object using adaptive thresholding. For the accurate tracking suggest a recursive particle filtering that depends on pixel difference. [12] Proposed a multi object tracking using traditional background subtraction and noise analysis. Morphology is used for noise removal to handle dynamic scene. Scale invariant feature transformation SIFT is used for accurate tracking of segmented object. It gives faster convergence with less tracking error.[13] Proposed an algorithm for real time object tracking using feature based selection method and background subtraction. Dynamic threshold and morphological process improves tracking efficiency and processing speed up to certain extent.

III. PROPOSED METHOD

For robust surveillance system one must have to segment object clearly that coupled also helps in accurate tracking and for that we have proposed the background estimation approach which can handle stationary and non-stationary backgrounds. It gives excellent tracking accuracy almost in all constraints like slow and fast moving objects, clutter background, sudden changes in illuminations, introducing and removing objects from the frames. The proposed algorithm is shown in Figure 2. It works on three different sections: Background estimation, Foreground detection and Object tracking.

Background Estimation

Background estimation is essential for moving object detection. Most research scholars give focus on adaptive background estimation. Averaging or smoothing filters are generally used for adaptive background estimation.

Probabilistic model

Adaptive algorithm gives fast convergence at the cost of poor accuracy because it will use single threshold for entire sequence, hence it leads to poor estimation. Statistical probabilistic estimation approach will overcome the demerit of poor estimation and it will improve accuracy and robustness. Gaussian mixture model [1] gives excellent statistical estimation for foreground segmentation. Gaussian mixture model defines each pixel as mixture of Gaussians and then update the model and learning parameters.

Uni-variant Gaussian

Each pixel in a video sequence uses Gaussian probability density function each Gaussian can be parameterized by mean and variance.

$$F(X/\mu, \sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad (1)$$

Simple thresholding is required to identify a pixel whether it is related to foreground or background. Single Gaussian is application under the assumption that background is almost static and the foreground is in motion. Such an constraint is an ideal situation for the surveillance but in real time processing the background is no longer be static but it is dynamic and the motion of the object is also remains complex. To handle such a dynamic background single Gaussian will not handle it with the single mean and variance so, instead of single or Uni-variant Gaussian, multi variant Gaussian is preferable.

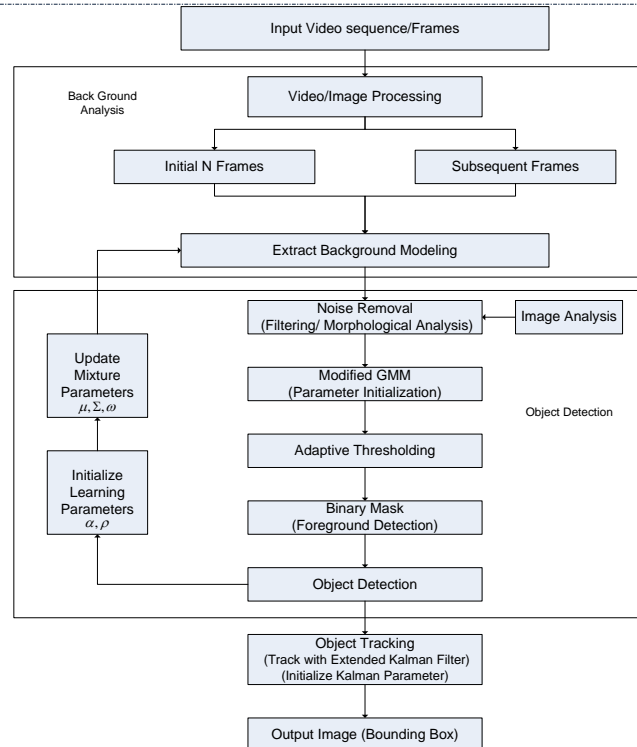


Figure 2 Flow Chart of the proposed method

Multi variant Gaussian: A Gaussian Mixture Model (GMM) is a parametric probability density function. It is represented by a weighted sum of Gaussian component densities. It is also called as a statistical background estimation as a parametric model of the probability distribution [1,14].

$$P(X_t) = \sum_{i=1}^k \omega_{i,t} \cdot \eta(X_t, \mu_{i,t}, \Sigma_{i,t}) \quad (2)$$

Where,

$\omega_{i,t}$ = weighted value of each single model..

k =no. of distributions.

$\mu_{i,t}$ = mean value of the single model (represents the centre of each single peak distribution).

Σ =covariance matrix of the pixel intensities,

$$\eta(X_t, \mu, \Sigma) = \frac{1}{(2\pi)^{n/2} |\Sigma|^{1/2}} e^{-\frac{1}{2}(X_t - \mu)\Sigma^{-1}(X_t - \mu)} \quad (3)$$

Where, $\Sigma_{i,t} = \sigma_{i,t}^2 I$

The mixture parameters are expected from training data using the continuous Expectation-Maximization (EM) algorithm or Maximum A Posteriori (MAP) estimation from a prior model [8].

$$\bar{\omega}_i = \frac{1}{T} \sum_{t=1}^T p_r(i | X_t, \lambda) \quad \text{----- Mixture weight} \quad (4)$$

$$\bar{\mu}_i = \frac{\sum_{t=1}^T p_r(i | X_t, \lambda) X_t}{\sum_{t=1}^T p_r(i | X_t, \lambda)} \quad \text{-----mean} \quad (5)$$

$$\bar{\sigma}_i^2 = \frac{\sum_{t=1}^T p_r(i | X_t, \lambda) x_t^2}{\sum_{t=1}^T p_r(i | X_t, \lambda)} - \bar{\mu}_i^2 \quad \text{---Variance} \quad (6)$$

The distribution of pixel in a frame is characterized by mixture of Gaussians. Now to identify the new pixels which enter into the existing frame is belonging to foreground and background. For that we have to find out the “match” against the existing K Gaussian distribution. This match component will give us an idea regarding the foreground and background pixel. A pixel match can be found by using the mahalanobis distance,

$$\sqrt{((X_{t+1} - \mu_{i,t})^T \cdot \Sigma_{i,t}^{-1} (X_{t+1} - \mu_{i,t}))} \langle k \sigma_{i,t} \quad (7)$$

Once the matched has been found our model parameters will be updated by means of learning parameters α and ρ .

$$\omega_{i,t+1} = (1 - \alpha) \omega_{i,t} + \alpha \quad \text{Updated weight} \quad (8)$$

$$\mu_{i,t+1} = (1 - \rho) \mu_{i,t} + \rho \cdot X_{t+1} \quad \text{Updated mean} \quad (9)$$

$$\sigma_{i,t+1}^2 = (1 - \rho) \sigma_{i,t}^2 + \rho (X_{t+1} - \mu_{i,t+1}) \cdot (X_{t+1} - \mu_{i,t+1})^T \quad \text{Updated Variance} \quad (10)$$

Once the all model parameters have been updated the mixture weight can be normalized by,

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

Foreground detection

The new pixel is still undetermined that whether it is belong to foreground or background. To identify the pixel as a foreground or background one must have to determine the component relevance.

$$B = \arg \min(\sum_{i=1}^b \omega_{i,t} \rangle Tn) \quad (12)$$

Tn = threshold value.

Above equation clearly defines that for the some of the Tn values the components are belonging to background and other than this will be considered as a foreground.

Object Tracking

Detection and tracking are important components in surveillance system. To develop robust video surveillance, robust foreground detection and precise object tracking using recursive approach is required. Tracking complexity can be increases due to different factor such as noisy image, occlusions, complex motion, illumination change, object silhouette etc. object representation and, feature, motion ,shape of the object and object environment will decide the filtering type such as template matching, mean-shift, Kalman filtering, particle filtering and silhouette tracking. [3]

Kalman filtering

Kalman filter is an iterative mathematical process that uses a set of equations and consecutive data inputs to quickly estimate the true position of the object being measured. Figure 3 shows the conceptual block diagram of the kalman filter.

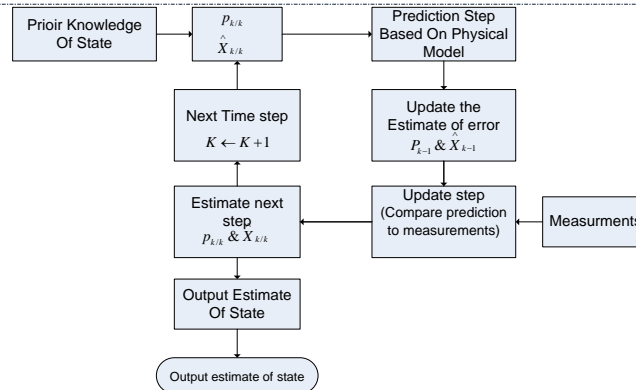


Figure 3 Block Diagram of Kalman filter

The state of the linear system can be estimated by the Kalman filter. Gaussian is used to distribute the same state. [3]. New state can be expected from the calculated value of the Kalman.

The Kalman filter estimates the complete process by means of feedback control. The Kalman estimates the process state at some time and get feedback in the form of measurement. The complete Kalman filter falls into two group(i)time update and (ii) measurement update. The time update is responsible for the projecting the current state and error covariance estimates to get the prior estimates for the next time step. The measurement update are responsible for incorporating a new measurement into a prior estimate to obtain an improved a posterior estimate [17].

Time update equation can be considered as prediction equations.

$$\hat{x}_k^- = A\hat{x}_{k-1} + Bu_{k-1} \quad (13)$$

$$\bar{P}_k = AP_{k-1}A^T + Q \quad (14)$$

The measurement equation is known as correction equations.

$$K_k = \bar{P}_k H^T (H\bar{P}_k H^T + R)^{-1} \quad (15)$$

$$\hat{x}_k = \hat{x}_k^- + K_k (z_k - H\hat{x}_k^-) \quad (16)$$

$$P_k = (I - K_k H)\bar{P}_k \quad (17)$$

IV. EXPERIMENT RESULT

The performance evaluation of our proposed algorithm has been tested against the different constraints like illumination variations, partial occlusion, suddenly entering or leaving moving objects from the current frame. We have used the publicly available datasets, video sequences available from the ViSOR [16] and PETS 2000 [15]. We have implemented proposed algorithm on some of our and standard datasets. All experiments are carried out in dual core 2.5GHz processor.





Figure 4 Outdoor sequence

Figure 4 shows the result of the outdoor sequence for Visor [16] dataset with single and multiple moving objects. Proposed algorithm is able to detect and track almost all the moving objects. In above figure the first row shows the original frames, the second row shows the binary foreground mask and the third row shows the tracking result. Figure 5 shows the result from standard dataset PETS 2000 [15] and Visor [16] for the outdoor sequences focuses on near, mid and far field objects. The first PETS sequence is considered as a near and mid field objects. The second sequence provides the near, mid and far filed moving objects. Our proposed algorithm gives better tracking results with near, mid and far field single and multiple moving objects.



Figure 5 near, mid and far filed detection

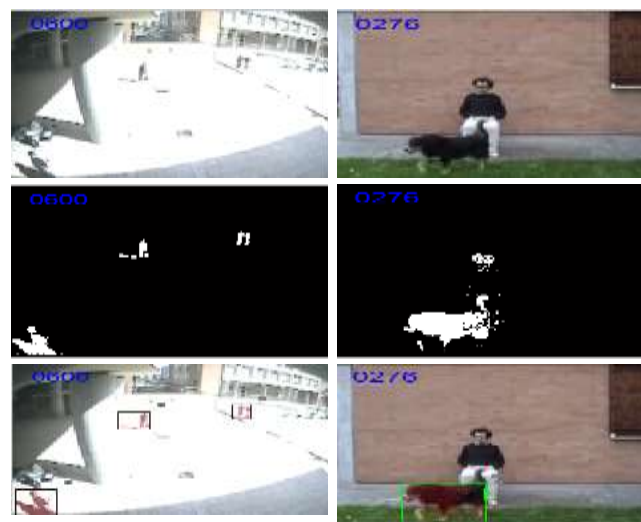


Figure 6 Outdoor sequence

Figure 6 is an outdoor sequence from Visor [16] dataset. First sequence provides large amount of illumination variations and the second sequence provides partial occlusions. Our proposed algorithm gives good tracking results and also able to handle light variations and occlusions accurately.

[Ghedia * et al., 6(7): July, 2017]
ICTM Value: 3.00

Figure 7 shows the plot of detected and tracked objects for outdoor ViSOR sequences with reference to the total object while Figure 8 shows the quantitative analysis of outdoor sequences, compared with the two well known algorithms GMM [1] and KDE [6], results shows that the significant improvement indicates in true positives as increases in FPPI.

$$\text{Recall} = \frac{T_p}{T_p + F_n}$$

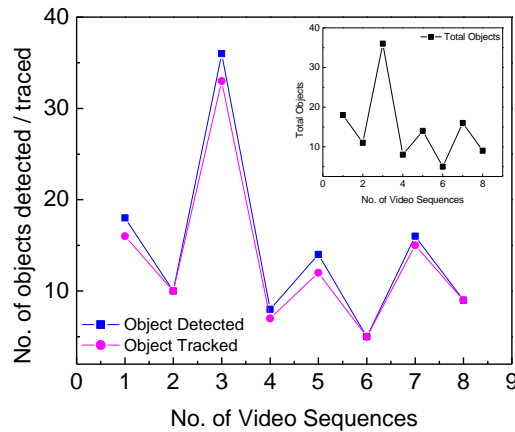


Figure 7 detected and tracked object(Outdoor sequences)

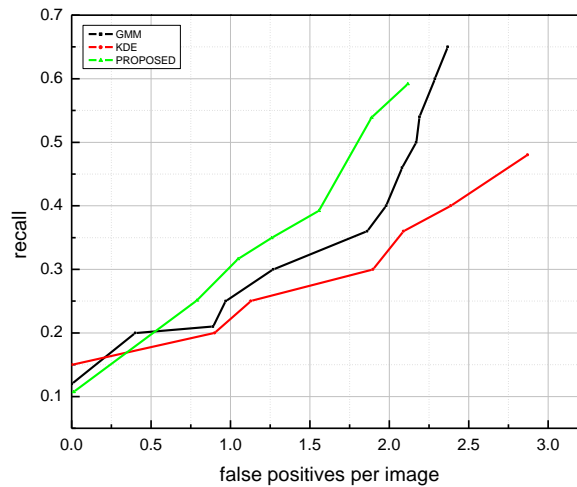


Figure 8 comparative analysis

V. CONCLUSION

This paper has shown the novel statistical approach for the background subtraction. This model can be updated with learning rate α and ρ . The slower learning rate will easily adopt any change in the background so the proposed model is able to handle any dynamic scene. This model gives quite a good result under all circumstances such as illumination variation, occlusions, different silhouette, complex motion etc. Proposed algorithm will detect and track almost all moving object in near, mid and far field. Present approach works well on outdoor environments with assumption that camera is static. If the camera is moving approach fails to detect object. Qualitative and Quantitative results show that proposed approach works well again dynamic background and provides robustness to visual monitoring system.

VI. REFERENCES

- [1] Chris Stauffer and W. Eric L. Grimson, "Learning Patterns of Activity Using Real-Time Tracking," IEEE Transactions On cePattern Analysis and Machine Intellege, Vol. 22, No. 8, pp 747-757 August 2000
- [2] Jong Sun Kim; Dong Hae Yoem; Young Hoo Joo. "Fast and robust algorithm of tracking multiple moving objects for intelligent video surveillance systems" IEEE transaction on Consumer Electronics. Vol . 57, Issue 3, pp. 1165-1170, August 2011.
- [3] A. Yilmaz, O. Javed and M. Shah, "Object Tracking: A Survey", ACM Journal of Computing Surveys, Vol. 38(4), , pp. 1-45,2006.
- [4] Thierry Bouwmans, Fida El Baf. Bertrand Vachon. "Background Modeling using Mixture of Gaussians for Foreground Detection – A Survey", Recent Patents on Computer Science, Bentham Science Publishers, 1 (3), pp. 219-237, 2008.
- [5] A Novel Approach to Detect Foreground in Video Sequences based on Mixture of Gaussians Navneet S. Ghedia, Dr. C.H. Vithalani , Dr. Mehul Raval, Dr. Kiran Parmar, IJSRD - International Journal for Scientific Research & Development| Vol. 3, Issue 06, 2015.
- [6] A Elgammal, D. Harwood, and L. Davis. Non-parametric model for background subtraction. European Conference on Computer Vision, pages 751–767, 2000.
- [7] Wang, Yang. "Real-time moving vehicle detection with cast shadow removal in video based on conditional random field." IEEE Transactions on Circuits and Systems for Video Technology, 19. 3 pp. 437-441(2009).
- [8] Douglas Reynolds, "Gaussian Mixture Models", MIT Lincoln Laboratory, 244 Wood St., Lexington, MA 02140, USA.
- [9] LI Xing-liang, WU Yu-bao, "Image Object Detection Algorithm Based on Improved Gaussian Mixture Model", Journal MULTimedia, Vol. 9, NO. 1, pp. 152-158 , January 2014.
- [10] Katharina Quast, Matthias Obermann ,Andr'e Kaup, "Real-Time Moving Object Detection in video sequences using spatio-temporal adaptive Gaussian mixture models", International Conference on Computer Vision Theory and Applications, pp. 413-418,2010.
- [11] Ka Ki Ng , Edward J. Delp, "Object Tracking Initialization Using Automatic Moving Object Detection", Video and Image Processing Laboratories (VIPER) ,School of Electrical and Computer Engineering, Purdue University West Lafayette, Indiana USA.
- [12] Md. Saidur Rahman, Aparna Saha, Snigdha Khanum, "Multi-Object Tracking in Video Sequences Based on Background Subtraction and SIFT Feature Matching", IEEE computer society, Fourth International Conference on Computer Sciences and Convergence Information Technology, pp 457-462, 2009.
- [13] K Susheel Kumar, Shitala Prasad, Pradeep K. Saroj, R.C. Tripathi, "Multiple cameras using real time object tracking for surveillance and security system", IEEE computer society, Third International Conference on Emerging Trends in Engineering and Technology, pp. 213-218. 2010.
- [14] Nima Seif Naraghi, "A Comparative Study of Background Estimation Algorithms", Eastern Mediterranean University pp. 16-43, September 2009.
- [15] PETS 2000 Dataset.
- [16] R.Vezzani,R.Cucchiara, "Video Surveillance Online Repository (ViSOR): an integrated framework" in Multimedia Tools and Applications, vol. 50, n. 2, Kluwer Academic Press, pp. 359-380, 2010.
- [17] Greg Welch and Gary Bishop. An introduction to the kalman filter, 1995 & 2006.

CITE AN ARTICLE

Ghedia, N. S., Vithalani, C. H., & Kothari, A., Dr. (2017). PERFORMANCE EVALUATION OF MOVING OBJECT DETECTION AND TRACKING ALGORITHM FOR OUTDOOR SURVEILLANCE USING MODIFIED GMM. INTERNATIONAL JOURNAL OF ENGINEERING SCIENCES & RESEARCH TECHNOLOGY, 6(7), 215-223. doi:10.5281/zenodo.823110