COMPARATIVE ANALYSIS OF GMM AND SVM FOR MULTI-PERSON DETECTION

Mrs.Khushali Sheth[1], Asst.Prof.Chita Jani[2]

[1] *ME Student, Computer Engineering, Kalol Institute Of Technology, Gujarat, India*
[2] *Assistant Professor, Computer Engineering, Kalol Institute Of Technology, Gujarat, India*

ABSTRACT

Person detection in a video surveillance system is major concern in real world. several application likes congestion analysis, abnormal event detection, human gait characterization, fall detection, person identification, gender classification and for elderly people. In this algorithm ,We use GMM method for multi object detection because of Gaussian Mixture Model (GMM) model is one such popular method compare to another methods like non-parametric model, optical flow, spatio-temporal etc are not compatible to detect the object due to its quality & limited computational time.GMM are used for detection of moving object. Foreground reconstructed using SVM(support vector machine).SVM are used for detect both moving and non moving object.

Keyword : - *Background subtraction, GMM ,moving and non moving object detection, SVM..*

1. INTRODUCTION

Moving object detection from video frame is one of the most active areas of research in Today world. It is a step for extracting information in various applications such as a automated traffic monitoring (pedestrian detection, vehicle detection),video surveillance (object detection, anomaly detection) etc. Computer vision based motion detection aims to track and detect pixels which is a corresponding to moving objects. The main challenges likes to detect presence of noise and dynamic background in video. Due to these challenges, some unwanted objects are also classified as part of namely, ghost, foreground object and shadows . As mentioned in [1], "A set of connected pixel that not belong to any real moving object but are detected in motion is known as ghost. Shadow is defined as a set of connected background pixels that are detected as moving object due to shadow over them by a moving object shadows and Ghosts may distort the shape of moving objects and lead to misclassification. Hence, both shadows and ghosts are unwanted. Moving object detection techniques are commonly known as Background Subtraction or foreground detection technique.

The Goal of background subtraction algorithm is to distinguish moving objects from static parts of the video frame (called background). In most of the cases, background is not already known that is generated automatically by the background subtraction algorithm. Commonly used background subtraction techniques are - non parametric model ,median filtering ,temporal filtering and Gaussian mixture model [1].When background image or video are available then moving objects can be obtained by subtracting background image from the existing frame. Gaussian mixture model has focus of attention in last few years. GMM has many advantage likes simple implementation, high performance and also give multi model scenario and real time tracking. Many methods for remove the shadows have been proposed in the recent literature [13][14]. In [14], Cucchiara et.al. proposed a method based on hue, motion and saturation for detecting shadows.. Cucchiara et al. [15] provide an good survey on shadow detection. Al-Najdawi et al. presented another survey about the detection of shadow methods in 2012.

After detect the object that are classify using some technique like SVM(support vector machine),HOG(histogram of oriented gradient),LBP(local binary pattern) etc. In this SVM is best compare to other because that provide high accuracy compare to another[1].

2. LITURATURE SURVEY

Paper presented by Shah M E Haque, Subrata Chakraborty and Manoranjan Paul that is proposed to improve human detection in video surveillance The detection process generally occurs in following steps: object detection and object classification. Object detection could be performed by background subtraction and spatio-temporal filtering optical flow[1]. Background subtraction is a popular method for detection of object where it is detect moving objects from the difference between the current frame and a background frame in a block-by-block or pixel by- pixel fashion. and another advantage is accuracy high compare to optical flow and spatio temporal. Background subtraction method divide into five methods: Adaptive Gaussian mixture Model, Non- parametric background Model, Temporal differencing Model, Wrapping background Model, Hierarchical background Model. Adaptive Gaussian mixture model has focus of attention in last few years. first the object are detected then classify the object using texture-based method and shape-based method, motion-based method. Regarding this paper texture based method is better than other because it's accuracy high compare to other method but here computational time is also high this is the disadvantage of this method. SVM,LBP and HOG are the part of texture based method.

Paper presented by Reecha P. Yadav, Vinuchackravarthy Senthamilarasuand Krishnan Kutty, Sunita P. Ugale[2] that proposed system are support vector machine which is used for classify the object. In SVM algorithm HOG are used for feature extraction. its advantage likes to capture the local edge/gradient information; along with a built-in invariance to local illumination condition. Almost all modern detectors employ HOG as a stand-alone detector or in combination with some other features. SVM are used for classify moving and non moving object both. In SVM training and testing set are used for classify the object.

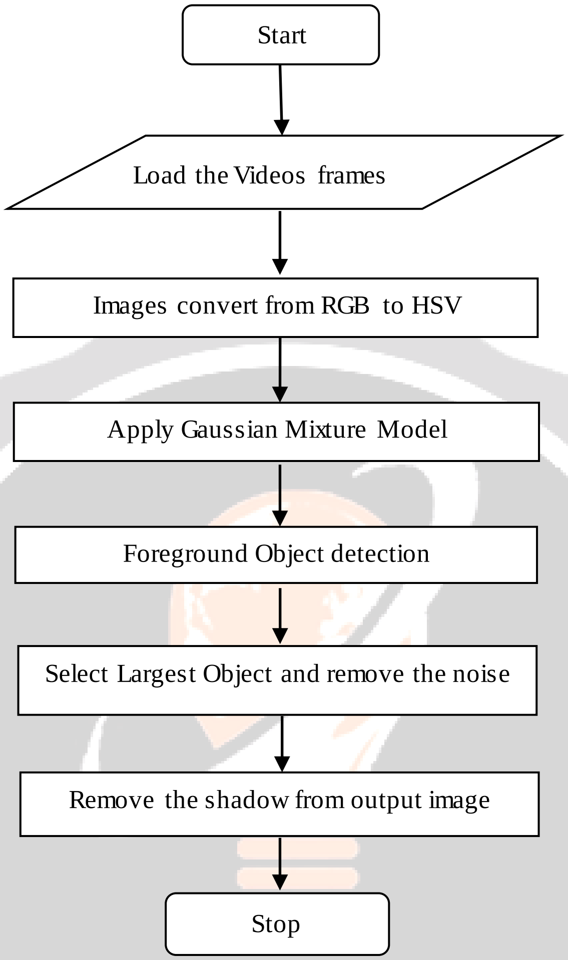
Paper presented by Deepak Kumar Panda, Sukadev Meher [3] that proposed system are Gaussian weight learning rate over a neighborhood which updated the parameters of GMM. The background pixel can be specially in outdoor environment, so in this paper exploited neighborhood correlation of pixels in foreground detection or object detection. Wren et al. [10] modeled background by a single Gaussian distribution methods. It works very well in indoor environment, can deal with gradual or small changes in the illumination and background variation. It fails in the outdoor ,when the background involves multi-modal distributions. To overcome these problem of multi-modal background, Stauffer and Grimson [7], [8] modeled each pixel intensity by a mixture of K adaptive Gaussian distributions method. Good improvements to the original GMM, have been proposed and a numerous survey of the related field and an classification of these improvement can be found in [9].

Paper Presented by Mohamed Taha, Hala H. Zayed, M. E. Khalifa and Taymoor Nazmy [4] That proposed method shows the elimination of the shadows in the frames of image sequences or videos. Its basic idea is based on used a common property of the shadow areas in outdoor scenes. In this, the pixels in shadow areas are darker than the rest of the image especially in the case of strong sunlight. In this algorithm use of Gamma decoding technique to focus the shadow areas and makes them separated through a thresholding process. so in this paper shadow removal technique are used.

Paper Presented by Jingwen Li, Lei Huang, Changping Liu [5] that is use to robust people counting in video. In this various techniques are used like Crowd Segmentation, Perspective normalization, Template Matching, Regression. using this techniques robustly people counting in images scene and video. this paper provide robustness for particular algorithm.

3. BACKGROUNG SUBTRACTION USING GMM

Background subtraction can be expressed in the three basic terms: Background Model Generation, Background subtracted, Background model updates[3]. In Background Model Generation, prepared a statistical model and generated a background image. In Background subtraction to remove the background using particular technique and in Background update model parameter are updated for generation of next background image.



(**Fig-1** Work flow Diagram of GMM)

3.1 Background Model Generation

In real word, Background is generated by system and it is not always know. In this phase the goal is to prepare background image. . In Gaussian Mixture Model based approach each and every pixel is modeled as K Gaussian distributions. The recent history of a pixel *X* at any time instant *t* can be written as {x1,x2,x3……,xt}. The probability of observing the value of current pixel in next frame can be written as [3]:

$$P(X_t)= \sum_{i=1}^{k} \omega_{i,t}\eta_i(X_t , \mu_{i,t},\Sigma_{i,t}) \quad (1.1)$$

where K is the number of Gaussian distribution, and the mean, the covariance matrix are included and is a Gaussian probability density function,

$$\eta\left(X_t , \mu_{i,t},\Sigma_{i,t}\right) = \frac{1}{(2\Pi)^{n/2}\Sigma^{1/2}} e^{\frac{1}{2}(X_t-\mu_t)^T} \Sigma^{(X_t-\mu_t)} \quad (1.2)$$

covariance matrix is assumed to be of the form,

$$\Sigma_k = \sigma_k^2 I \quad (1.3)$$

Parameters of the GMM *i.e.* variance, mean and weight as follows:

$$\mu_{i,t} = \frac{\sum_{i=1}^{t} x_t}{t} \quad (1.4)$$

$$\sigma_t^2 = \frac{1}{t}\sum_{k=1}^{t}(X_k - \mu_{i,t})^2 \quad (1.5)$$

$$\omega_{i,t} = \text{Frame Sequence Number} \quad (1.6)$$

After the initialization of that , initial moving object detection can be made.

3.2 Background Subtraction

These method classifying a pixel as background or foreground, Gaussian distributions of each and every pixel are ordered by ω/σ ratio in descending order. Background pixel occur more frequently than foreground pixels and Intensity value of Background pixel's are not change means remaining the constant. So, if a pixel not match any of the first B distributions it is classified as foreground pixel otherwise background pixel[3],

$$B = \operatorname{argmin}_b \left(\sum_{i=1}^{b} \omega_{i,t} > T\right) \quad (1.7)$$

For each pixel, first step is the i^{th}Gaussian distribution are identify whose mean is closest to X_t. The Gaussian distribution is match if Mahalanobi s distance,

$$\sqrt{\left(\left(X_{t+1} - \mu_{i,t}\right)^T . \Sigma_{i,t}^{-1}(X_{t+1} - \mu_{i,t})\right)} < k\,\sigma_{i,t} \quad (1.8)$$

where k defines a small positive deviation threshold. There may be two cases:

Case 1: In this case pixel is match with one of the K Gaussian and the pixel value is within threshold, then classify pixel as background, otherwise pixel as part of foreground.
Case 2: In this case, the pixel is classified as foreground because of the pixel value does not match with any Gaussian. The pixel's color value will be used in next frame when a pixel is classified as background .Chosen lowest variance and largest weight is as background pixel value when a pixel is classified as foreground. Result of background subtraction phase is a binary mask where foreground=1 and background = 0.

3.3 Background Model Updated

In this two cases are occur:

Case 1: Current pixel value Xt is matched with one of the K Gaussians. For matched components mean is brought closer to current pixel value ,weight is increased and variance is decreased this all to make particular distribution more relevant. These all updates are given by following equations respectively:

$$\omega_{i,t+1} = (1-\alpha)\omega_{i,t} + \alpha \quad (1.9)$$

Where α is a constant learning rate.

$$\mu_{i,t+1} = (1-\rho)\mu_{i,t} + \rho . X_{t+1} \quad (1.10)$$

$$\sigma_{i,t+1}^2 = (1-\rho)\sigma_{i,t}^2 + \rho(X_{t+1} - \mu_{i,t+1}).(X_{t+1} - \mu_{i,t+1})^T \quad (1.11)$$

Where $\rho = \alpha.\eta(X_{t+1}, \mu_i, \Sigma_i)$

Gaussian parameters mean and standard deviation are remains unchanged for unmatched components and only the weight is decreased as:

$$\omega_{j,t+1=}(1-\alpha)\omega_{j,t} \quad (1.12)$$

Case 2: If current pixel value Xt doesn't match with any of the K Gaussian. The parameters are updated as follows:

$$\omega_{k,t+1} = \text{ Low Prior Weight}$$

$$\mu_{k,t+1} = X_{t+1} \quad (1.13)$$

$$\sigma^2_{k,t+1} = \text{ Large Initial Variance}$$

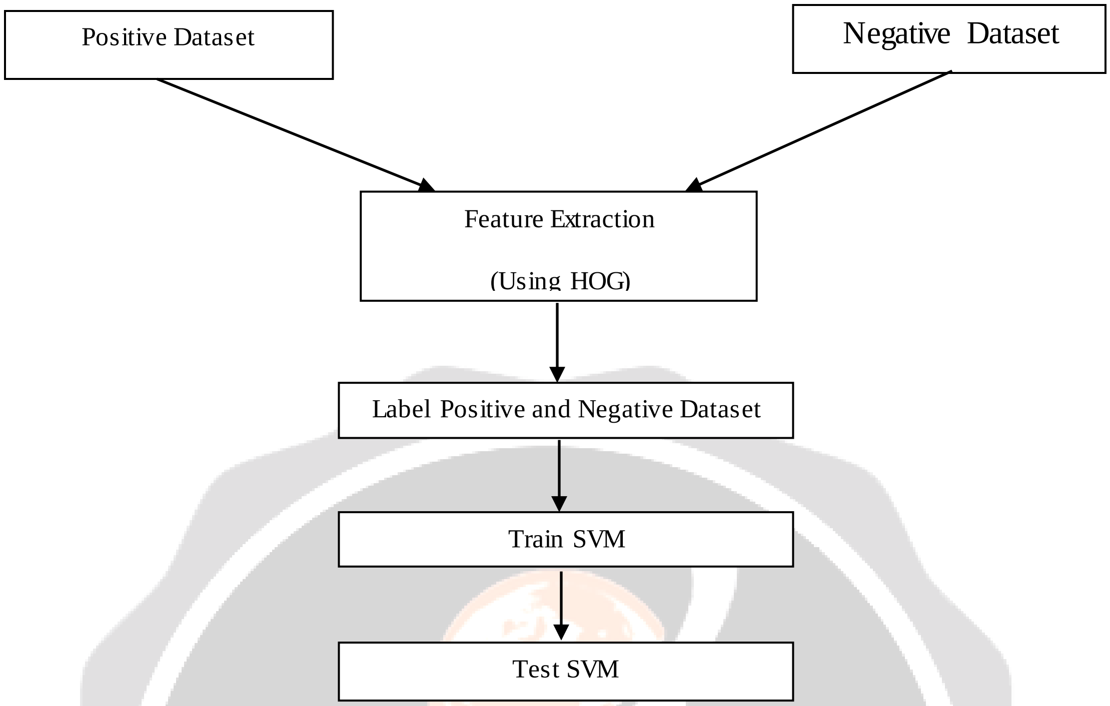
After update the parameters, foreground detection can be made in subsequent frames.

4. CLASSIFICATION USING SVM

SVM are used for detect and classify moving and non moving object. In this feature descriptor are used for extract features from images.

In this find the HOG features of positive and negative dataset. In HOG first find the gradient information of the images. The gradient information along horizontal and vertical direction of the image are calculated by convolving the given gray-scale image with the gradient kernel [-1, 0, 1] and [-1, 0, 1]T respectively. In second step the input image is divided into n cells of equal area. For every pixel in the cell, the magnitude and gradient orientation are calculated and are accumulated into a histogram. The histogram comprises of nine bins obtained by equally dividing the interval 0-180 into bins of interval 20 degree. In last step of HOG is normalization, in this second normal form are used for normalization[2]. After that implementation of SVM start.

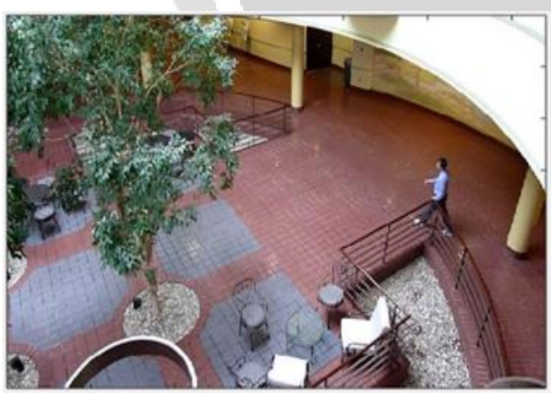
In SVM, positive and negative two dataset are used for classification. In that apply the HOG technique on that and then make training dataset for classify the images. SVM has two variants - Hard margin SVM and Soft-margin SVM. Hard margin SVM requires all data points to be classified correctly into their respective classes. However, the more popular soft-margin approach allow to controlled misclassification of difficult or noisy examples to achieve a maximum margin linear classifier by avoiding over-fitting and hence, the usage of kernels..In SVM apply training dataset on testing dataset and get appropriate output.



(**Fig-2** Workflow Diagram of SVM)

5.RESULT

5.1 Result of GMM Algorithm:



(**Fig-3** Result of video frame extraction)

(**Fig-4** Result of foreground object detection)

5.2 Result of SVM algorithm:

(Fig-5 Positive Dataset)

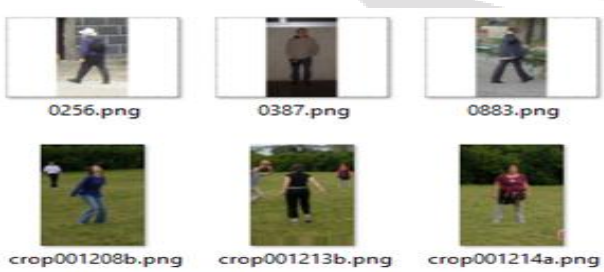
(Fig-6 Negative Dataset)

1) When Background subtracted images are give in input as positive dataset then not give the result

(**Fig-7** Apply Background Subtracted Result)

In this, when give the background subtracted results as input than not give any result means SVM only work with background images. because in this find the gradient of pixel with respect to background images.

2) When Background are not subtracted than SVM give correct result:



(Fig-8 Positive Dataset)

(Fig-9 Negative Dataset)

(Fig-10 SVM Result)

3) SVM also applied on non-moving object

(**Fig-11 Result of SVM apply on still images**)

6. CONCLUSIONS

here. One problem is identify when apply GMM on video and images than produce the background subtracted results.. but when the background subtracted results are give as input in SVM than not get any result. So SVM only work with non background subtracted images. GMM work for only moving object when SVM works for moving and non moving both objects.GMM computational time low compare to SVM but give less security.SVM gives high security compare to GMM and also provide high quality.GMM gives foreground object detection but in that shadow exist so my future work remove the shadow from foreground objects.

7. REFERENCES

[1] Manoranjan Paul*," Human detection in surveillance videos and its applications - a review", Springer open Journal,2013

[2] Reecha P. Yadav, Vinuchackravarthy Senthamilarasuand Krishnan Kutty, Sunita P. Ugale," Implementation of RobustHOG-SVM based Pedestrian Classification ",International Journal of Computer Applications,2015

[3] Deepak Kumar Panda, Sukadev Meher," A Gaussian Mixture Model with Gaussian Weight Learning Rate and Foreground Detection using Neighbourhood Correlation", IEEE Asia Pacific Conference on Postgraduate Research in Microelectronics and Electronics,2013

[4] Mohamed Taha, Hala H. Zayed, M. E. Khalifa and Taymoor Nazmy," Moving Shadow Removal for Multi-Objects Tracking in Outdoor Environments", International Journal of Computer Applications ,Volume 97 No.10, July 2014

[5] Jingwen Li, Lei Huang, Changping Liu," Robust People Counting in Video Surveillance: Dataset and System", IEEE International Conference on Advanced Video and Signal-Based Surveillance, 2011

[6] Tianyu Yang ,Baopu Li,Can Yang ,Max Q.-H. Meng,Guoqing Xu," Condensation-based Multi-person Tracking using an Online SVM Approach ",Proceeding of the IEEE International Conference on Robotics and Biomimetics (ROBIO) Shenzhen, China, December 2013

[7] C. Stauffer and W. E. L. Grimson, "Adaptive background mixture models for real-time tracking," in Computer Vision and Pattern Recognition. IEEE Computer Society, 1999

[8] "Learning patterns of activity using real-time tracking," IEEE Transactions on Pattern Analysis and Machine Intelligence, 2000.

[9] T. Bouwmans, F. E. Baf, and B. Vachon, "Background modeling using mixture of gaussians for foreground detection - a survey," Recent Patentson Computer Science, 2008.

[10] C. R. Wren, A. Azarbayejani, T. Darrell, and A. P. Pentl, "Pfinder: Real-time tracking of the human body," IEEE Transactions on PatternAnalysis and Machine Intelligence, 1997

[11] Cucchiara, R.; Grana, C.; Piccardi, M.; Prati, A., "Detecting moving objects, ghosts, and shadows in video streams," Pattern Analysis and Machine Intelligence, IEEE Transactions on , Oct. 2003

[12] J. Stander, R. Mech, and J. Ostermann. Detection of moving cast shadows for object segmentation. Trans,1999.

[13] Prati, A.; Mikic, I.; Trivedi, M.M.; Cucchiara, R., "Detecting moving shadows: algorithms and evaluation," Pattern Analysis and Machine Intelligence, IEEE Transactions on , July 2003

[14] Zhen Tang, Zhenjiang Miao," Fast Background Subtraction and Shadow Elimination Using Improved Gaussian Mixture Model" IEEE International Workshop on Haptic Audio Visual Environments and their Applications Ottawa Canada ,October 2007