Multiple Person Tracking with Shadow Removal using Adaptive Gaussian Mixture Model in Video Surveillance

Mrs. Khushali Sheth[1] Prof. Lokesh Gagnani[2] Asst. Prof. Chaita Jani[3]
[1]M.E. Student [2]Professor [3]Assistant Professor
[1,2]Department of Computer Engineering [3]Department of Information Technology
[1,2,3]Kalol Institute of Technology, GTU, Gujarat, India

Abstract— Person detection in a video surveillance system is major concern in real world. Several application likes abnormal event detection, congestion analysis, human gait characterization, fall detection, person identification, gender classification and for elderly people. In this algorithm, we use GMM method in background subtraction for multi person detection because of Gaussian Mixture Model (GMM) model is one such popular method this give a real time object detection. There is still not robustly but Multi person tracking with shadow removal fill this gap, in this work, HOG-LBP hybrid approach with GMM algorithm is presented for Multi person tracking with Shadow removal.

Key words: Background Subtraction, GMM, Shadow Removal, Moving Object Detection, HOG-LBP Algorithm

I. Introduction

Moving object detection from video frame is one of the most active areas of research in Today world. It is a fundamental step for extracting information in various applications such as a automated video surveillance (human detection, anomaly detection), traffic monitoring (pedestrian detection, vehicle detection) and control applications (human computer interaction, automated robot bodies).

Computer vision based motion detection aims to track and detect pixels which are a corresponding to moving objects.

The main challenges likes presence of noise and dynamic background (such as falling water and shaking leaves) and illumination changes in video. Due to these challenges, some unwanted objects are also classified as part of namely, ghost, foreground and shadows. As mentioned in [1], "A set of connected pixel that not belongs to any real moving object but is 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. The objectives of a background subtraction algorithm are 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 - 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. Advantages of this method are Simple implementation, good performance, computational time low, capture multi-modal scenarios.GMM is also provide the 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.

II. Literature Survey

Paper presented by Shah M E Haque, Manoranjan Paul and Subrata Chakraborty 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 optical flow, background subtraction and spatio-temporal filtering[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, Warping background Model, Hierarchical background Model. Adaptive Gaussian mixture model has focus of attention in last few years. First the object is detected then classifies 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.

Paper presented by Baopu Li, Can Yang , Qi Zhang, Guoqing Xu [2] they take research papers as documents ,as In this combine the HOG-LBP feature and apply the adaboost algorithm.HOG and LBP are the part of texture method which is used to classify the object. HOG features mainly describe the appearance of pedestrians, which performs poorly when the background is cluttered with noisy edges in video. LBP feature is used to describe the local texture based information, which has advantages of illumination invariance and handle invariance. Because of this, hybrid HOG-LBP algorithms have become the focus of attention in the last few years.

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. [9] modeled background by 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 this problem of multi-modal background, Stauffer and Grimson [10], [11] 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 [12].

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.

III. Background Subtraction using Adaptive GMM

Background subtraction can be expressed in the three basic terms: Background Model Generation: In this, prepared a statistical model and generated a background image. Background Subtraction: In this, generated a foreground mask for every frame. This step is simply performed by compare the background image from the recent frame. Foreground mask is take as a binary mask where foreground =1 and background = 0. Background Model Update: In this, background model parameters are updated for the generation of next background image.

Adaptive GMM can handle the complex background situations like dynamic background and illumination changes.

A. Background Model Generation

In real world, background image is not always known and it is to be generated by the system. Goal of this phase is to prepare a model which could be used for generate the 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:

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

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

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

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.

B. 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,

$$B= \operatorname{argmin}_b \left(\sum_{i=1}^{b}\omega_{i,t} > T\right) \quad (2.1)$$

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,

$$\sqrt{\left((X_{t+1} - \mu_{i,t})^T . \Sigma_{i,t}^{-1}(X_{t+1} - \mu_{i,t})\right)} < k\,\sigma_{i,t} \quad (2.2)$$

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.

C. Background Model Update

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:

$$\omega_{i,t+1} = (1-\alpha)\omega_{i,t} + \alpha \quad (3.1)$$

Where α is a constant learning rate.

$$\mu_{i,t+1} = (1-\rho)\mu_{i,t} + \rho . X_{t+1} \quad (3.2)$$

$$\sigma_{i,t+1}^2 = (1-\rho)\sigma_{i,t}^2 + \rho(X_{t+1} - \mu_{i,t+1}).(X_{t+1} - \mu_{i,t+1})^T \quad (3.3)$$

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

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

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

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

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

$$\mu_{k,t+1} = X_{t+1} \quad (3.5)$$
$$\sigma^2_{k,t+1} = \text{Large Initial Variance}$$

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

IV. SHADOW REMOVAL

In this shadow remove from foreground object using Proposed shadow detection algorithm. In this first normalize the frame then apply Gamma decoding technique on that. If the Gamma value large than one make a image darker and if the gamma value small than one make region lighter. In this whose pixel value larger than one then that is consider as shadow and that is easily separated. Result of this algorithm is shown in the Figure.

In Fig. 1 contain video frame with shadow. Then normalized video frame which result shown in Fig. 2. Then apply the gamma decoding on normalized image and converting the image into the grayscale image. Resulting frame after remove the shadow shown in Fig. 6.

Fig. 1: Video frame that contains the shadow

Fig. 2. Normalized video frame

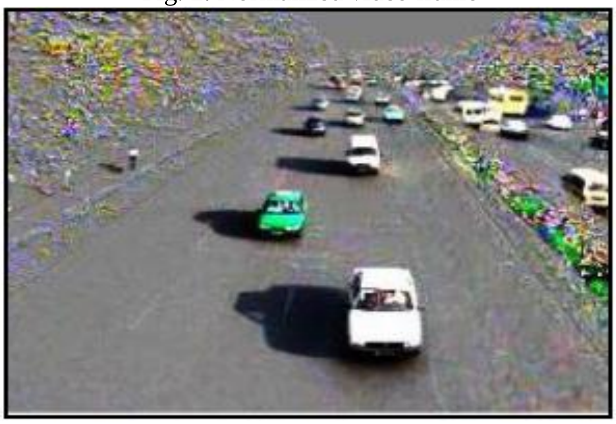

Fig. 3. Video Frame after Gamma Decoding

Fig. 4. converting the image to gray scale

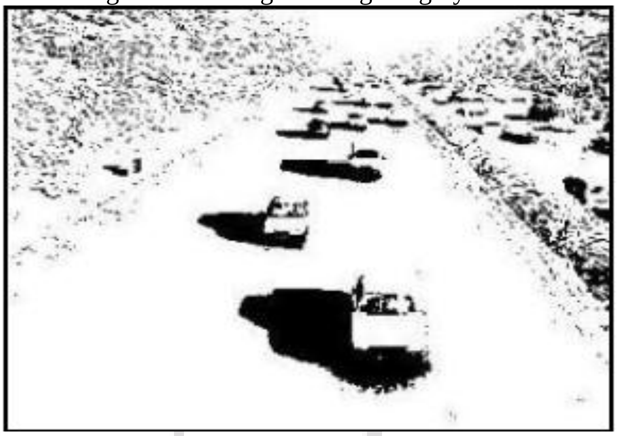

Fig. 5. white pixel indicate the pixel value greater than thresold

Fig. 6: resulting frame after remove the shadow

V. HOG-LBP BASED OBJECT DETECTION

In the first stage, Extract the features of objects. After that, start the training of a classifier. then use Adaboost algorithm to train a strong classifier by training a series of weak classifiers on different samples, which are obtained by changing each sample's weights during training. The strong classifier is then used to identify pedestrians in test sequences. [2] In this HOG , first divide a whole image into small cells. Then we compute the gradient and orientation in each cell. Then compute the LBP for each pixel. The main idea of LBP is that take the center pixel's gray value as a threshold, and compare it with its neighborhood to get the corresponding binary code to represent the local texture. After extracting HOG and LBP features, combine them into a HOG-LBP feature. Then use Adaboost algorithm to train a strong classifier for pedestrian detection. HOG give the global information, while LBP describes the detailed information. So these two feature descriptors can be complementary, which may be more robust to describe pedestrians.

VI. PROPOSED WORK

In proposed work we use background subtraction technique for object detection in that many method are available but we choose adptive gaussian mixture model because this method has many advantage compare to another method and for object classification we choose Texture methods.and after object are classify then remove the shadow using shadow removal algorithm.

Step of proposed work:

1) Input the video frame
2) Detection of object using GMM
3) Normalized the object
4) Use the HOG-LBP features to classify the object
5) After classify the image to detect the shadow from the image.

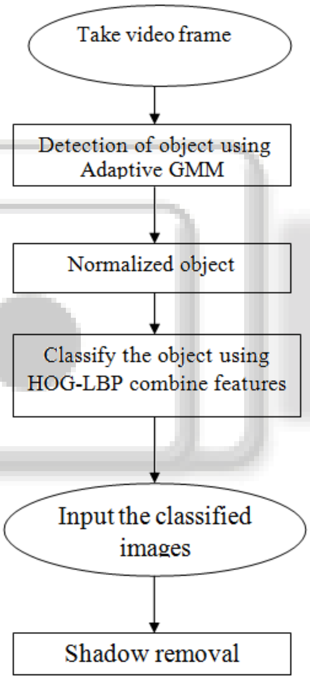

Fig. 7: Process of shadow detection

VII. CONCLUSION

In this paper, we are classifying the person using HOG-LBP based texture method. Then apply the shadow algorithm on images. HOG and LBP features gives high accuracy compare to another. Also use GMM method for object detection. GMM fundamental which is real time and multi modal and also give the good performance. so in this paper we get the good result for multi person detection with shadow removal with low computational ratio.

REFERENCES

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

[2] Baopu Li , Can Yang2,3,4 , Qi Zhang, Guoqing Xu,” Condensation-based Multi-person Detection and Tracking with HOG and LBP”, IEEE International Conference on Information and Automation,July 2014

[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] 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

[7] 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

[8] Subra Mukherjee1,Karen Das2,” an adaptive GMM approach to background subtraction for application in reai time surveillance

[9] 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.

[10] 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

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

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

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

[14] 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.

[15] 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

