A Novel Approach to Detect Foreground in Video Sequences based on Mixture of Gaussians

Prof. Navneet S. Ghedia[1] Prof. Dr. C.H. Vithalani[2] Prof. Dr. Kiran Parmar[3]
[1]Research Scholar [2]Professor and Head of the Department [3]Retired Professor
[2]Department of Electronics & Communication
[1]GTU [2]Government Engineering College, Rajkot [3]Government Engineering College

Abstract— Our aim is to develop a robust visual monitoring system that passively detects moving objects in a specified space and identify the activities of those objects. In the paper, we focus on foreground detection. Mixture of Gaussian is a very well know and friendly approach for background modelling to detect moving objects. Our method improves Gaussian mixture model by continuously updating the mixture parameters. It gives faster update and a smoother object mask. The Gaussian mixture model approach consists of different Gaussian distributions, mean, standard deviation, weight. So, focus is to develop a robust visual monitoring system which can work successfully against illumination variations, clutter background, slow moving objects, sawing trees and objects being introduced or removed from the frame.

Key words: video surveillance and monitoring, adaptive Gaussian mixture model, background model, Gaussian mixture density, foreground detection

I. INTRODUCTION

In the field of image processing and computer vision the concept of higher level application is visual monitoring system for real time object detection and tracking in indoor as well as outdoor environment. Video can be defined as a collection of image frames in sequence. Moving object detection and tracking in video sequence has been increasingly focused in computer vision research due to its wide range of real world applications. Video analysis employs three basic steps to locate, identify and study the behavior of objects in video: detection of moving objects, tracking of detected objects in successive frames and finally recognition of objects behavior. Object tracking is very much useful in different important application areas including motion based recognition, automated surveillance, video indexing, human-computer interaction, traffic monitoring etc.[3].

Tracking is basically the task of estimating the motion path of an object in successive frames of a video. This includes two basic steps: detection of objects and tracking. Object tracking is still a challenging task especially when dealing with objects of complex shapes with complex motion and occlusions. Point detection, background subtraction, segmentation and supervised classifiers are four detection approaches, where point detection means finding interest points in images that have an expressive texture in their respective localities, background subtraction means subtracting current frame from background frame to find the changing region. Segmentation means partitioning the image into perceptually similar regions. Object detection in supervised learning mechanism is carried out by learning different object that can be viewed automatically from a set of examples [3],[12]

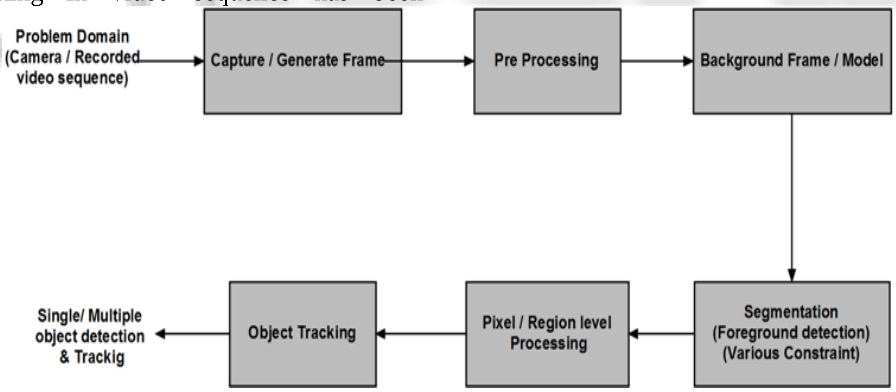

Fig. 1: Basic steps for visual monitoring system

Followings are the key components for the real time tracking, extracting a feature, background image Subtraction and identification of extracted feature. There are two types of approach using for moving object detection first is the region-based approach and second is boundary based approach. The most admired region-based approaches are optical flow and background subtraction. Background subtraction approach detects real time moving objects by subtracting expected background models from images. But background subtraction method takes long time to estimating the background models.[13].

Moving object detection aims at extracting moving objects that are of interest in video sequences with background which can be static or dynamic. Some examples of interesting objects include walking pedestrians and running vehicles, but usually not waving tree leaves. Most algorithms use either temporal or spatial information in the image sequence to perform moving object detection, and the most commonly used feature is pixel intensity. [11]

II. RELATED WORK

In [1] they have used the values of a particular pixel over time are modelled as a mixture of Gaussian distributions. Thus, the background can be modelled by Gaussian mixture model (GMM). Once the pixel wise GMM likelihood is obtained, the final binary mask is generated by thresholding.[8]. In [2] have used RGB background modelling for the real time moving object detection and used morphology for removing noise and blob labelling for real time moving object detection. They predicts the velocity of the moving objects and detects it.in [7] they have used the suitability to background changes is not as satisfying as them especially to some phenomenon like sudden illumination changes. In [9] they have proposed GMM as a background modelling and update the learning parameters. They have proposed algorithm which works well on against illumination variations. Due to adaptive GMM and updating the parameters the convergence speed has shown very much improvement. .[10] proposed spatio-temporal adaptive GMM by using traditional GMM and spatial and temporal dependency. Further the quality of the foreground detection can be improved by removing the shadows. [11] Introduced tradition approach for the foreground detection. They used background subtraction, GMM and use learning parameters that are depends on pixel difference. And get the precise results. [12] Proposed novel approach for the foreground detection using traditional background subtraction and SIFT. [13] have proposed morphology analysis for point processing (noise removal, connectivity) and feature analysis for background detection.

III. PROPOSED METHOD

Our system robustly deal with different challenges like, illumination variations, detection through clutter background, introducing or removing objects from the scenes. Slowly moving objects the longer time to be a part of background and it will preserve the background model. However if objects is remain stationary for the longer duration it will become a part of background without affecting the background model. Fig.2 represents the proposed algorithm for the foreground detection.

A. Segmentation

Segmentation is one of the most crucial and important task in image processing. Segmentation is a fundamental low-level operation on images. If an image is already partitioned into segments, where each segment is a “homogeneous” region, then a number of subsequent image processing tasks become easier. A homogeneous region refers to a group of connected pixels in the image that share a common feature. This feature could be brightness, color, texture, motion, etc. Segmentation is defined as the process of partitioning an image into a set of non-overlapping regions whose union is the entire image. These regions should ideally correspond to objects and their meaningful parts, and background. Most image segmentation algorithms are based on one of two basic properties that can be extracted from pixel values—discontinuity and similarity—or a combination of them. Segmentation of nontrivial images is a very hard problem—made even harder by non-uniform lighting, shadows, overlapping among objects, poor contrast between objects and background, and so on that has been approached from many different angles, with limited success to this date. Many image segmentation techniques and algorithms have been proposed and implemented during the past 40 years and yet, except for relatively “easy” scenes, the problem of segmentation remains unsolved.[5],[6]

There is no underlying theory of image segmentation, only ad hoc methods, whose performance is often evaluated indirectly, based on the performance of the larger system to which they belong. Even though they share the same goal, image segmentation techniques can vary widely according to the type of image (e.g., binary, gray, color), choice of mathematical framework (e.g., morphology, image statistics, graph theory), type of features (e.g., intensity, color, texture, motion), and approach (e.g., top-down, bottom-up, graph-based). There is no universally accepted taxonomy for classification of image segmentation algorithms either. In this chapter, we have organized the different segmentation methods into the following categories:

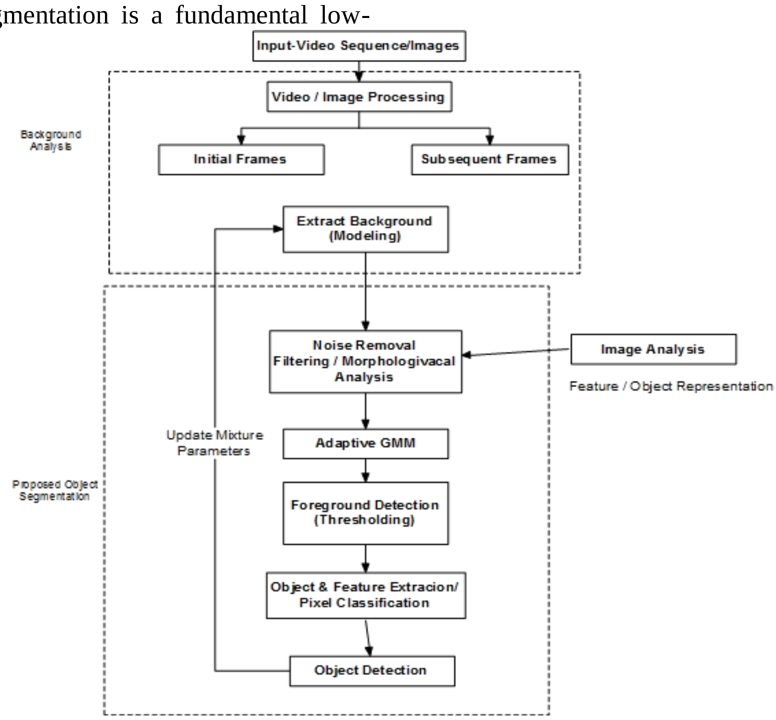

Fig. 2: Proposed algorithms for foreground detection

- Intensity-based methods also known as non-contextual methods, work based on pixel distributions.
- Region-based methods also known as contextual methods, rely on adjacency and connectivity criteria between a pixel and its neighbors.
- Other methods, where we have grouped relevant segmentation techniques that do not belong to any of the two categories above. These include segmentation based on texture, edges, and motion, among others.[5],[6]

B. Sementation Statistics

To understand the role of statistics in image segmentation, let us examine some preliminary functions that operate on images. Given an image f_o that is observed over the lattice Ω, suppose that $\Omega 1 \leq \Omega$ and fi is a restriction of f_o to only those pixels that belong to Ω1. Then, one can define a variety of statistics that capture the spatial continuity of the pixels that comprise fi.
Some of the common statistics are,

- Gaussian Statistics
- Fourier Statistics
- Covariance Statistics
- Label Statistics

Computation of image statistics of type tremendously facilitates the task of image segmentation. [5]

C. Gaussian Mixture Model

Single Gaussian: calculating the average image of a sequence of frames and then subtracting each new input frame and checking the difference values against a predefined threshold is one of the simplest background removal techniques. It requires t frames to estimate the mean μ and the standard deviation σ in each color component separately,

$$\mu(x,y,t) = \sum_{i=1}^{t} \frac{p(x,y,i)}{t}$$

$$\sigma(x,y,t) = sqrt\left(\sum_{i=1}^{t} \frac{p^2(x,y,i)}{t} - \mu^2(x,y,t) \right)$$

Here, p(x,y,t) is the pixel's current intensity value at the location (x,y) at a given time t. After computing the parameters, a pixel is considered as a part of the foreground object based on the following formula,

$$\left|\mu(x,y,t) - p(x,y,t)\right| \langle c.\sigma(x,y,t)$$

Where, c is a constant. Even though this method is capable of adapting to indoor environments with gradual illumination changes, it's not able to handle moving background objects like trees, flags, etc.[14]

D. Gaussian Mixture Model

A Gaussian Mixture Model (GMM) is a parametric probability density function represented as a weighted sum of Gaussian component densities. GMMs are commonly used as a parametric model of the probability distribution of continuous measurements or features in a biometric system, such as vocal-tract related spectral features in a speaker recognition system. GMM parameters are estimated from training data using the iterative Expectation-Maximization (EM) algorithm or Maximum A Posteriori (MAP) estimation from a well-trained prior model.[8].

A Gaussian mixture model is a weighted sum of M component Gaussian densities as given by the equation,

$$p(X/\lambda) = \sum_{i=1}^{M} w_i g(X/\mu_i, \Sigma_i)$$

where x is a D-dimensional continuous-valued data vector (i.e. measurement or features), w_i, i = 1, . . . ,M, are the mixture weights, and $g(X/\mu_i, \Sigma_i), i$ = 1, . . . ,M, are the component Gaussian densities.

The complete Gaussian mixture model is parameterized by the mean vectors, covariance matrices and mixture weights from all component densities. These parameters are collectively represented by the notation,

$$\lambda = \{w_i, \mu_i, \Sigma_i\} \text{ i = 1, . . . ,M.}$$

It is also important to note that because the component Gaussian are acting together to model the overall feature density, full covariance matrices are not necessary even if the features are not statistically independent. The linear combination of diagonal covariance basis Gaussians is capable of modeling the correlations between feature vector elements. The effect of using a set of M full covariance matrix Gaussians can be equally obtained by using a larger set of diagonal covariance Gaussians.[8]

E. Foreground Detection

After re estimating the parameters of the mixture, it is sufficient to sort from the matched distribution toward the most probable background distribution because only the matched models relative value will have changed. This ordering of the model is effectively an ordered, open-ended list, where the most likely background distributions remain on top and the less probable transient background distributions gravitate toward the bottom and are eventually replaced by new distributions. Then, the first B distributions are chosen as the background model, where EQ

where T is a measure of the minimum portion of the data that should be accounted for by the background. This takes the ªbestº distributions until a certain portion, T, of the recent data has been accounted for. If a small value for T is chosen, the background model is usually unimodal. If this is the case, using only the most probable distribution will save processing.[1]

IV. EXPERIMENT RESULT

To validate the algorithm we have verify it on the comparison experiment on several video sequences in different environment (indoor and outdoor). The experiment is carried out in dual core 2.5GHz processor.

Frame no. 46,508,525

Fig. 3: Indoor Scenes

Fig. 3 is our indoor database which is taken against illumination variations and cluttered background. An algorithm works successfully and detects foreground objects in all cases.

Frame no. 773,789,1455,1932

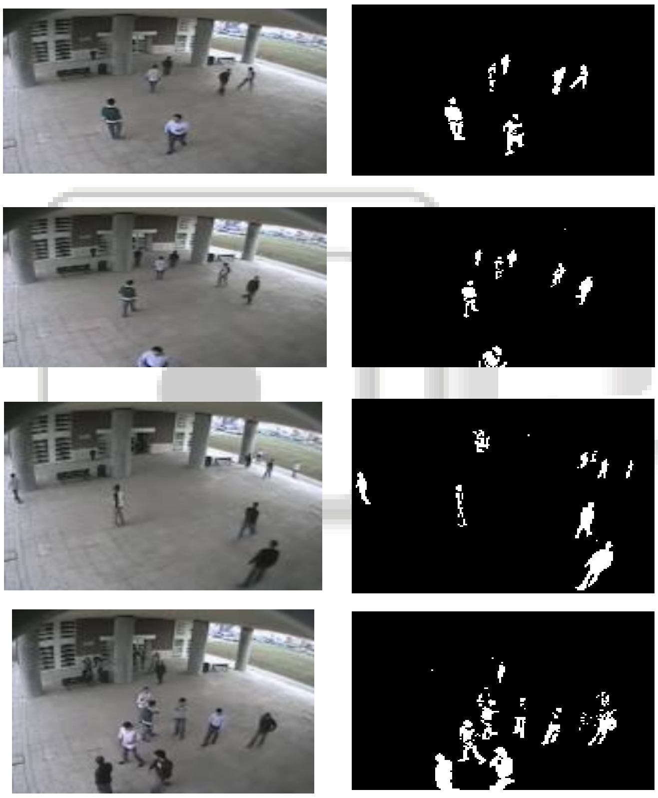

Fig. 4: outdoor scenes [16]

Fig. 4 is standard dataset for the outdoor scene with the different influencing parameters are there like illumination variation in space, clutter background and at some where possibilities of occlusions are there but an algorithms works satisfactorily and it can detects and segments almost every objects. So algorithm can works well on multiple objects.

PETS (Dataset)– Frame no. 872

Fig. 5: Outdoor scene
(PETS 2000 Dataset)

Fig.5 is a very famous standard outdoor scene from PETS dataset. Our algorithm produces good results by segmenting the vehicle as well as pedestrian simultaneously.

V. CONCLUSION

This paper has shown the probabilistic method for background modeling. All the environments involved with different spaces, different static camera, different illumination and various types of objects. This approach works with slow lighting changes by slowly adapting the values of the Gaussians. This approach works successfully used to track people in indoor environment and people and vehicle in outdoor environments. When background is moving or the cameras are moving algorithm fails to detect the objects. So, proposed algorithm gives robustness and stability to visual monitoring system against still background.

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] Hand book of Image and Video Processing, Editor AL BOVIK, Academic Press, USA, ISBN 0-12-119790-5, 2000.
[6] Practical Image and video processing Using MATLAB-O. Marques, IEEE press, John wiley sons inc., 2011.
[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