International Journal of Applied Engineering Research
ISSN 0973-4562 Volume 10, Number 11 (2015) pp. 28031-28040

http://www.ripublication.com

Patch Based Prior Occlusion Detection In Vital Sites With Performance Analysis

Smitha Suresh
Research Scholar, Department of Computer Science
Bharathiar University, Coimbatore, Tamilnadu, Pin: 641046, India
Assoc. Professor,SNGCE,KadayiruppuErnakulam
Email: smithadeepak2006 @gmail.com
Dr.K Chitra
Asst.Professor, Department of Computer Science
Govt. Arts College, Melur, Tamilnadu, Pin: 625106, India
Email: manikandan.chitra@gmail.com

Abstract

Video surveillance systems are increasingly widespread apparatus for security systems to monitor and manage suspicious behaviors or activities in the public areas. Object tracking is the procedure of extracting an object of interest from a video scene and keeping track of its action, orientation, occlusion etc. in order to extract valuable information. This paper tackles the problem of analyzing multiple human motions under occlusion. A novel PPOD system is used in the proposed system for the prior detection of occlusion in parts and to obtain cost effective solutions over the existing approaches of occlusion detection in video surveillance. Background modeling is the first processing step in video surveillance system. We attempt to model the background using Gaussian Mixture Model (GMM). Robustness and efficiency are the two striking issues of occlusion detection algorithms in video surveillance systems. Our experimental results and evaluation have been indicated that the proposed method is efficient and robust for the prior detection of occlusion during multiple human tracking in real world surveillance sites with very good accuracy.

Keywords: Video surveillance; Occlusion; Background model; Gaussian Mixture Model; Patches; Centre of mass

I. Introduction

Automated video surveillance deals with real time observation of people or vehicles in restricted environments. As the increase of deployment of more and more surveillance cameras, the pressure of automatic video processing methods is also growing. Human detection in video image frames is the essential step to be carried out in video surveillance system. The objective of video based surveillance system is to identify and monitor human for security purposes in a crowded environment such as airports, railway stations, shopping malls, banks etc. Video surveillance can be manual, semi autonomous or fully autonomous. In the earlier stages, human operators are constantly watching the videos in a monitor to detect any unusual activities. So these passive systems are to be substituted with efficient systems. A well-organized video surveillance system must be rapid, consistent and the algorithms are robust for moving object detection, classification, tracking, activity analysis. The system has to generate essential warning if any suspicious event arises. The performance of visual surveillance system mainly depends on the accuracy of object detected from the scene. The proposed self-sufficient system inputs the video sequence taken from a scene where surveillance is a necessity. In this system there is no human intervention and the system takes low level efforts in moving object detection, tracking, and also takes high level decision making task like occlusion detection. Recent methods to track people [1, 2] take up people detectors to make initial tracking hypotheses, and often include complex strategies to link people tracks across occluded events. However, they usually fail to track people that stay behind occluded at a considerable time in the entire sequence. State -of –the- art approaches to people detection [3, 4] are clever to detect people under a mixture of imaging circumstances, people poses and appearance. This approach becomes successful while people are fully observable. Black et.al, proposed [5] an approach for similar and tracking of articulated object using a view based design. In [6] C.Wojek et.al, proposed a bank of partial detectors to produce the primary proposals that are developed based on the 3D scene layout and temporal reasoning. In work [7], Daw-Tung Lin promotes a method called component shape template partial filter for object tracking and it can anticipate occlusion by Kalman filter. Ablavsky et.al, [8],devised a layered graphic model for the tracking of moderately occluded objects, defining an occluders centric illustration as a first order Markov process on activity zones with respect to the occlusion mask of the re locatable object. Chong et.al, [9], mutually joins multiple color and edge feature cues in a particle filter framework to resolve object tracking difficulties when occlusion happens. Führ et.al, [10], presented an approach combining vertical standing patch matching and a weighted pedestrian detection vector. This approach is skilled to detect occlusions and video sequences with brawny appearance variations. Our motivation for studying human object tracking under occlusion is that there exists an opportunity to build up a fully automated occlusion detection system to reduce the human interactions associated with the system and alerts are produced automatically based on the events detected with affordable computational costs.

This paper presents our work through the following section. Sections II express the system overview of the proposed occlusion detection system during tracking. Subsection 1 reveals moving object detection to extract the foreground object. Object tracking was illustrated in subsection 2 and a novel PPOD moving object tracking method for prior occlusion detection and occlusion detection using centre of mass was described in subsection 3 .The experimental dataset with the obtained results and analysis are presented in section III. Finally, we conclude our work in section IV.

II. System Overview

When the object present in a scene comes behind another object in the same scene, some parts may be left undetected due to occlusion. Moving object detection is the essential phase in occlusion detection system for the realistic analysis of the videos in further steps. Any abnormal events are probable to take place in surveillance sites because of the intentionally formed occlusion. So an automated monitoring system with low cost is essential in cope with the full occlusion or occlusion in parts. The following are the main contributions to develop a robust human detection framework to detect human under occlusion in a surveillance environment. To achieve this goal, the following objectives have been set out.

- Implementation of a video surveillance framework combining different image processing techniques.
- Implementation of occlusion detection using PPOD technique based on patches and Centre of Mass.
- Organization of a framework for the researchers who are running research in the same plain.

1. Moving Object Detection

The Moving object detection in a video stream is an indispensable step in video surveillance applications. Many algorithms for moving objects extraction are easy to available in recent years. Background subtraction, optical flow, and temporal differencing are some of the moving object detection algorithms. Among these moving object detection methods, the most common and effectual approach to localize moving objects is background subtraction and is for detecting moving regions in an image, especially under those situations with a relatively static background[11]. The name “Background subtraction” comes from by taking the absolute difference between current image and the reference background image and then thresholding the result to generate the objects of interest [12].The process that segment moving objects from the stationary scene is termed as background subtraction. In real environments, due to various background types such as waving tree branches, waterfalls, and flickering monitors, background subtraction is an extremely difficult task. To overcome this problem, background is updated by background modeling. Even if numerous background subtraction algorithms have been proposed in the literature, a problem of identifying moving objects in complex environments is still an obscure issue. The generally used background subtraction models are non recursive and recursive approaches. Compared with non recursive techniques, recursive techniques require a reduced amount of storage space. The frequently used non recursive background subtraction methods are frame differencing, average filtering and linear predictive filter. The existing Gaussian Mixture Model (GMM) is an example of recursive technique. The results of three background subtraction techniques are analyzed by determining the PSNR of the data samples chosen for our proposed work are shown in fig. 1. According to the result, a higher PSNR is obtained in GMM method. So GMM is chosen in the proposed PPOD method for object detection.

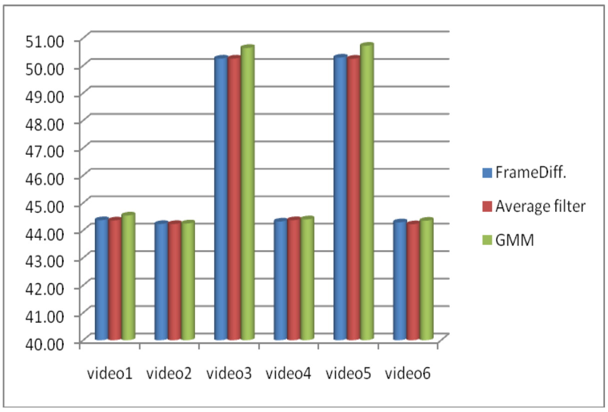

Figure 1: Analysis of GMM with Non Recursive methods

GMM based method was first launched by Stauffer and Grimson in 1999. Now it is the most extensively used method for background subtraction due to its speed, simplicity and the easiness of execution [13].The GMM is a single expansion of Gaussian probability density function. In this method, Gaussians are calculated and decide which Gaussians may match up to the background based on the resolution and the variance of all Gaussians. If the pixel intensity values that do not equivalent to the background distributions, then that pixel will be taken as foreground until there is a Gaussian that includes them. Gaussians are renewed later. The pixel intensity values that do not match with one of the pixel's background, Gaussians are grouped using connected components. For every pixel, the predetermined number of K states, typically between 3 and 5 is defined. The probability of incidence of a color at a given pixels is symbolized as:

$$P(X_t) = \sum_{i=1}^{K} \omega_{i,t}.\eta(Xt,\mu_{i,t},\Sigma_{i,t} \qquad (1)$$

Where $\omega_{i,t}$ is the weight constraint that is used to position by the i^{th} Gaussian distribution and K stand for the number of Gaussian distributions. η is a Gaussian distribution with the two parameters μ_i is the mean of the Gaussian distribution at time t and $\sum_i$ is the covariance matrix at time t. For computational ease, the covariance matrix $\Sigma_{i,t}$ can be assumed to be diagonal,

$$\Sigma = \sigma^2_{i,t} I \qquad (2)$$

After determining the equivalent components, the parameters of the equivalent components are renewed as follows:

Each new pixel value, X_t, is checked with the existing K Gaussian distribution until a match is obtained. A match is defined as a pixel intensity value that is less than or equal to 2.5 standard deviations of a distribution. If no one of the K distributions match the current pixel intensity value, the least possible distribution is go out. A new distribution with the present value as its mean value, an originally high variance, and low prior weight, is entering. The prior weight of K distributions at time t is adjusted as follows:

$$\omega i,s,t=(1-\alpha)\omega i,s,t-1+\alpha \quad (3)$$

The μ and σ parameters for unmatched distributions remain same. The parameters of the distribution which matches the most recent examination are updated as follows:

$$\mu i,s,t = (1-\rho).\mu i,s,t-1 + \rho.I,s,t \quad (4)$$

$$\sigma^2 i,s,t = (1-\rho).\sigma^2 i,s,t-1 + \rho.(Ii,s,t - \mu i,s,t)^2 \quad (5)$$

Where α is a user defined learning rate and ρ is a 2^{nd} learning rate. The Gaussians are ordered by the value of ω/σ.Then, the first B distributions are chosen as the background mode, where

$$B=\mathrm{argmin}_b\left(\sum_{K=1}^{b}\omega k>T\right) \quad (6)$$

Where T is a measure of the minimum portion of the data should be accounted for by the background.

2. Object Tracking

The binary image obtained after applying GMM often has discrete noises and holes in the object region [14]. In the proposed system the mathematical morphological processing such as erosion and dilation was used to remove the isolated noise and fill the gap in the object region. Object separation is performed in the system after extracting the foreground object. In human monitoring, the surveillance systems are used for tracking many humans in small groups in indoor and outdoor sites. Segmenting and tracking people in real time environment is a challenging and important problem in video surveillance. Video tracking algorithms is mainly used for finding the path for the targets in following video frames. The factors that are affecting the tracking are the number of objects in the frame, density of background objects, and type of objects in the frame and occlusion. Moving object tracking is a pre-requisite for many computer vision applications. To track objects and analyze their behaviors, it is essential to correctly classify the moving objects. Shape based classifiers such as points, boxes, silhouettes and blobs are available for classifying moving objects .The bounding box is used in the proposed work to track the object blobs. To detect and track people in a video sequence, group the pixels that represent individual people and calculate the appropriate bounding box for each person. The tracked objects are labeled with number and enclosed by bounding box. To form bounding box around each object, the algorithm as follows:

- Let B_tX and B_tY be the maximum value of row and column minus 0.5 gives staring
- B_1=Max(col)-Min(col)+1;
- L=Max(row)-Min(row)+1;
- B_{Box}=[B_tX B_tY B_1 L];
- Display B_{Box}

3. PPOD and Occlusion Detetion

In the proposed PPOD (Patch based Prior Occlusion Detection) system, patches are used to identify the parts in the objects that are going to make occlusion between any two objects. Patch is an identical sized small units or template that is created around each object in an image frame. To create patches, a mask with the same size of current frame was created first. Then find the edges of each object and the row and column values are stored in an array. Patch size in the proposed system was chosen as 4x4 pixel size. By utilizing the size of the patch and the number of rows, column values in the edges, find the total number of patches that is possible to create around the objects in the image frame. To create patches, find a suitable location in the object boundary to fit the patch. For this, search the appropriate position in all directions. After finding a suitable location to fix the patch, the patches are placed in that direction one by one. The details of patches are called one patch set p_i .The parameters in one patch set is as follows:

$$Pi= \{S1\ (Rw_1, Cl1),\ PN\ ((Rw_1, Cl1) \ldots (Rw_i, Cl_i)),\ IL\}$$

Where, $S1(Rw_1,Cl_1)$ is the starting index of the first patch, the patch data index is represented as $PN((Rw_1,Cl_1),(Rw_2,Cl_2)\ldots(Rw_i, Cl_i))$,it stand for the coordinates from starting point to the patch size-1. IL is the index label for each patch. In this new approach, patches are created only at the boundary regions in an object and check the distance between the patches in different objects in an image frame. If the measured distance between the patches is greater than the predefined threshold, then the color of the corresponding patches in that objects is altered. The centres of mass for all objects in image frames are also calculated for occlusion detection. During occlusion, the distance between centre of mass point in any two object is zero.

Analysis of the proposed experiment was completed in different video datasets to check the robustness. To show the precision and robustness of our algorithm, it was tested in many occluded image sequences. Simulation for the proposed occlusion detection algorithm was done in MATLAB 8.2. We compiled a set of 10 challenging tracking video sequences to evaluate PPOD system. The video sequences are collected from PETS dataset which contains a large number of well organized real world occlusions. Apart from this, our system was tested in many real time environments. To show the effectiveness and robustness of the proposed system in real time scenarios, sample occlusion video sequences were collected from the Govt. Medical College Hospital at Aalpuzha district, Kerala, India, a beach at the coastal outskirts of Arabian ocean and the Alwaye railway station which is a major one in Kerala. The overall performance of the system is quite acceptable in finding the occluded part and occlusion detection. Foreground object detection, Occluded patch identification and Occlusion detection are the important phases of performance analysis and is carried out as the part of research. Some of the tested video sample and its obtained results are shown as follows:

III. Experimental Results and Analysis

The proposed PPOD system has the capability of tracking multiple persons in complex scenes. The important features of the PPOD system when compared with the existing work is that, the parts that are going to occlusion are detected in advance. To perform prior occlusion detection, patch based model was created in the proposed work and the size of the patch is an important factor.

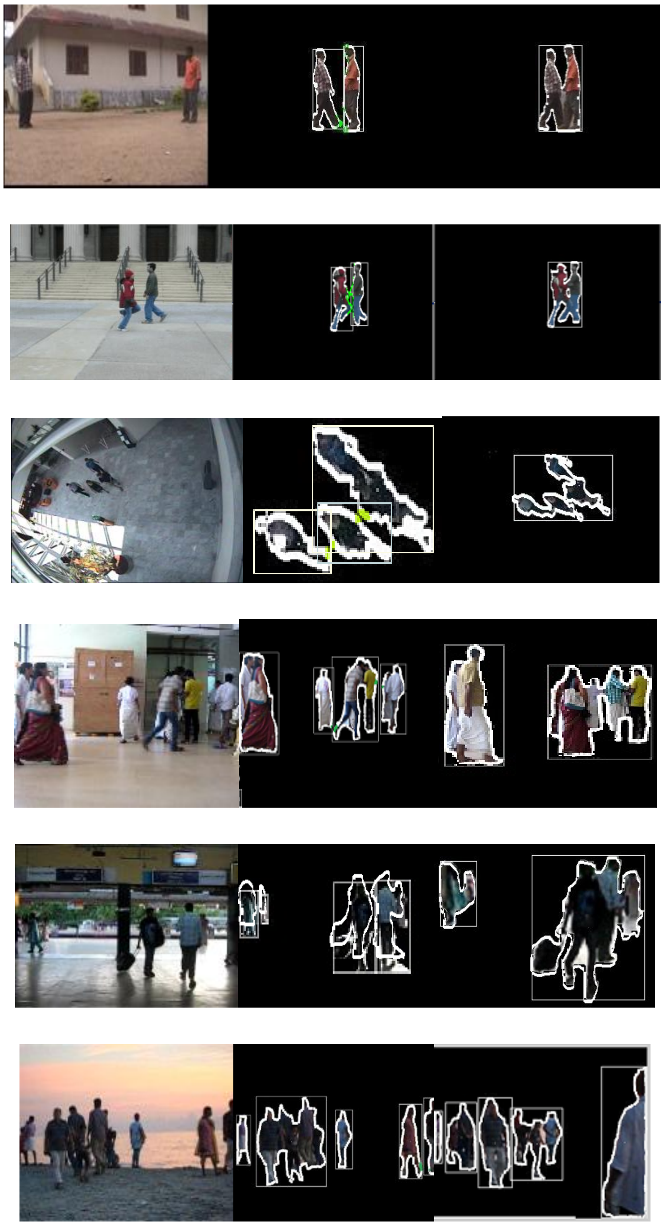

Figure 2: Occlusion Detection Data Sets: occluded patch identification and Occlusion Detection

The experimental results for occlusion detection for different videos are shown in fig 2.We analyzed the prior detection of occlusion by measuring the distance between the patches in different directions. From the analysis of the prior detection of occlusion, better result was not obtained if the patch size is too small or big. In our analysis, better results were obtained against first 4 videos for a patch size of 4 X 4 is shown in fig.3.

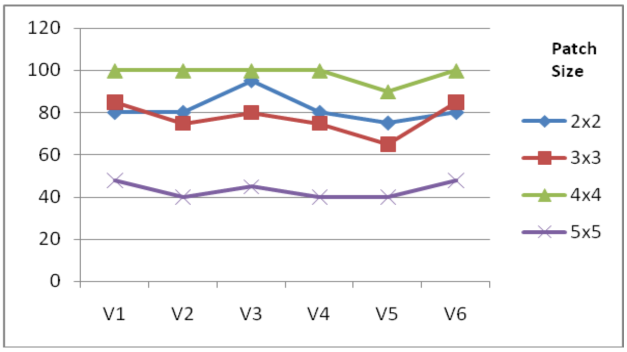

Figure 3: Prior occlusion detection sites based on patch size

The threshold level comparisons of some videos at four time events sequences are done based on distance of centre mass between objects and corresponding histograms are plotted. Here T_a corresponds to reference frame. Similarly T_b, T_c, and T_d represents threshold levels of before occlusion frame, during occlusion frame and after occlusion frame respectively. As per our research inventions, distance between the centre of mass in different objects will be zero at time event T_c. It is observed from the time frame T_c that, the histogram band became narrowed than other events at the fully occluded image frame. In Fig 4, the histograms of first video in the specified events are shown.

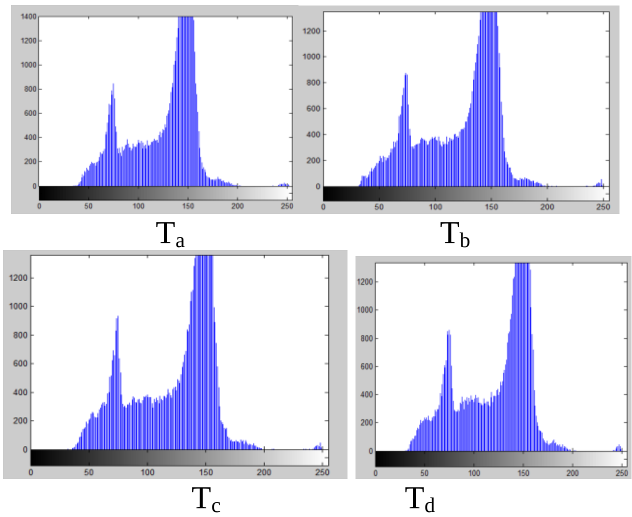

Figure 4: Histogram Analysis based on distance Threshold

Occlusion happens when the objects comes closer. The green colored units in the tracked frames indicate the occluded patches that are identified by calculating the distance between patches in different direction. We track the occluded object by calculating the distance between the centre of mass of objects in both direction. For calculating this distance, one predefined threshold value is used. In the proposed system, a threshold value of 50 is selected. During occlusion, the width of the bounding box is increased and in full occlusion, the width of the bounding box suddenly decreased and only one centre of mass is present between the occluded objects. But just after occlusion, again the width of the bounding box is increased and after occlusion, the bounding boxes are seemed separately in the frame. Similarly the centre of mass is also seemed as twice. The prior occlusion results of all the datasets are compared with its corresponding salt and pepper image noise pixels values and the mean Square error (MSE) of each video in the datasets are shown in fig 5. The MSE of fifth video is higher than other videos.

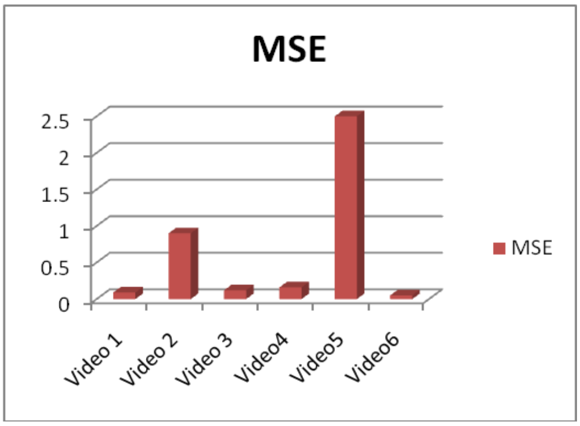

Figure 5: MSE for six Datasets

IV. Conclusion and Future Work

Today video surveillance system is used in many sites for human detection and identification. In the proposed PPOD work, background subtraction based on the GMM model used as our main object detection algorithm. The detected objects are tracked using bounding boxes. The strength of this paper is the prior detection of occlusion. The occlusion detection by centre of mass gives best result during human tracking. Based on patch based frame work and centre of mass concepts, occlusion was detected successfully. In the future works, the scope of present work can be extended to develop a simple model to detect and handle occlusion with added features of human density estimation in occlusion.

References

[1] A. Andriyenko, K. Schindler, and S. Roth, "Discrete- Continuous optimization for multitarget tracking", Proc. of the IEEE conf.,CVPR'June 2012.

[2] M.D. Breitenstein, F. Reichlin, B. Leibe, E. Koller-Meier, and L. Van Gool,"Robust tracking-by-detection using a detector confidence particle filter", IEEE ICCV'09.

[3] P. Dollár, C. Wojek, B. Schiele, and P. Perona," Pedestrian Detection: A Benchmark", CVPR'09.

[4] P. F. Felzenszwalb, R. B. Girshick, D. McAllester, and D. Ramanan.,"Object detection with discriminatively trained part-based models", PAMI'10.

[5] Michael Black & Allan Jepson ,"Eigen Tracking: Robust Matching and Tracking of Articulated Object using a View based Representation", Proc. European Conf. Computer Vision, Cambridge, UK, 1996.

[6] C. Wojek, S. Walk, S. Roth, and B. Schiele, "Monocular 3d scene understanding with explicit occlusion reasoning",CVPR'11.

[7] Lin, D.-T., Chang, Y.-H," Occlusion handling for pedestrian tracking using partial object template-based component particle filter", IADIS International Journal on Computer Science and Information Systems 8(2), 40–50, ISSN: 1646-3692.

[8] Ablavsky, V. and Sclaroff, S., "Layered Graphical Models for Tracking Partially Occluded Objects",IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 33, No. 9, pp. 1758-1775, 2011.

[9] Chong, Y., Chen, R., Li, Q., & Zheng, C. H," Particle Filter Based on Multiple Cues Fusion for Pedestrian Tracking", In Emerging Intelligent Computing Technology and applications,pp. 321-327, 2012.

[10] Führ, G., & Jung, C. R., "Combining Patch Matching and Detection for Robust Pedestrian Tracking in Monocular Calibrated Cameras", Pattern Recognition Letters, in press, 2013.

[11] K.Srinivasan, K.Porkumaran, G.Sainarayanan, "Improved Background Subtraction Techniques for Security in Video Applications". Alan M, "Background Subtraction techniques"

[12] C. Stauffer and W. Grimson,"Adaptive background mixture models for real-time tracking",Proceedings CVPR, pp. 246.252, 1999

[13] Ye, Q., Gu, R., Ji, Y.: Human detection based on motion object extraction and head– sholder fearure. Optik 124, 3880–3885, 2013.