

Development of Human Tracking in Video Surveillance System for Activity Analysis

Neelam V. Puri¹ and Prof. P. R. Devale²

¹PG Scholar, Department of Information Technology, Bharti Vidyapeeth Deemed University, Pune, India.)

²Professor & Head, Department of Information Technology, Bharti Vidyapeeth Deemed University, Pune, India.)

Abstract : In current era of digital technology visual surveillance systems are persistently in pursuance of being easier to use, versatile, inexpensive and very fast. Continuous video capturing systems are the replacement for human watch, because as we know human can be easily distracted and one mistake may lead to big disaster. So video surveillance systems make this kind of work very easier for user and it provides security and control where all time watch is required. Proposed algorithm will helpful for to detect moving object and classify it as human being and keep track of moving human. This procedure is done without getting help of any additional sensing device. In this paper proposed system can classify in three steps detection, tracking and action analysis. Detection of human being is done by combination of morphological procedure and feature extraction method. Tracking of same human and occlusion handling is done in second phase. At last phase activity analysis is done and in case of any abnormal activities, an alert should be issued.

Keywords: Automated Video surveillance system, Human Detection, Human Tracking, Motion Detection , Occlusion handling.

I. INTRODUCTION

There are several studies on the human tracking especially human detection using various methods like template matching, skin color recognition, gesture modeling etc. Human body gestures can be finding out by modeling human body and creating consolidated body of knowledge. But this kind of techniques has some drawbacks like highly sophisticated and very complex for implementation of the procedures. Solution for this kind of problem is pattern recognition and matching methods.

In video surveillance system to provide better solution for security and monitoring problems new technologies are introduced. Input device just capture the video streams and store it for the future purpose. Detecting, tracking and monitoring different activities by the same object are the more complicated and difficult job for the same capturing device. Every capturing device will not going to have the same configuration and will not be affordable by each user. Now we know that these video streams are always watched on the computer system or on same kind of the electronic devices which has some processing capability. While recording video stream this algorithm may helpful to analyze different actions taken by the moving object (i.e. human being). This paper is further subdivided as, Section 2 provides an over view of previous research done related to Automatic Video surveillance. Section 3 discusses an overview of the architecture of the proposed system in step by step procedure. Section 4, section 5 and section 6 have been dedicated to brief discussion of three major steps of proposed algorithm that is detecting object as human being, tracking same human being with occlusion handling and activity analysis. And conclusion work in section 6. Due citation are given in the references section.

II. PREVIOUS WORK DONE

Normally Background extraction technique is used for the foreground segmentation. There is another method called as optical flow computation [1]. Optical flow method is computationally expensive method and it can provide better performance. But for the real time system like video surveillance optical flow method are not so much suitable. To provide better computational consumption, these two techniques are combined in [2], to detect movement area they used background subtraction method, and they used optical flow computation method to remove unnecessary pixels from motion area. Subtracting current image from the reference background image motion area information can extracted. In recent foreground segmentation method to handle modelling multiple background objects like tree shaking, water waves etc. mixture of codebook [19], nonparametric kernel [18], different Gaussian distribution adaptive models [20] are used. To modelling multiple backgrounds more memory and more complicated algorithms are required even if these are more effective algorithms. Different Gaussian models considered with environments change like tree shaking, shadow, light change etc.

After detecting object from motion area, object can be classified as human being using different approaches like model based technique [3], Shape-based [4], neural network based technique [5] or blob based approaches. Shape based approach highly used in still images, in this shapes are modelled as local curves [6, 7], then implicitly represented by local or global descriptors in [8]. For pattern recognition neural network models are very much useful. Mostly part-based modelling is very efficient for detection of articulated objects like humans. For example, [9] use local features for part detection. [4] Introduce Hierarchical Part-Template Matching features for human detection.

Then object tracking algorithms in video surveillance system are established to track same object from two different frames. By creating human model which includes the human features tracking is possible. In human tracking system occlusion handling is difficult part. To solve this problem, Kalman filter based method was proposed [10]. Let study different techniques that are used to detect human and next phase human tracking.

III. ARCHITECTURE

Visual surveillance process that takes video as an input, processes the video and performs actions accordingly. The process of video surveillance consists of many phases.

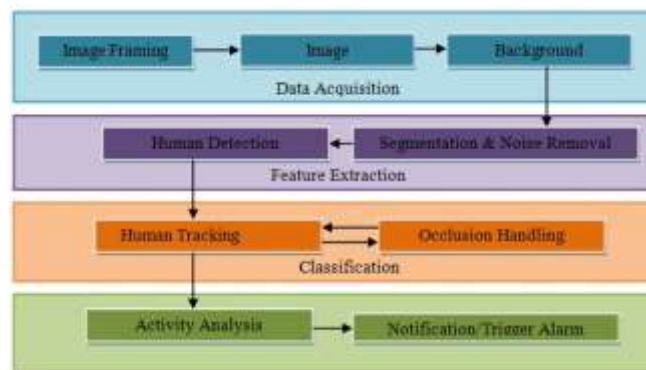


Figure 1. Architecture of Proposed System

This algorithm approach specifies three sub processes as follows and as shown in Fig. 1.

Detection of Moving Object: This algorithm takes any video stream as an input with static background. Then detect any motion as an object and then it will identify that object as human being.

Tracking of an object: After detection of an object as human being same object will track using its speed and velocity. Also in this phase occlusion handling is done.

Activity Analysis: At last phase using gathered information analyze different activities of tracked human recorded and if any sign of unusual activity or abnormal behaviour of human noticed then alarm will be triggered.

IV. HUMAN BEING DETECTION

4.1 IMAGE FRAMING AND IMAGE ACQUISITION

In Data acquisition phase different kind of sensing and capturing input devices used for data gathering from the working environment. For example video capturing and monitoring devices used to record video stream. For Image framing video must be divided into sequence of frames. Video must be divided mostly into 20-30 frames which are sent to the next phases for further processing.

Most probably background subtraction method applies on the static background images [11]. System can detect ‘motion area’ (only useful data) by comparing reference frame with current frame and get the extraction of a (new) moving object from the background.

This phase attempt background subtraction comparing reference image and current image in a pixel-by-pixel fashion as shown by equation below:

$$| P(X_1, Y_1)_C - P(X_2, Y_2)_{bk} | > Th \tag{1}$$

Where, $P(X_1, Y_1)_C$ is current frame pixel present at (X_1, Y_1) co-ordinate $P(X_2, Y_2)_{bk}$ is reference frame pixel present at (X_2, Y_2) co-ordinate and this threshold value of difference.

Here foreground pixels are extracted from the static background image so the threshold is fixed. When the threshold value is less than the pixel difference, it is considered as foreground image. In this way the foreground objects present in the scene are detected. Only patterns are preprocessed for selection of significant features.

4.2 BACKGROUND SUBTRACTION

Basically background subtraction done subtracting reference image from current image, here in this algorithm temporal differencing [12] is used to perform background subtraction. In this method the reference image is the previous images, and it is subtracted from the current image. In order to produce difference between two images the subtraction value must be greater than a threshold value:

$$I_{Th}(X, Y) = 1, I_{Th}(X, Y) > Th$$

$$= 0, I_{Th}(X, Y) \leq Th$$

From the above three human detection methods, the background subtraction is the simplest and mostly used when the background is stationery. In our implementation we will be keeping static background and hence will use the simple background subtraction method of human detection.

4.3 SEGMENTATION USING OPTICAL FLOW

Selected feature then passed through the classification phase where classified objects are compared with the features of object using any mapping function. Optical flow segmentation can be used to provide the velocity of the moving object is distinguish from that of the background, and has expected characteristics. Optical flow is the amount of image movement within a given time period [21], [22], [23]. Optical flow method also known as Lucas-Kanade Method [21],[23]. Assume reference image captured at time T and current image capture at time ΔT then the optical flow method calculate the motion between these two images for every pixel position as shown in Fig. 2.

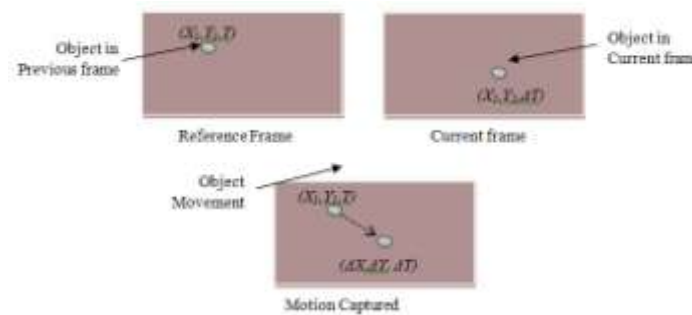


Figure 2. Motion of the moving Object

In reference image pixel present at location (X_1, Y_1, T) with intensity $I(X_1, Y_1, T)$ will have moved in current frame by $\Delta X, \Delta Y$ and ΔT the following image constraint equation can be given:

$$I(X_1, Y_1, T) = I(X_1 + \Delta X, Y_1 + \Delta Y, T + \Delta T)$$

Here, the movement is small enough. Equation (2) can be further written as:

$$\frac{\Delta I}{\Delta X} V_x + \frac{\Delta I}{\Delta Y} V_y + \frac{\Delta I}{\Delta T} = 0$$

Where, V_x and V_y are optical flow of $I(X, Y, T)$ or the X and Y components of the velocity and $\Delta I / \Delta T$ at a given pixel is just how fast the intensity is changing with time.

4.4 MORPHOLOGICAL OPERATIONS FOR NOISE REMOVAL

Due to environment factors (for example, humidity or fog in the area under surveillance, during transmission of video from the camera to the processing unit the illumination changes) the image is expected to contain noises. To get proper boundary lines (edges) without any noise portion morphological operations [1] like opening and closing on the subtracted image are very useful. Opening is a combination of erosion and dilation operations with erosion followed by dilation whereas closing is dilation followed by erosion [13]. Human detection is based on the Navneet Dallal and Bill Triggs' Histogram of Oriented Gradients (HOG) for Human Detection [14]. They used small region (cell) structure which accumulates a local 1-D histogram of edge orientations of the pixel value. A unique representation for the image and better invariance to illumination changes is given by combine entries of the histogram with contrast normalization.

By simple 1-D masking the Gradients in the image is calculated, in the vertical (\uparrow) and the horizontal (\rightarrow) direction. A histogram of gradient direction is computed for multiple regions of fixed rectangular size which generates a 9 component feature vector. Gradient sign shows the contrast difference between the foreground and the background in the image so it is ignored. The histograms clearly capture the edge or gradient structure that is very characteristic of local shape. The SVM classifier decides whether the region of interest contains a human figure or not.

V. HUMAN TRACKING AND HANDLING OCCLUSION

To track same object between two consecutive frames some correspondence should be there, if correspondence found then mark that object with same color rectangle and classify that object as human being. To find correspondence few features [15] of previous frame are stored and compared with the next frame, if features matches then same object found in both frames. And these features can be color, orientation, speed, intensity etc.

For tracking any object these features plays an important role. Tracking can be categories in two types [16] one is region based tracking and second one is contour based tracking. In region based tracking features of the blob detected from two consecutive images if both matches then these two frame object gets related. In contour based tracking [2] the energy of the boundary/ contour of blob detected from both frames and if that energy of the boundary matches then same object present in both images is declared. In this paper static background images are used and region tracking is very efficient with stationary background. So for tracking human being region tracking method is used. Region tracking stores the features of whole object for matching the features with the features of the object in the next frame.

But in this procedure required so much time for comparison, to reduce this wastage of time another approach can be applicable that is salient region tracking. In most salient region method only matches the most salient region of the previous frame with the most salient region of the current frame. Hence, in less amount of time required to match the whole image.

The most salient region tracking fetches the basic features from the image like color, orientation and intensity [17]. Then from these fetched features, the feature vector is calculated for color, orientation and intensity using centre-surround method [10]. After the feature map has been created to find out which feature more uniquely identifies the object 3 features are weighted. Using their weight the saliency map created for the detected object. Finally from the previous image feature weight vector is calculated for this most salient region and this feature weight vector is matched with the subsequent next frame's feature weight vector. If match is above the threshold value then there is a match and can predict the same object is present in both images. If match not found the search area is doubled and the same salient feature matching procedure is repeated. If the detected object match is not found after doubling the search area, then object is expected to be occluded by some other object or any stationary background object.

As seen in Figure1 Occlusion handling is not the separate process. Occlusion handling is done with the tracking of object. When motion area of the moving object has been detected then its centroid of an object can easily obtained from both images. Using centroid of these two images speed or velocity of an object can be obtained:

$$\begin{aligned}V_{x1} + V_{y1} &= S \\V_{x2} + V_{y2} &= S\end{aligned}\tag{5}$$

Even after doubling the search area if object not found then algorithm can conclude that object is occluded by some other object area. Using previously calculated velocity and direction of motion system can predict the next possible area of an object, minimum 5 frames should be processed to calculate the velocity and the direction of motion.

VI. ACTIVITY ANALYSIS AND ALARM TRIGGERING

After the detecting and tracking object as a human being activity analysis provides intelligence to this whole process. Algorithm stores the information about the detected human which will study its position, pose, motion and analyse them. After analysis, in case of any abnormal activity is found then an alarm is triggered.

Any action like moving into any highly secure area, moving with speed more than a limit, any typical pose that is not normal, mass movement and many other actions can refer as abnormal action and for this alarm will be triggered. The list of abnormal activities varies according to place and users.

In addition, any sign of unusual activity notification for this is done by making entry into the data base, generating reports, actually ringing any alarm [15], or sending message like email, SMS etc. Thus in this stage deal with utilizes of information has been received from last stages and notify unusual activities are involved in.

VII. CONCLUSION AND FUTURE WORK

While performing most salient region tracking using the direction and the speed of an object then object can be identified as human being and tracked with activity analysis. If this method is not able to find the human in initial search window, then just doubling the search window algorithm again search for human being and even if this fails too then the position of the human is predicted using previous stage information. In case the object is hidden and as algorithm is not searching whole frame instead just doubling the search area this process reduces the processing time. For proposed algorithm only static background is considered, in future it can be enhanced for moving background or on changed background.

REFERENCES

- [1] S.S. Beauchemin and J.L. Barron, "The Computation of Optical flow", *ACM Computing Surveys*, Vol.27, 1995, pp 433 - 466
- [2] Y.L. Tian and A. Hampapur, "Robust Salient Motion Detection with Complex Background for Real-time Video Surveillance", in *Proc. Of IEEE Computer Society Workshop on Motion and Video Computing*, January, 2005
- [3] I. Haritaoglu, D. Harwood and L.S. Davis, "W4: Who? When? Where? What? A Real Time System for Detecting and Tracking People", in *Proc. Of the International Conference on Face and Gesture Recognition*, April, 1998
- [4] Zhe Lin, Larry S. Davis, David Doermann, and Daniel DeMenthon, "Hierarchical Part-Template Matching for Human Detection and Segmentation", *University of Maryland*, College Park, MD 20742
- [5] L. Zhao and C.E. Torpe, "Stereo- and Neural Network-Based Pedestrian Detection", *IEEE Trans. Intelligent Transportation System*, Vol.1, No.3, Sept., 2000
- [6] Bo Wu, and Ram Nevatia, "Detection and Tracking of Multiple, Partially Occluded Humans by Bayesian Combination of Edgelet based Part Detectors.", *International Journal of Computer Vision*, 75(2): 247-266, 2007.
- [7] A. Opelt, A. Pinz, and A. Zisserman, "A Boundary-Fragment- Model for Object Detection". In *Proc. Of the European Conference on Computer Vision*, 2006.
- [8] E. Seemann, B. Leibe, and B. Schiele, "Multi-Aspect Detection of Articulated Objects", *IEEE Conference on Computer Vision and Pattern Recognition*, June 2006.
- [9] K. Mikolajczyk, C. Schmid, and A. Zisserman, "Human Detection based on A Probabilistic Assembly of Robust Part Detectors". In *Proc. Of the European Conference on Computer Vision*, 2004.
- [10] J.W. Lee, M.S. Kim, I.S. Kweon, "A Kalman filter based visual tracking algorithm for an object moving in 3D", *IEEE International Conference on Intelligent Robots and System*, Vol.1, Aug., 1995
- [11] J. Connell, A.W. Senior, A. Hampapur, Y-L. Tian, L. Brown, and S. Pankanti, "Detection and Tracking in the IBM People Vision System", *IEEE International Conference on Multimedia and Expo.*, June 2004
- [12] L.M. Fuentes and S.A. Velastin, "People Tracking in surveillance application", in *Proc. 2nd IEEE International Workshop on Performance Evaluation of Tracking and Surveillance*, Dec. 2001
- [13] M. Balcells-Capellades, D. DeMenthon and D. Doermann, "An Appearance-based Approach for Consistent Labeling of Humans and Objects in Video". *Pattern Analysis and Applications*, November 2004.
- [14] N. Dalal and B. Triggs, "Histograms of oriented gradients for human Detection", in *Proc. Of the Conference on computer Vision and Pattern Recognition*, San Diego, California, USA, pp. 886-893, 2005.
- [15] Suzaimah Bt Ramli, Kamarul Hawari B Ghazali, Muhammad Faiz Bin Mohd Ali, Zakaria Hifzan B Hisahuddin, "Human Motion Detection Framework", *IEEE 2011*, pp. 158-161.
- [16] A. Senior, A. Hampapur, Y.L. Tian, L. Brown, S. Pankanti and R. Bolle, "Appearance Models for Occlusion Handling", in *Proc. Of Second International workshop on Performance valuation of Tracking and Surveillance systems in conjunction with Computer Vision and Pattern Recognition*, Dec. 2001
- [17] C. BenAbdelkader and L. Davis, "Detection of People Carrying Objects: a Motion-based Recognition Approach", *5th IEEE International Conference on Automatic Face and Gesture Recognition*, May, 2002
- [18] A. Elgammal, R. Duraiswami, D. Harwood and L. Davis, "Background and Foreground Modeling Using Nonparametric Kernel Density Estimation for Visual Surveillance", in *Proc. Of the IEEE*, Vol.90, No.7, July, 2002
- [19] Kim, T. H. Chalidabhongse, D. Harwood and L. Davis, "Adaptive background mixture models for real tracking", *IEEE International Conference on Image Processing (ICIP)*, 2004.
- [20] C. Stauffer and W.E.L. Grimson, "Adaptive background mixture models for real tracking". *Int. Conf. Computer Vision and Pattern Recognition*, Vol.2, 1999, pp.246-252
- [21] Nan Lu, Jihong Wang, Q.H. Wu and Li Yang, "An improved Motion Detection Method for Real-Time Surveillance", *International Association of Engineers (IAENG) International Journal of Computer Science*, 35:1, IJCS_35_1_16
- [22] Denman, Simon P. and Chandran, Vinod and Sridharan, Sridha (2007) "An adaptive optical flow technique for person tracking systems". *Pattern Recognition Letters* 28(10): pp. 1232-1239
- [23] Yokoyama, M. and T. Poggio, "A Contour-Based Moving Object Detection and Tracking." In: *Proceedings of Second Joint IEEE International Workshop on Visual Surveillance and Performance Evaluation of Tracking and Surveillance (in conjunction with ICCV 2005)*, Beijing, China, October 15-16, 2005. pp. 271-276