International Journal of Science and Research (IJSR)
ISSN (Online): 2319-7064
Index Copernicus Value (2013): 6.14 | Impact Factor (2013): 4.438

Automatic Video Based Surveillance System for Abnormal Behavior Detection

Divya J

M.E. (VLSI and Embedded system), G. H. Raisoni College of Engineering & Management,, Wagholi, Pune
(Under the guidance of Prof. Dr. R.S.Bichkar)

Abstract: ***Security and surveillance are important issues in today's world. Any behavior which is uncommon in occurrence and deviates from normally understood behavior can be termed as suspicious. This model aims at automatic detection of abnormal behavior in surveillance videos. We have targeted to create a system for the recognition of human activity and behavior, and extract new information of interest for end-users in highly secured indoor surveillance system. The objective for this project is to design a model for detection of abandoned objects and track abnormal human behaviors. The multi-object detection is done by background subtraction with the help of appropriate model and then recognizing the person using HOG feature and SVM classifier. Face of a detected person can be captured using Viola–Jones algorithm. Finally, anomaly detection is done for tracking persons based on their individual appearance using Mean Shift Technique. For change detection, Mean-ratio and Log Ratio operators are used.***

Keywords: Video Surveillance, abnormal behaviour, Meanshift, Thresholding, Abandoned object

1. Introduction

An effective video surveillance system depends on detection of suspicious activities. In recent times, detecting abnormalities in human behavior using such systems gained more importance as it can provide clues while preventing breaches in security. Suspicious behavior detection is one of the paramount goals in surveillance systems along with abandoned object detection.

However, the major constrained posed here are the result of human factors. Such constraints face challenges towards effective utilization of the systems, especially as crime-fighting tools. An important factor to be considered here is the fatigue limits of the human operators. It is significantly more while several scenes monitored by the operators for prolonged period of time. This result in a considerable degrades in the monitoring performance on surveillance areas. Another factor is that there is a limitation in the number of monitors that a human operator can manage simultaneously. This would add more complexity in large-scale surveillance systems having many cameras with a limited number of security staff. Therefore, an automatic process for detecting suspicious behavior is a necessity.

2. Literature Review

Video surveillance, which involves acquiring and processing visual data from a scene, to detect target(s) along time and space for purpose of recognizing interesting situations and perhaps generate alarms, has been a particularly hot topic. It typically begins with change detection and motion information capture for moving targets (using tracking or non-tracking methods), to enable successive high-level event analysis. Oftentimes, people pose it as a pattern-learning problem that deals with the classification of video object behavior by finding good matches either with *a priori* known templates of behavior or learning and forming statistical models of the behavior types from time varying feature data.

In [1], *Brezeale et al.* explored the video classification literature. They found that features are drawn from three modalities–texts, audio, and visual. Also combinations of these features along with classifications have been explored. While the more focused topic of understanding specific events in video data was addressed in a review by *Lavee et al.* [2]. In their survey the two main components of the event understanding process: Abstraction and Event modeling has been explained. *Buxton* [3] presented a survey on understanding dynamic scene activity using intelligent cognitive vision systems which can compute conceptual descriptions by analyzing activity in dynamic scenes from motion trajectories of moving people and the objects in particular scene. *Hu et al.* [4] produced a survey of automated visual surveillance which focuses on the two areas of motion detection and object tracking, offering only a short overview of behavior understanding. They reviewed recent developments, basic strategies along with possible research directions.

Teddy Ko [5] presented a survey for homeland security applications on behavior analysis in video surveillance. Also he exploited developments and a general strategy of stages involved in video surveillance, and analyzes the challenges and feasibility for combining object tracking, motion analysis, behavior analysis, and biometrics for stand-off human subject identification and behavior understanding. *Dee and Hogg*, in their review of *real-world surveillance* [6], include a section reviewing anomaly detection. In [7], *Niels Haering et al.* enabled the expansion of the vocabulary of video surveillance systems paving the way for more general automated video analysis.

In [8], *Minh et al.* proposed efficient algorithm for still images, which detects various human actions by analyzing silhouette and the upper body of the human being. *Dorin et al.* in [9] proposed a robust approach using mean shift algorithm for the analysis of complex multimodal feature space and to delineate arbitrarily shaped clusters in it.

3. System Model

Fig. 1 below represents the actual work-flow implemented in our work.

Video Input
While the application is loaded, as part of the initialization process we scan for the video input devices attached to the system. Further these devices are made available for selection by the user; so that the most appropriate area to be captured for a given instance of time can be chosen.

Background Image Acquisition
This section basically consists of capturing a frame as the reference image or the desired ideal background condition which thus considered as the reference for any further processing. Once the background is set, camera is then programmed for capturing live video footage of the monitored area.

Image pre-processing
Pre-processing is performed on the acquired images for enhancing the quality of the frames. The video frames have a lot of noise due to camera, illumination and reflections etc. This can be removed and quality of images can be enhanced with the help of preprocessing stages. The suitable steps are carried out in this stage. Now-a-days, video pre-processing is performed in the digital domain, which is carried out after the digital video capture, giving full play to the convenience of the digital signal, efficient, flexible and consistent superiority.

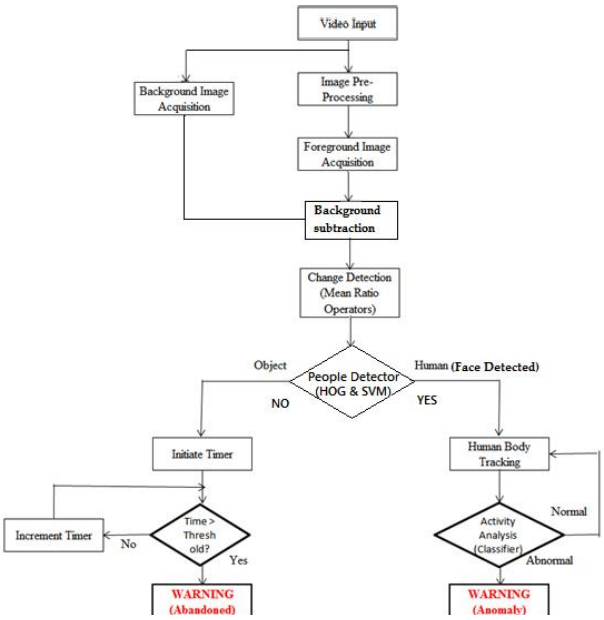

Figure 1: System Model

Change Detection
The captured video is to be processed for detecting any change with the reference frame taken and processed in the previous step, as the reference condition. The *Image Acquisition Toolbox* available within Matlab is used, for camera controlling, which is connected to the personal computer. By using these functions camera can be operated to capture desired snapshots and even a sequence of frames which forms a part of the video. In this stage, we separate foreground images from background images. For change detection *Mean-ratio operators* are used.

i. Foreground Extraction
This is one of the significant sections of the project. Here the foreground object is reconstructed by removing the background elements. To accomplish this, the background image obtained previously is taken as a perfect reference image for the processing of image segmentation.

Before proceeding to identify the foreground section of the image with reference to the background image, the enlisted challenges needs to be considered:

i. It is always very difficult to construct a robust system which will suit various light conditions.
ii. The noise in the images needs to be removed.
iii. The comparison among the background and acquired images must be performed as quickly as possible, so that the system works in a real time manner.

To address these challenges, each image frame is initially converted to grayscale image. This makes the task easier because when a pixel to pixel comparison is made for identifying the foreground image, a 2-D unit that is grey scale image of data will be faster and easier to process upon when compared to a RGB image which is 3-D unit.

In our algorithm, we performed morphological operations to remove further noise and then binary image is dilated using the structuring element, which returns the dilated image. So that maximum information can be gathered in the smallest span of time. This results in decreased amount of the processing time along with an increase in the accuracy level of the image processing which is very much required.

ii. Algorithm for Change Detection
The detailed algorithm of background model construction and change detection is summarized as follows:

i. The probability that the same value of pixel is present in the background image is taken into account when a new pixel in the new image is observed.
ii. If the pixel value is greater than the tolerance range i.e. threshold, then the pixel is recorded and marked it as a part of the foreground image. This is done as a background image pixel is expected to repeat its value for a long time. At least it is expected to be within a tolerance range. A very productive change marks the pixel as a part of the change. That has to be tracked and segmented out.
iii. If the observed pixel lies within the threshold value, it is considered as a part of the background and is replaced with a zero. At the same time if any significant change in the pixel value is observed the pixel particular value is retained as it is assumed to be the part of the change at the foreground.

This probability algorithm allows us to reconstruct an image which retains the pixel values of the pixels which is the part of the foreground change and removing the pixels which is a part of the background image.

Abandoned object detection

In case an object is detected, the timer is initiated and continuously incremented until it is found that the object is static. The incremented timer is consistently compared against a pre-defined *Threshold timer value*. Once the incremented timer value equals or exceeds the threshold value, it is considered that the object is abandoned. The next action would be to raise an alarm and notify about the abandoned object found. Additionally, the system highlights the object by adding a rectangle around it on the screen that helps in identifying it in the monitored area.

Here, the threshold value is of significant importance and hence needs to be determined carefully. It forms a basis for deciding whether the object is in a dormant state or not. The lower value may lead to false alarms, while the greater value may result in the object going undetected for a considerably larger amount of time. In case of surveillance systems this may prove fatal, since the abandoned object might contain a timed bomb.

Activity Analysis

If the detected set of pixels represents a human being then his behavior is tracked for any sort of anomalies. For detecting human being in our algorithm we are utilizing the people detector object, which detects people in an input image using the Histogram of Oriented Gradient (HOG) features and a trained Support Vector Machine (SVM) classifier.

HOG is a mechanism for feature description. The need for a feature descriptor is to generalize a given object in such a way that the object (in this case a person) exhibits a close match to this feature descriptor when observed under various conditions. This makes the classification process simple.

In our algorithm we have used trained Support Vector Machine (a type of machine learning algorithm for classification), or "SVM", to recognize HOG descriptors of people. It utilizes a sliding detection window which is moved across the image. A HOG descriptor is computed at each position of the detector window. This descriptor is then given to the trained SVM, which classifies it whether it is person or not.

The face of the detected person can be captured using *Viola–Jones object detection framework*. Which can be stored are stored in disk location specified by the user which can be utilized for further use if required.

Once a person is detected, next task is to track his behavior for abnormalities. For this, in our model we are using widely used *Mean-shift algorithm* [10]. The mean shift algorithm is a *nonparametric clustering* technique that does not require prior knowledge of the number of clusters, and does not constrain the shape of the clusters. It is basically an iterative expectation maximization-clustering algorithm executed within local search regions.

The motion tracking continues throughout the video with respect to the standard background which changes as per the motion of the moving object in the captured video. This satisfies the aim of the project where we consider an indoor surveillance system. This system aims to find any kind unwanted disturbances in a highly secure zone such as bank vaults or military weapon hanger, and many other such highly secure zones.
Using the geometry of the frame associated with the detected motion of the recognized object, we may categorize some basic activities like running/walking, jumping and bending:

Running/Walking: Running is detected when the speed calculated with the change in the centroids X-axis across the subsequent frames goes beyond a predetermined threshold. In fact, the speed value less than this threshold characterize "normal" walking. This is shown in Fig (b) & Fig (c) below.

Jumping: This case is detected when the position of the centroid changes along Y-axis. Moreover, the speed of this up and down motion should be greater than a predetermined threshold, shown in Fig (e) below.

Bending: This behavior can be detected when the aspect ratio of the detected object decreases across the frames, beyond a pre-defined threshold value. This is shown in Fig (d) below.

The distance travelled by a person is calculated by using the Euclidean distance formula using the centroid. It is. The variables for this are the pixel positions of a person at initial stage to the final stage.

Distance = $\sqrt{(X2 - X1)^2 + (Y2 - Y1)^2}$

Where *X1* is previous pixel position and *X2* is present pixel position in width. *Y1* represents previous pixel position and *Y2* is present pixel position in height. Velocity of moving object is determined by using the distance travelled by the centroid to the frame rate of the video.

Velocity = Distance travelled / Frame rate

4. Result

The proposed work has been developed using MATLAB 8.1(R2013a) on Intel dual core processor, 2GB RAM and Windows 7 operating system. The real time video sequences are acquired at the rate of 2 frames/second with the frame size of 320×240 pixels resolution.

Figure (a): Abandoned Object Detection

Figure (b): Walking

Figure (C): Running

Figure (d): Bending

Figure (e): Jumping

5. Conclusion

In this paper, Automatic real time Video Based Surveillance system for abnormal behavior is proposed, based on background subtraction, mean shift algorithm and Thresholding. For abandoned object detection, we use threshold method to detect object based on the background captured when the application is started. Further, the abnormal activity is detected using Mean-shift algorithm by finding the centroid and velocity. A benefit of this method is that it is time efficient, and it works well in artificial light environment as well. The results of our method, indicates that the method is effective in detection of abandoned object as well as abnormal behaviors in secure zone.

References

[1] D. Brezeale and D. J. Cook, "Automatic video classification: A survey of the literature," *IEEE Trans. Syst., Man, Cybern. C, Appl. Rev.*, vol. 38, no. 3, pp. 416–430, May 2008

[2] G. Lavee, E. Rivlin, and M. Rudzsky, "Understanding video events: A survey of methods for automatic interpretation of semantic occurrences in video," *IEEE Trans. Syst., Man, Cybern. C, Appl. Rev.*, vol. 39, no. 5, pp. 489–504, Sep. 2009

[3] H. Buxton, "Learning and understanding dynamic scene activity: A review," *Image Vision Comput.*, vol. 21, no. 1, pp. 125–136, 2003

[4] W. Hu, T. Tan, L. Wang, and S. Maybank, "A survey on visual surveillance of object motion and behaviors," *IEEE Trans. Syst., Man, Cybern.C, Appl. Rev.*, vol. 34, no. 3, pp. 334–352, Aug. 2004

[5] T. Ko, "A survey on behavior analysis in video surveillance for homeland security applications," in *Proc. Appl. Imagery Pattern Recognit.Workshop*, Oct. 15–17, 2008, pp. 1–8

[6] H.M. Dee and D. C. Hogg, "On the feasibility of using a cognitive model to filter surveillance data," in *Proc. IEEE Int. Conf. Adv. Video Signal Based Surveillance*, Sep. 15–16, 2005, pp. 34–39

[7] N. Haering, P. L. Venetianer, and A. Lipton, "The evolution of video surveillance:Anoverview,"*Mach.VisionAppl.*, vol. 19, no. 5-6, pp. 279–290, 2008

[8] Hoai, Ladický, Zisserman, "Action Fromweak Alignment Of Body Parts", *British Machine Vision Conference*, 2014

[9] Dorin comaniciu, Peter meer, "Mean Shift: a Robost approach Toward Feature Space Analysis", *IEEE Trans.*

on Pattern Analysis and machine intelligence, vol. 24, No. 5, may 2002

[10] Shehan Fernando, T M J A Cooray, "Mean Shift Kalman Object Tracking for Video Surveillance", *National Engineering Conference, University Of Moratuwa, Sri Lanka 2013*

[11] Oluwatoyin P. Popoola and Kejun Wang, "Video-Based Abnormal Human Behavior Recognition—A Review", *IEEE Transactions On Systems, Man, And Cybernetics—Part C: Applications And Reviews*, 2011

[12] Weilun Lao, Jungong Han, and Peter H.N. de With, *Fellow*, IEEE, "Automatic Video-Based Human Motion Analyzer for Consumer Surveillance System", *IEEE Trans. on Consumer Electronics*, Vol. 55, No. 2, MAY 2009

[13] Chen-Chiung Hsieh and Shu-Shuo Hsu, "A Simple and Fast Surveillance System for Human Tracking and Behavior Analysis", *Third International IEEE Conference on Signal-Image Technologies and Internet-Based System*, 2008