International Journal of Engineering and Techniques - Volume 1 Issue 3, May – June 2015

RESEARCH ARTICLE OPEN ACCESS

Survey on Moving Body Detection in Video Surveillance System

Prof. D.S.Patil[1], Miss. R.B.Khanderay[2], Prof.Teena Padvi[3]
[1]Associate Professor, SSVPS, Dhule (North maharashtra university)
[2](ME persuing,SSVPS,Dhule)
[3](Govt polytechnic,Avasari,pune)

Abstract:

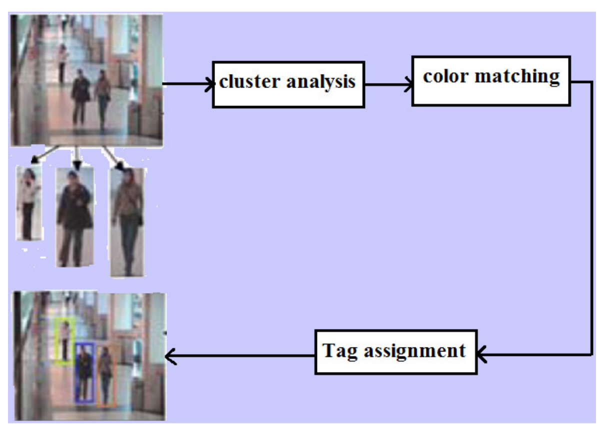
This paper represents a survey of various methods of video surveillance system which improves the security. The aim of this paper is to review of various moving object detection technics. This paper focuses on detection of moving objects in video surveillance system. Moving body detection is first important task for any video surveillance system. Detection of moving object is a challenging task. Tracking is required in higher level applications that require the location and shape of object in every frame. In this survey,paper described about optical flow method, Background subtraction, frame differencing to detect moving object. It also described tracking method based on Morphology technique.

Keywords -- Frame separation, Pre-processing, Object detection using frame difference, Optical flow, Temporal Differencing and background subtraction. Object tracking

I. INTRODUCTION

There is an increasing desire and need in video surveillance applications for a proposed solution to be able to analyze human behaviors and identify subjects for standoff threat analysis and determination. The main purpose of this survey is to look at current developments and capabilities of visual surveillance systems and assess the feasibility and challenges of using a visual surveillance system to automatically detect abnormal behavior, detect hostile intent, and identify human subject. Visual (or video) surveillance devices have long been in use to gather information and to monitor people, events and activities. Visual surveillance technologies, CCD cameras, thermal cameras and night vision devices, are the three most widely used devices in the visual surveillance market. Visual surveillance in dynamic scenes, especially for humans, is currently one of the most active research topics in computer vision and artificial intelligence. It has a wide spectrum of promising public safety and security applications, including access control, crowd flux statistics and congestion analysis, human behavior detection and analysis, etc.

Video surveillance in dynamic scene with multiple cameras, which may detect, recognize and track specific objects from image sequences, and most important to understand and describe object behavior. The main stratergy of video surveillance is to develop intelligent visual surveillance to replace the traditional passive video surveillance that is proving ineffective as the number of cameras exceed the capability of human operators to monitor them. The main odject of visual surveillance is not only to put cameras in the place of human eyes, but also to accomplish the entire surveillance task as automatically as possible. The capability of being able to analyze human movements and their activities from image sequences is crucial for visual surveillance.



generally,the processing framework of an automated visual surveillance system includes the following stages: Moving object detection, object classification, object tracking, behavior and activity analysis and understanding, person identification, and camera handoff and data fusion. Most of the every visual surveillance system starts with motion and object detection. Motion detection aims at segmenting regions corresponding to moving objects from the rest of an image. Subsequent processes such as object tracking and behavior analysis and recognition are greatly dependent on it. The process of motion/object detection usually involves background / environment modeling and motion segmentation, which intersect each other during the processing. Motion segmentation in image sequences aims at detecting regions corresponding to moving objects such as humans or vehicles. Detecting moving regions provides a focus of attention for later processes such as tracking and behavior analysis as only these regions need be considered and further investigated. After motion and object detection, surveillance systems generally track moving objects from one frame to another in an image sequence. The tracking algorithms usually have considerable intersection with motion detection during processing. Tracking over time typically involves matching objects in consecutive frames using features such as points, lines or blobs. Behavior understanding involves analysis and recognition of motion patterns, and the production of high-level description of actions and interactions between or among objects. In some circumstances, it is necessary to analyze the behaviors of people and determine whether their behaviors are normal or abnormal.

The problem of who enters the area and/or engages in an abnormal or suspicious act under surveillance is of increasing importance for visual surveillance. Human face and gait are now regarded as the main biometric features that can be used for personal identification in visual surveillance systems.

II. MOVING OBJECT DETECTION:

Each application that benefit from smart video processing has different needs, thus requires different treatment. However, they have something in common: moving objects. Thus, detecting regions that correspond to moving objects such as people and vehicles in video is the first basic step of almost every vision system since it provides a focus of attention and simplifies the processing on subsequent analysis steps. Due to dynamic changes in natural scenes such as sudden illumination and weather changes, repetitive motions that cause clutter (tree leaves moving in blowing wind), motion detection is a difficult problem to process reliably. Frequently used techniques for moving object detection are frame difference, background subtraction, Temporal Differencing and optical flow whose descriptions are given below.

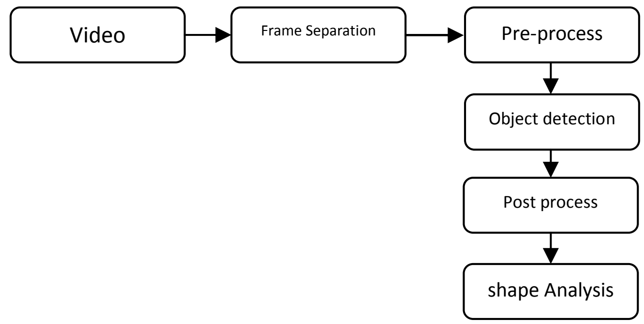


Fig2. A generic framework for smart video processing algorithms.

2.1.1 Preprocessing:

In Preprocessing of the proposed system the following steps namely Gray scale conversion, Noise removal is involved. In computing, a gray scale digital image is an image in which the value of each pixel is a single sample, it carries only intensity information. Images of this sort, also known as black-and-white, are composed exclusively of shades of gray, varying from black at the weakest intensity to white at the strongest. Gray scale images are distinct from one-bit bi-tonal black-and-white images, which in the context of computer imaging are images with only the two colors, black, and Gray scale images have many shades of gray in between. Gray scale images are often the result of measuring the intensity of light at each pixel in a single band of the electromagnetic spectrum, and in such cases they are monochromatic proper when only a given frequency is captured. And the gray scale conversion of image is given by [17].

$$\mathrm{gray}(i,j) = \{0.29 * \mathrm{rgb}(:,:,1) + 0.59 * \mathrm{rgb}(:,:,2) + 0.11 * \mathrm{rgb}(:,:,3)\}; \quad (1)$$

Generally we are using median filter to suppress the noise. The procedures are

(i) Arranging matrix pixel value in the form of ascending order.
(ii) Find the median value of that matrix.
(iii) Replace that value into that noisy pixel location.

2.1.2. Histogram Equalization

Basically the histogram equalization spreads out intensity values along the total range of values in order to achieve higher contrast. This method is especially useful when an image is represented by close contrast values, such as images in which both the background and foreground are bright at the same time, or else both are dark at the same time. Here are the steps for implementing this algorithm.

1. Create the histogram for the image.
2. Calculate the cumulative distribution function histogram.
3. Calculate the new values through the general histogram equalization formula.
4. Assign new values for each gray value in the image.

2.2 Frame Differencing:

Detection of moving object from a sequence of frames captured from a static camera is widely performed by frame difference method. The objective of the approach is to detect the moving objects from the difference between the existing frame and the reference frame. The frame difference method is the common method of motion detection. This method adopts pixel-based difference to find the moving object.

Difference of Two Consecutive Frames

Ik is supposed to be the value of the kth frame in image sequences. Ik+1is the value of the (k+1)th frame in image sequences. The absolute differential image is defined as follows:

$$Id(k, k+1) = |Ik+1 - Ik| \quad (2)$$

Limitations

The proposed method also detects the motion due to the movement in air. As the air moves, the camera not remains in the position of static so when there is no movement of object then also it results motion and shows holes in the binary output image.

2.3 Optical Flow

Optical flow is the distribution of the apparent velocities of objects in an image. By estimating optical flow between video frames, you can measure the velocities of objects in the video. In general, moving objects that are closer to the camera will display more apparent motion than distant objects that are moving at the same speed.

Optical flow estimation is used in computer vision to characterize and quantify the motion of objects in a video stream, often for motion-based object detection and tracking systems.

Optical flow analysis is also useful to distinguish rigid and non-rigid objects. A. J. Lipton proposed a method that makes use of the local optical flow analysis of the detected object regions. It is expected that non-rigid objects such as humans will present high average residual flow whereas rigid objects such as vehicles will present little residual flow. Also, the residual flow generated by human motion will have a periodicity. By using this cue, human motion, thus humans, can be distinguished from other objects such as vehicles.

2.4 Background Subtraction:

Object detection can be achieved by building a representation of the scene called the background model and then finding deviations from the model for each incoming frame. Any significant change in an image region from the background model signifies a moving object. The pixels constituting the regions undergoing change are marked for further processing. Usually, a connected component algorithm is applied to obtain connected regions corresponding to the objects. This process is referred to as the background subtraction [6].

Heikkila and Silven [6] presented this technique. At the start of the system reference background is initialized with first few frames of video frame and that are updated to adapt dynamic changes in the scene. At each new frame foreground pixels are detected by subtracting intensity values from background and filtering absolute value of differences with dynamic threshold per pixel [8] .The threshold and reference background are updated using foreground pixel information. It attempts to detect moving regions by subtracting the current image pixel-by-pixel from a reference background image that is created by averaging images over time in an initialized period [6]. The pixels where the difference is above a threshold are classified as foreground. After creating foreground pixel map, some morphological post processing operations such as erosion, dilation and closing are performed to reduce the effects of noise and enhance the detected regions. The reference background is updated with new images over time to adapt to dynamic scene changes.

Pixel is marked as foreground if the inequality is satisfied [3],

$$| It (P , Q) - Bt (P , Q) | > T \quad (3)$$

Where T is a pre-defined threshold. The background image Bt is updated by the use of a first order recursive filter as shown in Equation

$$Bt+1 = \square\ It + (1 - \square\ ) Bt \text{ alpha} \quad (4)$$

Where alpha is an adaptation coefficient. The basic idea is to provide the new incoming information into the current background image. After that, the faster new changes in the scene are updated to the background frame. However, alpha cannot be too large because it may cause artificial "tails" to be formed behind the moving objects. The foreground pixel map creation is followed b morphological closing and the elimination of small-sized regions.

2.5 Temporal Differencing:

Temporal differencing method uses the pixel-wise difference between two or three consecutive frames in video imagery to extract moving regions. It is a highly adaptive approach to dynamic scene changes however, it fails to extract all relevant pixels of a foreground object especially when the object has uniform texture or moves slowly [3]. When a foreground object stops moving, temporal differencing method fails in detecting a change between consecutive frames and loses the object. Let In(x) represent the gray-level intensity value at pixel position x and at time instance n of video image sequence I, which is in the range [0, 255]. T is the threshold initially set to a pre-determined value. Lipton et al.[3] developed two-frame temporal differencing scheme suggests that a pixel is moving if it satisfies the following [3]:

$$| In (x) - In-1 (x) | > T \quad (5)$$

This method is computationally less complex and adaptive to dynamic changes in the video frames. In temporal difference technique, extraction of moving pixel is simple and fast. Temporal difference may left holes in foreground objects, and is more sensitive to the threshold value when determining the changes within difference of consecutive video frames [5]. Temporal difference require special supportive algorithm to detect stopped objects.

III. Object Tracking:

Moving object tracking is the important part in human motion detection. It is high level computer vision problem. Tracking includes matching detected foreground objects between consecutive frames using different feature of object like motion, velocity, color, texture. Object tracking is the process to track the object over the time by locating its position in every frame of the video in surveillance system. It may also complete region in the image that is occupied by the object at every time instant [7]. In tracking approach, the objects are represented using the shape or appearance models [9]. The model selected to represent object shape limits the type of motion. For example, if an object is represented as a point, then only a translational model can be used. In the case where a geometric shape representation like an ellipse is used for the object, parametric motion models like affine or projective transformations are appropriate [15]. These representations can approximate the motion of rigid objects in the scene. For a non rigid object, silhouette or contour is the most descriptive representation and both parametric and nonparametric models can be used to specify their motion. Different object tracking methods are described as follows.

Centroid parameter : The Center of mass is vector of 1-by-n dimensions in length that specifies the center point of a parameter. For each point it is worth mentioning that the first element of the centroid is the horizontal coordinate (or x-coordinate) of the center of region, and the second element is the vertical coordinate (or y-coordinate) [16].

Point Tracking: Point tracking is robust, reliable and accurate tracking method developed by Veenman. This method is generally used to track the vehicles on road. In this approach requires good level of fitness of detected object.object which is tracking is based on point which is represented in detected object in consecutive frames and association of the points is based on the previous object state which can include object position and motion. This approach requires an external mechanism to detect the objects in every frame.

VI. Conclusion:

To recognise images and extract its information, image enhancement, motion detection, object tracking and behavior understanding researches have been studied. In this paper, we have studied and presented different technics of moving body detection, used in video surveillance. We have described background subtraction, temporal differencing, optical flow and frame differencing methods. Here research on object tracking can be classified as point tracking; Centroid based tracking according to the representation method of a target object.

VI. Reference

[1] M. Kass, A. Witkin, and D. Terzopoulos. Snakes: active contour models. Int. J. Comput. Vision 1, 321–332, 1988.

[2] V. Caselles, R. Kimmel, and G. Sapiro. Geodesic active contours. In IEEE International Conference on Computer Vision . 694–699, 1995.

[3] N. Paragios, and R. Deriche.. Geodesic active contours and level sets for the detection and tracking of moving objects. IEEE Trans. Patt. Analy. Mach. Intell. 22, 3, 266–280, 2000.

[4] Comaniciu, D. And Meer, P. 2002. Mean shift: A robust approach toward feature space analysis. IEEE Trans. Patt. Analy. Mach. Intell. 24, 5, 603–619.

[5] S. Zhu, and A. Yuille. Region competition: unifying snakes, region growing, and bayes/mdl for multiband image segmentation. IEEE Trans. Patt. Analy. Mach. Intell. 18, 9, 884–900, 1996.

[6] Elgammal, A. Duraiswami, R.,Hairwood, D., Anddavis, L. 2002. Background and foreground modeling using nonparametric kernel density estimation for visual surveillance. Proceedings of IEEE 90, 7, 1151–1163.

[7] Isard, M. And Maccormick, J. 2001. Bramble: A bayesian multiple-blob tracker. In IEEE International Conference on Computer Vision (ICCV). 34–41.

[8] S. Y. Elhabian, K. M. El-Sayed, "Moving object detection in spatial domain using background removal techniques- state of the art", Recent patents on computer science, Vol 1, pp 32-54, Apr, 2008.

[9] Yilmaz, A., Javed, O., and Shah, M. 2006. Object tracking: A survey. ACM Comput. Surv. 38, 4, Article 13,December 2006

[10] In Su Kim, Hong Seok Choi, Kwang Moo Yi, Jin Young Choi, and Seong G. Kong. Intelligent Visual Surveillance - A Survey. International Journal of Control, Automation, and Systems (2010) 8(5):926-939

[11] A. M. McIvor. Background subtraction techniques. Proc. of Image and Vision Computing, 2000.

[12] C. Stauffer and E. Grimson, "Learning patterns of activity using real time tracking," IEEE Trans. On Pattern Analysis and Machine Intelligence, vol. 22, no. 8, pp. 747-757, August 2000.

[13] K. Suganya Devi, N. Malmurugan, M. Manikandan, Object Motion Detection in Video Frames Using Background Frame Matching, International Journal of Computer Trends and Technology, Vol.4, Issue 6, pp 1928-1931, June 2013.

[14] Nan Lu, Jihong Wang, Q.H Wu, Li Yang, An Improved Motion Detection Method for Real-Time Surveillance, International Journal Of Computer Science, vol.1, Issue 6, 2008.

[15] Guangping Hu, HuijieLian, Pei Wang, Design and Research on the Motion Target Detection and Tracking Algorithm, (978-1-4244-76183)Information Science and Engineering, 2nd International conference, pp 1114-1117, Dec 2010.

[16] Xiaoping Li, Tao Li, Feng Ling, YinxiangLi, Research on Motion Detection and Target Contour Extraction in Wearable Video System, Asia Pacific Conference on Wearable Computing Systems, vol. 1, pp 90-93, 2010.

[17] C.LakshmiDevasena, R.Revathi, M.Hemalatha, Video Surveillance Systems, International Journal of computer science, (1694-0814), Vol. 8, Issue 4, No 1, pp 635-641, July 2011.