A Review on Object Detection in Video Processing

Kauleshwar Prasad, Richa Sharma and Deepika Wadhwani

BIT, Durg, India
kauleshwarprasad@gmail.com, sharma.richa157@gmail.com, deepika0141@gmail.com

Abstract

This paper initially proposes a technique for identifying a moving object in a video clip of stationary background for real time content based multimedia communication systems [2]. It deals with identifying an object of interest. Dynamic objects are identified using both background elimination and background registration techniques. Post processing techniques are applied to reduce the noise. The background elimination method uses concept of least squares to compare the accuracies of the current algorithm with the already existing algorithms. The background registration method uses background subtraction which improves the adaptive background mixture model and makes the system learn faster and more accurately, as well as adapt effectively to changing environments.

***Keywords**: Background Elimination, Background Registration, Post Processing*

1. Introduction

Video surveillance has long been in use to monitor security sensitive areas such as banks, department stores, highways, crowded public places and borders. The advance in computing power, availability of large-capacity storage devices and high speed network infrastructure paved the way for cheaper, multi sensor video surveillance systems. Traditionally, the video outputs are processed online by human operators and are usually saved to tapes for later use only after a forensic event. The increase in the number of cameras in ordinary surveillance systems overloaded both the human operators and the storage devices with high volumes of data and made it infeasible to ensure proper monitoring of sensitive areas for long times. In order to filter out redundant information generated by an array of cameras, and increase the response time to forensic events, assisting the human operators with identification of important events in video by the use of "smart" video surveillance systems [3] has become a critical requirement. The making of video surveillance systems "smart" requires fast, reliable and robust algorithms for moving object detection [4], classification, tracking and activity analysis.

The procedure of moving object tracking is to decide whether there exist objects moving in video and to position the target basically and recognize it. A video sequence is made of basically a series of still images at a very small interval time between each capture. As video sequence consists of frame sequences which have certain temporal continuity, the detection for moving object in video is conducted in a way that frame sequences are extracted from the video sequence according to a definite cycle.

Therefore, moving object detecting has something similar to object detection in still images. Only moving object detecting is more relying on the motion characteristics of objects, i.e. the continuity of time, which is the difference between moving object and object detection in still images. The need of real-time object detection for video surveillance has spawned a huge amount of our daily life, especially in some domains where it has received considerable attention, for instance: criminology, sociology, statistic, traffic accident detection and military applications. Moving object detection is considered to be the most important task in automated video surveillance systems. It represents the low level image processing technique which is the basic of automated video surveillance.

2. Methodology

In the present algorithm, we assume that the background is stationary for the video clips considered. The architecture and modeling of the proposed algorithm is shown in Figure 1.

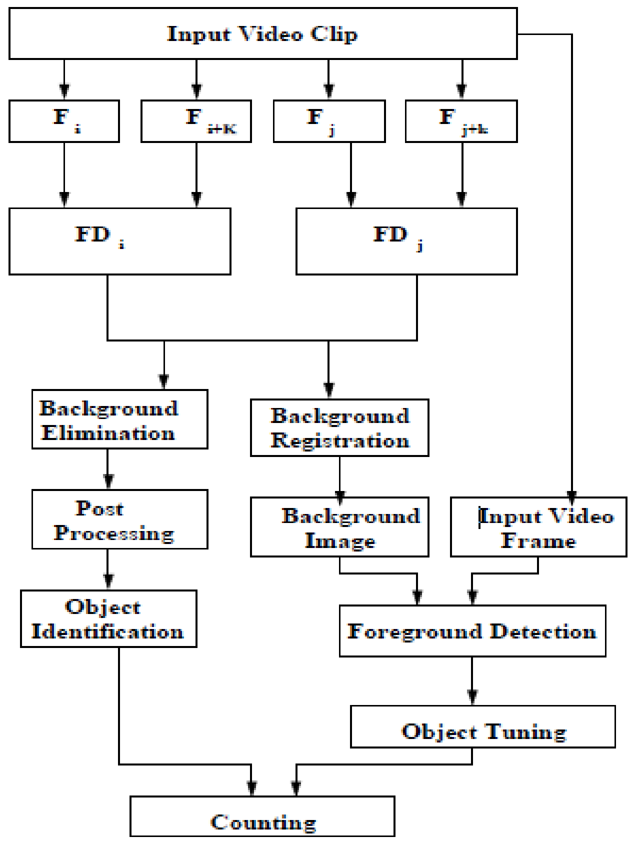


Figure 1. Architecture for Object Identification

The first step is to calculate the frame difference [3] mask by thresholding the difference between two consecutive input frames. At the same time, the background difference mask is generated by comparing the current input image and the background image stored in the background buffer. This background difference mask is our primary information for object shape generation.

In the second step, according to the frame difference mask of past several frames, pixels which are not moving for a long time are considered as reliable background in the background registration. This step maintains an up-to-date background buffer as well as a background registration mask indicating whether the background information of a pixel is available or not.

In the third step, an initial object mask is constructed from the background difference mask and the frame difference mask. If the background registration mask indicates that the background information of a pixel is available, the background difference mask is used as the initial object mask. Otherwise, the value in the frame difference mask is copied to the object mask. The initial object mask generated in the third step has some noise regions because of irregular object motion and camera noise. Also, the boundary region may not be very smooth.

In the last step, these noise regions are removed and the initial object mask is filtered to obtain the final object mask.

3. Moving Object Detection

3.1 Frame Difference

Frame differences method is a technique for detecting the motion area by making the difference between the current frame and the background frame. An image is divided into foreground and background in this method. The background is modeled, and the current frame and the background model are compared pixel by pixel. If there are n frames, then we will get (n/k) frame differences (FD). The differencing includes frame differencing and background differencing.

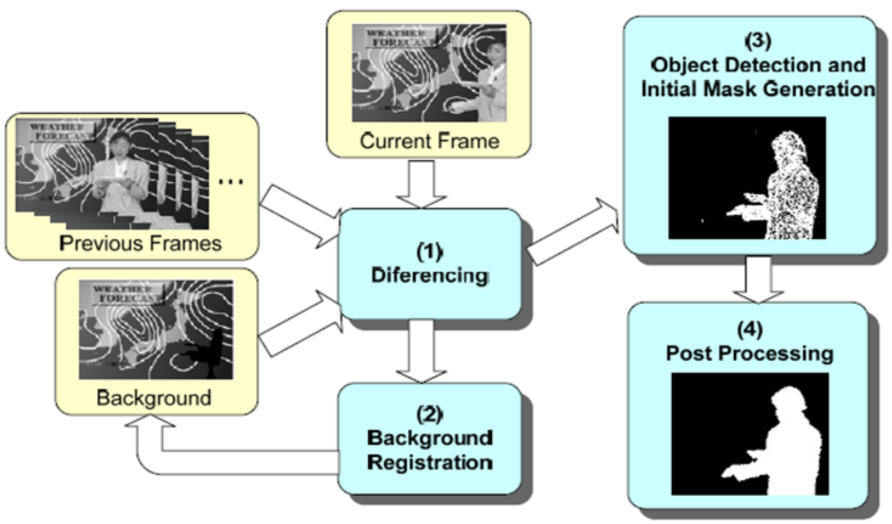


Figure 2. Block Diagram of Video Processing Using Background Registration

Thus, in the frame difference method, moving objects are extracted according to the differences among two or three continuous frames. The method is the most simple and direct, with which the changing part in video can be quickly detected.

The method using frame differences can better adapt to environment in intensive fluctuation, and can easily detect those pixels causing images to change distinctly when the target moves.

$$FD(x, y, t) = |I(x, y, t)-I(x,y,t-1)|$$

$$FDM(x, y, t) = 0 \text{ if } FD \geq TH \text{ or } 1 \text{ if } FD > TH \quad (1)$$

Where I is frame data, FD is Frame Difference, and FDM is Frame Difference Mask.

3.2 Background Elimination

Background elimination is a common method in moving object tracking algorithm, which is used more often in situations with relation still background. Once the frame differences are computed the pixels that belong to the background region will have a value almost equal to zero, as the background is assumed stationary. Many a times because of camera noise, some of the pixels belonging to the background region may not tend to zero. These values are set to zero by comparing any two frame differences, say, FD_i and FD_j. Thus, the background region is eliminated and only the moving object region will contain non-zero pixel values.

3.3 Background Registration

The goal of background registration step is to construct reliable background information from the video sequence. According to FDM, pixels not moving for a long time are considered as reliable background pixels. The procedure of Background Registration can be shown as where SI is Stationary Index, BI is Background Indicator, and BG is the background information. The initial values of BI, BG, and BI are all set to “0.” Stationary Index records the possibility if a pixel is in background region. If is high, the possibility is high; otherwise, it is low. If a pixel is “not moving” for many consecutive frames, the possibility should be high. When the possibility is high enough, the current pixel information of the position is registered into the background buffer. In addition, Background indicator is used to indicate whether the background information of current position exists or not.

Figure 3. Illustration of Background Registration

3.4 Post Processing

Many a times due to camera noise and irregular object motion, there always exists some noise regions both in the object and background region. Moreover the object boundaries are also not very smooth; hence a post processing technique is required. Most of the post processing techniques are applied on the image obtained after background elimination. Initially, order-statistics filters are used, which are the spatial filters and whose response is based on ordering (ranking) the pixels contained in the image area encompassed by the filter. The response of the filter at any point is then determined by the ranking result. The current algorithm uses Median filter which is the best-known order-statistics filter. This filter replaces the value of a pixel by the median of the gray levels in the neighborhood of that pixel. The formula used is

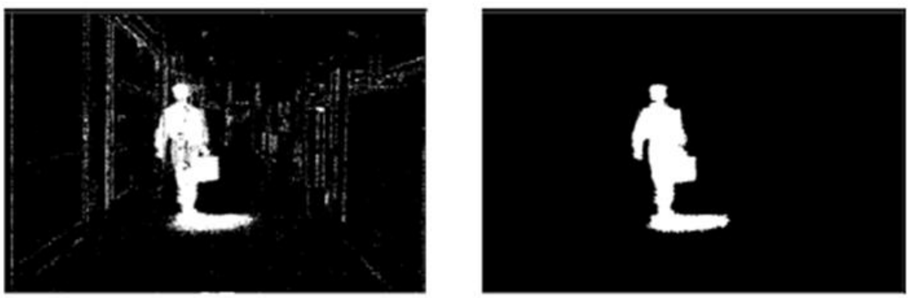
$$f(x, y) = \text{median}\{g(s, t)\} \tag{2}$$

After applying the median filter, the resulting image is converted into a binary image. The morphological opening technique is applied on this binary image. The opening of A by B is simply erosion of A by B followed by dilation of the result by B. This can be given as

$$A \circ B = (A \Theta B) \oplus B \tag{3}$$

Here, A is the image and B is a structuring element.
Some of the post processing techniques are:

3.4.1 Object Tuning: This is a post processing technique where we use a median filter for noise elimination in both the object and background. As the object boundaries are not very smooth, a post processing technique is required on the foreground image. The final output of the object tuning phase is binary image of the objects detected.



(a) (b)

Figure 4. Illustration of Post-processing. (a) Initial Object Mask; (b) After Object Tuning

3.4.2 Object Identification: The image obtained after the pre-processing step has relatively less noise, so, the background area is completely eliminated. Now, if the pixel values of this image are greater than a certain threshold, then, those pixels are replaced by the pixels of the original frame.

3.4.3 Object Counting: The tracked binary image mask1 forms the input image for counting. This image is scanned from top to bottom for detecting the presence of an object. Two variables are maintained, one that keeps track of the number of objects and the other contains the information of the registered object. When a new object is encountered, it is first checked to see whether it is already registered in the buffer, if the object is not registered then it is assumed to be a new object and count is incremented, else it is treated as a part of an already existing object and the presence of the object is neglected. This concept is applied for the entire image and the final count of objects is present in variable count. A fairly good accuracy of count is achieved.

4. Conclusion

In this paper, we propose an efficient algorithm for detecting a moving object using background elimination technique. It can be seen from analysis and examples that the computer language Matlab has the characteristics of simple programming, easy operation and high processing rate, etc. when used in series of processing of moving object detecting algorithm. Initially, we calculated the frame difference mask by thresholding the difference between two consecutive input frames. A background registration technique is used to construct reliable background information from the video sequence. Then, each incoming frame is compared with the background image. Finally, a post-processing step is used to remove noise regions and produce a more smooth shape boundary. As a result, the implementation of the algorithm becomes very fast.

References

[1] X. Chen, "Application of Matlab in Moving Object Detecting Algorithm", Department of Electronic Information Engineering Nanjing University of Information Science, **(2001)**.

[2] M. M. Velammal, "Vision Based Moving Object Tracking through Enhanced Color Image Segmentation using Haar Classifiers", **(2004)**.

[3] M. L. Lemmou and N. Dahnoun, "Real time object detection for a panning/tilting surveillance camera using an embedded dsp platform".

[4] R. Venkatesan, "Real Time Implementation On Moving Object Tracking and Recognisation Using Matlab", **(2001)**.

[5] P. Sharma, C. Huang and R. Nevatia, "Unsupervised Incremental Learning for Improved Object Detection in a Video".

[6] R. Jain, P. Sankar and C. V. Awahar, "Interpolation Based Tracking for Fast Object Detection In Videos", **(2011)**.

[7] M. Ammar, S. Le Hegarat-Mascle and R. Reynaud, "Video Scene Object Detection Using An A Contrario Approach".