International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 3 Issue 1, January - 2014

Background Subtraction Algorithm with Post Processing in Video Surveillance

(1)Ms. Rucha D. Pathari, (2)Asst. Prof. Sachin M. Bojewar
(1)*PG Scholar, Alamuri Ratnamala Institute of Engineering and Technology, Mumbai University*
(2)*Assistant Professor, Vidyalankar Institute of Technology, Mumbai University*

Abstract

In the field of motion estimation for video surveillance many techniques have been used . One of the common approach is to use generic method for background subtraction algorithm. This method has phases like preprocessing the video input file then use backgroungd subtraction algorithm onto it and then go for further operations. In this paper in generic method we have added a new phase called as post processing which will help to remove noise from the output video before it has been sent to display output. Using filters like Kalman filter or enhanced Kalman filter helps to remove noise from the video output file. Background subtraction is the one of the crucial step in detecting the moving object. Many techniques were proposed for detected moving object.

1. Introduction

Background subtraction, also known as Foreground Detection, is a technique in the fields of image processing and computer vision wherein an image's foreground is extracted for further processing (object recognition etc.). Generally an image's regions of interest are objects (humans, cars, text etc.) in its foreground. After the stage of image preprocessing (which may include image denoising etc.) object localisation is required which may make use of this technique. Background subtraction is a widely used approach for detecting moving objects in videos from static cameras. The rationale in the approach is that of detecting the moving objects from the difference between the current frame and a reference frame, often called “background image”, or “background model”.[1] Background subtraction is mostly done if the image in question is a part of a video stream.

Detection of moving object from a video sequence is crucial task in video surveillance. In literature many techniques was proposed for object extraction which can be classified into two categories: automatic and semiautomatic [2]. The former works without any human intervention while the latter requires user’s interaction. A common approach to identifying the moving objects is by using background subtraction techniques [3], The idea of background subtraction is consists of extracting moving objects as the foreground elements obtained from the image difference between each frame and the so-called background model of the scene. This model is used as a reference image to be compared with each recorded image. Consequently, the background model must be an accurate representation of the scene after removing all the non-stationary elements. It must be permanently updated to take into account the change in the lighting conditions or any change in the background texture. All this should be attainable even when foreground elements are permanently moving around in the scene at different speeds from the beginning of the video stream [4]. Many background subtraction algorithms are proposed from the literature however a problem of detecting moving object in complex environment is still not yet completely solved. Any good background algorithm must have the following characteristics (1) should adapt to various levels of illuminations at different times of the day (2) It must handle adverse weather condition such as fog or snow that modified the background. (3) It must handle the moving objects that first merge into background and then become foreground at a later time [3].

2. Steps for Background Subtraction

Most of the background subtractions algorithms follow a simple flow diagram defined by Cheung and Kamath [4] shown in Figure 1. The four major steps in a background subtraction algorithm are preprocessing, background modeling, foreground detection and data validation.

2.1. Preprocessing: In most computer vision systems, simple temporal and/or spatial smoothing is used in the early stage of processing to reduce camera noise. Smoothing can also be used to remove transient environmental noise such as rain and snow captured in outdoor camera. Forreal-time systems, frame-size and frame-rate reduction are commonly used to reduce the data processing rate. If the camera is moving or multiple cameras are used at different locations, image registration

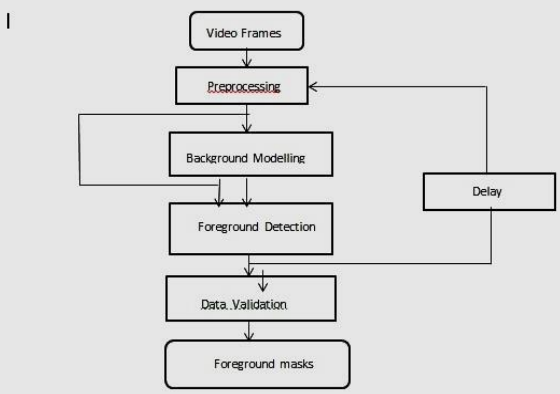

fig 1. Generic of Background and subtraction algothim [4]

between successive frames or among different cameras is needed before background modeling [5]. Another key issue in preprocessing is the data format used by the particular background subtraction algorithm. Most of the algorithms handle luminance intensity, which is one scalar value per each pixel. However, color image, in either RGB or HSV color space, is becoming more popular in the background subtraction literature [5].

2.2 Background Modeling: Background modeling is at the heart of any background subtraction algorithm. Much research has been devoted to developing a background model that is robust against environmental changes in the background, but sensitive enough to identify all moving objects of interest. In [4] background modeling techniques are classified into two categories: Non- recursive and recursive.

A non-recursive technique uses a sliding-window approach for background estimation. It stores a buffer of the previous L video frames and estimates the background image based on the temporal variation of each pixel within the buffer [3]. Non-recursive techniques are highly adaptive as they do not depend on the history beyond those frames stored in the buffer [3]. Frame difference, median filter, linear predictive filter and Non parametric model are some examples of Non recursive algorithms.

For recursive techniques, it does not maintain a buffer for background estimation. Instead, they recursively update a single background model based on each input frame. As a result, input frames from distant past could have an effect on the current background model [4]. Compared with non-recursive techniques, recursive techniques require less storage, but any error in the background model can stay for a much longer period of time. Some examples of algorithms found in this category are: Approximated median filter, Kalman filter and Mixture of Gaussians.

2.3.Foreground Detection: It identifies pixels in the video frame that cannot be adequately explained by the background model and outputs them as a binary candidate foreground mask. Foreground detection compares the input video frame with the background model and identifies candidate foreground pixels from the input frame. The most commonly used approach for foreground detection is to check whether the input pixel is significantly different from the corresponding background estimation [5]:

$|It(x,y) - Bt(x,y)| > T$

Another popular foreground detection scheme is to threshold based on the normalized statistics:

$\underline{|It(x,y) - Bt(x,y) - \mu_d|} > Ts\ \lambda_d$

where μ_d and λ_d are the mean and the standard deviation of $It(x,y) - Bt(x,y)$ for all spatial locations (x,y). T or Ts are foreground experiment which most schemes determined it experimentally.

2.4. Data Validation: It examines the candidate mask, detection algorithm where decisions are made independently at each pixel will generally be noisy, with isolated foreground pixels, holes in the middle of connected foreground components and jagged boundaries. Cheung and Kamath [5] define data validation as the process of improving the candidate foreground mask based on information obtained from outside the background model. Data validation phase is sometimes referred to as the post-processing phase of the foreground mask (pixels).

Table 1: Shows advantages and Disadvantages of difference backgr subtraction algorithm

Method	Advantages	Disadvantages
Frame Differencing	Easiest method. It performs well for static background.	It requires a background without any moving objects
Mixture Gaussian	Low memory requirement	It does not cope with multimodal backgrounds
Approximation Median	It performs better than Running Gaussian Average	Its computation requires a buffer with the recent pixel values

3. Background Subtraction Algorithms

A robust background subtraction algorithm should be able to handle lighting changes, repetitive motions from clutter and long-term scene changes. Although different, most BS techniques share a common denominator: they make the assumption that the observed video sequence I is made of a static background B in front of which moving objects are observed. With the assumption that every moving object is made of a color (or a color distribution) different from the one observed in B, numerous BS methods can be summarized by the following formula:

$$X_t(s) = \begin{cases} 1 & \text{if } d(I_{s,t}, B_s) > \tau \\ 0 & \text{otherwise} \end{cases}$$

where τ is a threshold, Xt is the motion label field at time t (also called motion mask), d is the distance between Is,t, the color at time t and pixel s, and Bs, the background model at pixel s. The main difference between several BS methods is how B is modeled and which distance metric d they use. In the following subsection, various BS techniques are presented as well as their respective distance measure.

Using frame differencing

Frame difference (absolute) at time *t* + 1 is

$$D(t+1) = |V(x,y,t+1) - V(x,y,t)|$$

The background is assumed to be the frame at time *t*. This difference image would only show some intensity for the pixel locations which have changed in the two frames. Though we have seemingly removed the background, this approach will only work for cases where all foreground pixels are moving and all background pixels are static.[6]

A threshold "Th" is put on this difference image to improve the subtraction.

$$|V(x,y,t) - V(x,y,t+1)| > \mathrm{Th}$$

(this means that the difference image's pixels' intensities are 'thresholded' or filtered on the basis of value of Th) The accuracy of this approach is dependent on speed of movement in the scene. Faster movements may require higher thresholds.

Mean filter

For calculating the image containing only the background, a series of preceding images are averaged. For calculating the background image at the instant *t*,

$$B(x,y) = \frac{1}{N}\sum_{i=1}^{N} V(x,y,t-i)$$

where *N* is the number of preceding images taken for averaging. This averaging refers to averaging corresponding pixels in the given images. *N* would depend on the video speed (number of images per second in the video) and the amount of movement in the video [7]. After calculating the background *B(x,y)* we can then subtract it from the image *V(x,y,t)* at time *t*=t and threshold it. Thus the foreground is

$$|V(x,y,t) - B(x,y)| > \mathrm{Th}$$

where *Th* is threshold. Similarlywe can also use median instead of mean in the above calculation of *B*(*x*,*y*).

Usage of global and time-independent Thresholds (same Th value for all pixels in the image) may limit the accuracy of the above two approaches.

4. Comparison of Generic of Background Subtraction Algorithm v/s Generic of Background Subtraction Algorithm with Post processing

In generic method we have added a new phase called as post processing which will help to remove noise from the output video before it has nbeen sent to display output. Using filters like Kalman filter or enhanced Kalman filter helps to remove noise from the video output file. Background subtraction is the one of the crucial step in detecting the moving object. Many techniques were proposed for detected moving object. Following figure gives us the idea to add this phase:

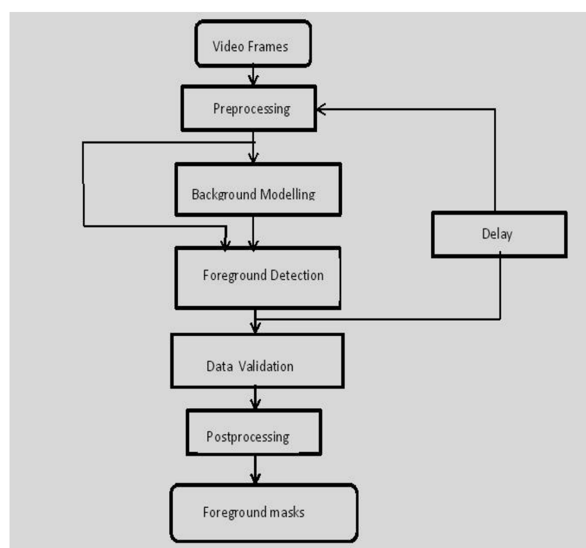

Fig 2: Proposed System

The differences of typical Kalman filter and enhanced kalman filter are highlighted by the involvement of occlusion scene determination and occlusion rate calculation which is making tracking of occlusion. The occlusion rate feature will be activated once the detection showing that the moving object is fully occluded. Besides that the object recognition capability is also added to enhanced Kalman filter so that the tracking target can be recognised from others. This feature is running all the time to recognise targeted object all the time.

1.The typical Kalman Filter:

The Kalman filter is a recursive estimator. This means only estimated state from previous time step and current measurement are needed to compute estimate for the current state. Therefore no history of observations or estimates is required.

The Kalman filter has two main features: a) its mathematical model is described in terms of state-space concepts. b) the solution is computed recursively.

Kalman filter is decribed by system state model and measurement model as shown in (1) and (2) respectively:

s(t) = O(t-1)s(t-1)+w(t) (1)

z(t) = H(t)s(t)+v(t) (2)

where, O(t-1) = state transition matrix

and H(t) = measurement matrix

w(t),v(t) = white Gaussian noise with zero mean

The Kalman filter has two phases:

a) Prediction Step – It is projecting forward the current state, obtaining prior estimate of the state $S^-(t)$.

b) Correction Step – This Phase is for feedback. It incorporates an actual measurement into prior estimate to obtain an improved posterior estimate $S^+(t)$, as shown below:

$$S^+(t) = S^-(t) + k(t)[z(t) - H(t)S^-(t)] \quad (3)$$

Where k(t) is weighting and described as shown in following equation

$$k(t) = p^-(t)H(t)^T[H(t)p^-(t)H(t)^T + R(t)]^{-1}$$

Where $p^-(t)$ = priori estimate error covariance and is defined by as follows :

$p^-(t)$ = E[$e^-(t)e^-(t)^T$]

Where, $e^-(t) = s(t) - s^-(t)$ …….prior estimate error

The posteriori estimate error will be given by -

$p^+(t)$ = E[$e^+(t)e^+(t)^T$] where $e^+(t) = s(t) - s^+(t)$ …...posteriori estimate error The prediction and correction steps will be executed recursively in the following sequence –

Prediction step:

$$s^-(t) = O(t-1)s^+(t-1)$$

$$p^-(t) = O(t-1)\,p^+(t+1)O(t-1)^T + Q(t-1)$$

Correction Step:

$$k(t) = p^-(t)H(t)^T[H(t)p^-(t)H(t)^T + \mathrm{R}(t)]^{-1}$$

$$s^+(t) = s^-(t)\,k(t)[z(t) - H(t)s^-(t)]$$

$$p^+(t) = [\,1 - k(t)\,H(t)]p^-(t)]$$

The prediction-correction cycle is repeated. From above equations we can say that measurement error R(t) and Kalman gain k(t) are in inverse ratio. However as priori estimate error $p^-(t)$ approaches zero the gain k(t) weights residual less heavily.

2. The Enhanced Kalman Filter

The algorithm is same as that of Kalman Filter just that occlusion rate is added into the correction step once the moving object detection in consecutive frame calculated is zero, i.e., the moving object is being occulted. Once moving object is being occluded, its consecutive predicted position will be relying on latest last two frames which are used for calculating rate of occlusion. Based on the rate of occlusion next position of occluded moving object will be predicted until the object is moving put from the occluded area. The occlusion rate is given as follows: OcclRate = x(t-1) +d(t)

Object recognition feature has been also added into the enhanced Kalman filter for tracking targeted moving object. This feature is useful for tracking condition that has multiple moving objects. The main advantage is estimated state from the previous time step and current measurement are needed to calculate estimate for current state. So, no history observations or estimates are required.

5. Conclusion

In object tracking using typical generic method can cause difficulties in tracking because of abrupt object motion, changing the appearance pattern, camera motion and object structures. And even it will not be able to track the system where objects are occluded. The post processing phase with Kalman filtering method successfully tracks the moving object position in moving object occlusion as well. It can handle occlusion scene handling, occlusion rate calculation and recognition feature. This off course leads into performance improvement and giving tracking result more accurate and precise.

6. References

1. M. Piccardi (October 2004). "Background subtraction techniques: a review". IEEE International Conference on Systems, Man and Cybernetics 4. pp. 3099–3104.doi:10.1109/icsmc.2004.1400815

2. Yoshida, T., 2004. Background differencing technique for image segmentation based on the status of reference

ISSN: 2278-0181
Vol. 3 Issue 1, January - 2014

pixels.” International Conference on Image Processing ICIP’04, pages 3487–3490, Singapore, October 24-27, 2004.

3. Elhabian, S., K. El-Sayed and S. Ahmed, 2008 Moving object detection in spatial domain using background removal techniques - State-of-Art. Recent Pat on Computer Sci., 1(1): 32-54.

4. Vargas, M., F. Barrero and J.M. Milla, 2008. An Enhanced Background Estimation Algorithm for Vehicle Detection in Urban Traffic Video Proceedings of the 11 International IEEE Conference on Intelligent Transportation Systems, Beijing, China, October 12-15, 2008.

5. Sen-Ching,S. Cheung and C. Kamath, 0000. Robust techniques for background subtraction in urban traffic video Center for Applied Scientific Computing Lawrence Livermore National Laboratory 7000 East Avenue, Livermore, CA 94550.

6. B. Tamersoy (September 29, 2009). "Background Subtraction – Lecture Notes". University of Texas at Austin.

7. Benezeth; B. Emile; H. Laurent; C. Rosenberger (December 2008). "Review and evaluation of commonly-implemented background subtraction algorithms". 19th International Conference on Pattern Recognition. pp. 1–4. doi:10.1109/ICPR.2008.4760998.