Hybrid Algorithm for Detection of Moving Object

Nikunj D. Jagani1

1PG Student

1,2Department of Computer Engineering

B. H. Gardi College of Engineering and Technology, Rajkot, India

Abstract— Motion detection is the process of detecting a change in position of an object relative to its surrounding or the change in the surroundings relative to an object. Detecting the moving objects relative to the whole image is the major task of it. Detecting moving objects is the foundation of other advanced applications, such as target tracking, targets classification and target behavior understanding. Input video sequence is given to the background subtraction model which updates background continuously. Here median filter and morphological operation is also done to remove the noise from the background. After doing this foreground is extracted from the video scene. After extracting the foreground, applied optical flow method and do analysis of motion vector of object. After motion vector analysis, calculate the threshold value using mean, standard deviation etc. After doing this detection of moving object is extracted from video scene. This both methods give better accuracy of object detection.

Key words: Background Subtraction, Block Matching, Frame Differencing Gaussian Mixture Model, Moving Object, Motion Vector, Optical Flow

I. INTRODUCTION

Nowadays, it is seen that surveillance cameras are already prevalent in commercial establishments, with camera output being recorded to tapes that are either rewritten periodically or stored in video archives. To extract the maximum benefit from this recorded digital data, detect any moving object from the scene is needed without engaging any human eye to monitor things all the time. Real-time segmentation of moving regions in image sequences is a fundamental step in many vision systems. Motion detection and object tracking algorithms are an important research area of computer vision and comprise building blocks of various high-level techniques in video analysis that include tracking and classification of trajectories. It is an obvious and biologically motivated observation that the main clue for detection of moving objects is the changing texture in parts of the view field.

Motion detection in consequent images is nothing but the detection of the moving object in the scene. In video surveillance, motion detection refers to the capability of the surveillance system to detect motion and capture the events. Motion detection is usually a software-based monitoring algorithm which will signal the surveillance camera to begin capturing the event when it detects motions. This is also called activity detection. An advanced motion detection surveillance system can analyze the type of motion to see if it warrants an alarm. In this a camera fixed to its base has been placed and is set as an observer at the outdoor for surveillance. Any small movement with a level of tolerance it picks is detected as motion.

There are mainly four methods to detect moving object from the video sequences i.e. 1) Background subtraction, 2) Optical flow, 3) Frame differencing, 4) Block Matching.

Background subtraction method uses the current frame minus the reference background image. The pixels where the difference is above a threshold are classified as the moving object. The Mixture of Gaussians method is widely used for the background modeling since it was proposed by Friedman and Russell. Stauffer presented an adaptive background mixture model by a mixture of K Gaussian distributions. Optical flow method can detect the moving object even when the camera moves, but it needs more time for its computational complexity, and it is very sensitive to the noise. The motion area usually appears quite noisy in real images and optical flow estimation involves only local computation [1].

Block matching techniques match blocks from the current frame with blocks from a reference frame. The displacement in block location from the current frame to the location in the reference frame is the motion vector. In fixed sized blocks, the main disadvantages that large blocks may fail to match the actual motion in a sequence, particularly along the moving edges, while small blocks require more overhead information [block matching]. Frame differencing attempts to detect moving regions by making use of the pixel-by-pixel difference of consecutive frames (two or three) in a video sequence. The temporal differencing approach [6] involves three important modules: block alarm module, background modeling module and object extraction module.

Here, the combination of two methods is used, first is Adaptive Gaussian Mixture Modeling, which is background subtraction method and second is Optical Flow. Adaptive GMM can be used in the context of a complex environment while Optical Flow can be used for quick calculation with simple background. GMM is not a complete object tracking while Object Flow provides complete computation tracking [1], so the complexity regarding the combination of two method is also need to take care.

II. DETECTION METHODS

A. Frame Differencing

Frame differencing is the simplest form of the background subtraction. The current frame is subtracted from the previous frame. The difference in pixel values for a given pixel is greater than a threshold Ts, the pixel is considered as a part of the foreground [2].

\[ |frame_i - frame_{i+1}| > T \]  

(1)

B. Gaussian of Mixture:

The Gaussian mixture model(GMM) is a single extension of the Gaussian probability density function(Gaussian PDF). As the GMM can approximate any
smooth shape of the density distribution. It is often used in image processing in recent years for good results. Suppose the Gaussian mixture model consists of and the combination of Gaussian probability density function, the Gaussian probability density function of each has its own standard deviation, mean, and weight. The weights can be interpreted by the corresponding Gaussian model of the frequency, they more often appear in the Gaussian model the higher the weight. The higher frequency of occurrence, then find the maximum weight on the Gaussian probability density function. Finally, we get the Gaussian probability density function of the means pixel value is background image [5].

Background subtraction is one of the most common methods of object segmentation. This process contains two steps: Background and Update model [6][8].

1) **Background Model:**
The basic concept of the Gaussian mixture model is as long as the number of Gaussian of mixtures, an arbitrary distribution can be in any of the precision is mixed with a weighted average of Gaussian approximation.

2) **Background Update:**
The known algorithms, if it is not updated, the step operation time will be very long, we must use the iterative method to update the standard deviation, mean and the weight to reduce the time required. In next steps before must set the basic parameters, as the number of Gaussian components, Number of background components are positive deviation threshold, learning rate between 0 to 1.

Although the computational complexity of the GMM is high but it can provide better results. If new entrants cannot be matched to any pixel of a Gaussian probability density function, update the pixel value of mean, then initialize the standard deviation and the weights [4][6].

3) **Algorithm of Gaussian Mixture Model:**
For the desired result, following steps were adopted and background subtraction methods are used for better understanding:

1) **Step-1** Initially, we compare each input image of pixels to the mean ‘µ’ of the associated components. If the value of a input image’s pixel is close enough to a chosen component’s mean, then that component is considered as the attached component. In order to be a matched component, the difference between the pixel and mean must be less than compared to the component's standard deviation scaled by factor D in the algorithm.

2) **Step-2** In the next step, update the mean, Gaussian weight and standard deviation to reflect the new obtained pixel values. In relation to non-matched components of the weights ‘w’ decreases whereas the standard deviation and mean stay the same. It is dependent upon the learning component ‘p’ in relation to how fast they change.

3) **Step-3** In the third step, we identify which components are parts of the background model. To do this a threshold value is applied to the component weights ‘w’.

4) **Step-4** In the final step, we determine the foreground pixels. Here the pixels that are identified as foreground do not match with any components determined to be the background[5][11].

4) **General Formula of Gaussian Mixture Model:**
The weighted sum of the means of the component densities. Where be the variable which represents the current pixel in frame, K is the number of distributions, and t represents time is an estimate of the weight of the ith Gaussian in the mixture at time t, µ is the mean value of the ith Gaussian in the mixture at time t, ω represents the current pixel in frame, K is the number of distributions, and t represents time (i.e., the frame index), ω is an estimate of the weight of the ith Gaussian in the mixture at time t, µ is the mean value of the ith Gaussian in the mixture at time t, ∑ is the covariance matrix of the ith Gaussian in the mixture at time t.

A Gaussian mixture model can be formulated in general as follows:

\[
P(\mathbf{X}_t) = \sum_{i=1}^{K} \omega_i \mathbf{N}(\mathbf{X}_t; \mu_i, \Sigma_i)
\]

Where, obviously,

\[
\sum_{i=1}^{K} \omega_i = 1
\]

\[
\mu_t = \sum_{i=1}^{K} \omega_i \mu_i
\]

That is, the weighted sum of the means of the component densities. Where be the variable which represents the current pixel in frame, K is the number of distributions, and t represents time (i.e., the frame index), ω is an estimate of the weight of the ith Gaussian in the mixture at time t, µ is the mean value of the ith Gaussian in the mixture at time t, ∑ is the covariance matrix of the ith Gaussian in the mixture at time t.

C. **Block Matching Algorithm:**
The block matching algorithm (BMA) is a standard technique for encoding motion in video sequences [8]. It aims at detecting the motion between two images in a block-wise sense. The blocks are usually defined by dividing the image frame into non-overlapping square parts. Each block from the current frame is matched into a block in the destination frame by shifting the current block over a predefined neighborhood of pixels in the destination frame. At each shift, the sum of the distances between the gray values of the two blocks is computed. The shift which gives the smallest total distance is considered the best match.

In the ideal case, two matching blocks have their corresponding pixels exactly equal. This is rarely true because moving objects change their shape in respect to the observer's point of view, the light reflected from objects' surface also changes, and finally in the real world there is always noise. Furthermore, from semantic point view, in scenes containing motion there are occlusions among the objects, as well as disappearing of objects and appearing of new ones.

Despite the problems of pixel by pixel correspondence, it is fast to compute and is used extensively for finding matching regions. Some of the most often used matching criteria based on pixel differencing are mean squared distance (MSD), mean absolute distance (MAD) and normalized cross-correlation(NCC) [7].

Block matching techniques can be divided into three main components as shown in Figure 1: block determination, search method, and matching criteria.

The first component, block determination, specifies the position and size of blocks in the current frame, the start location of the search in the reference frame, and the scale of the blocks. We focus on fixed size, disjoint blocks spanning the frame, with initial start location at the corresponding location of the block in the reference frame.
In tracking, a predictive method may be used. In tracking, a predictive method may be used to improve the start location of the search.

The search method is the second component, specifying where to look for candidate blocks in the reference frame. A fully exhaustive search consists of searching every possible candidate block in the reference frame. This search is computationally expensive and other search methods have been proposed to reduce the number of candidate blocks and/or reduce the processing for all candidate blocks. In this report, we concentrate on search methods that reduce the number of candidate blocks.

The third component is the matching criteria. The matching criteria are a similarity metric to determine the best match among the candidate blocks.

![Block Matching Flowchart]

**Fig. 1: Block Matching Flowchart**

In faster search methods, the best match so far will also determine the direction of the search (choice of next candidate blocks).

The motion vectors are fed to the block determination to implement multi resolution blocks. A coarse to fine resolution of the blocks is generated. The start location of the search at each resolution is the location of the best match (motion vector) from the previous coarser resolution.

The implementation of block matching using components allows for flexibility; interchanging components produces a large variety of block matching techniques. Based on the application, components which provide the best results can be chosen with ease.

**D. Optical Flow:**

The optical flow describes the direction and time rate of pixels in a time sequence of two consequent images. A two dimensional velocity vector, carrying information on the location and the velocity of motion is assigned to each pixel in a given place of the picture.

For making computation simpler and quicker we may transfer the real world three dimensional (3-D+time) objects to a (2-D+time) case. Then we can describe the image by means of the 2-D dynamic brightness function of location and time \( I(x, y, t) \). Provided that in the neighborhood of a displaced pixel, change of brightness intensity does not happen along the motion field, we can use the following expression:

\[
I(x+\delta x, y+\delta y, z+\delta z) = I(x, y, t) + \frac{\partial I}{\partial x} \delta x + \frac{\partial I}{\partial y} \delta y + \frac{\partial I}{\partial z} \delta z + H.O.T. \quad (6)
\]

From (5) and (6), with neglecting higher order terms (H.O.T.) and after modifications we get

\[
I_x V_x + I_y V_y = -I_t \quad (7)
\]

or in formal vector representation

\[
\nabla I \cdot \vec{v} = -I_t \quad (8)
\]

Where \( \nabla I \) is so-called the spatial gradient of brightness intensity and \( \vec{v} \) is the optical flow (velocity vector) of image pixel and \( I_t \) is the time derivative of the brightness intensity.

Equation (6) is the most important equation for optical flow calculation and is called 2-D Motion Constraint Equation or Gradient Constraint. It represents one equation with two unknown quantities.

Optical flow estimation is computationally demanding. At present there are several groups of methods for its calculation. All the methods come from (7) and consequently the presumption of conservation of brightness intensity. In this article our interest is concentrated in the differential methods.

The optical flow determination is solved by the calculation of partial derivatives of the image signal.

**E. Approximation Median:**

Assuming that the background is more likely to appear in a scene, we can use the median of the previous \( n \) frames as the background model:

\[
B(x, y, t) = \text{median}\{ I(x, y, t-1) \} \quad \text{and} \quad \text{median}\{ I(x, y, t-j) \} > Th \quad (9)
\]

**III. FLOWCHART OF HYBRID ALGORITHM**

Figure 1.2. shows the flow chart of hybrid algorithm for detection of motion object. As shown in figure 1.2, first of all input video sequence is given to the background subtraction model which updates background continuously. Here median filter and morphological operation is also done to remove the noise from the background. After doing this foreground is extracted from the video scene.

After extracting the foreground, applied optical flow method and do analysis of motion vector of object. After motion vector analysis, calculate the threshold value using mean, standard deviation etc. After doing this detection of moving object is extracted from video scene. This both method gives better accuracy of object detection.
As shown in figure 1.3, given input as an original video for detection of motion object. This original video is converted into the frames.

Figure 1.4 shows foreground is extracted from the image. Moving parts become white i.e. 1 in binary and stationary part becomes black i.e. 0 in binary.

As shown in figure 1.5 shows the calculation of the motion vector using optical flow method. It detects the movement of the pixels as a yellow spot using optical flow method.
As shown in figure 1.6, it gives the threshold output for calculating the motion object.

As shown in figure 1.7, it detects the motion object from the video frame with green rectangle boundary using Hybrid of both GMM method and Optical Flow method.

IV. CONCLUSION
Motion detection becomes attractive and crucial topic for computer vision community. Many researchers do a research related to moving objection detection in video.

There are many methods for object detection. In this research, combined approach of Gaussian mixture modeling and optical flow is successfully implemented. The detection is more accurate than the traditional background subtraction and optical flow method.

V. REFERENCES