RESEARCH ARTICLE OPEN ACCESS

A Comprehensive Review on Object Detection in Video Surveillance

Mamta Chopade [1], Pooja Chordiya[2], Shubhangi Godse[3] , Prof. Anjali Shejul [4]
Department of Information Technology
MITCOE, Savitribai Phule Pune University
Pune – India

ABSTRACT

Object detection and tracking are two fundamental tasks in computer vision applications such as surveillance, vehicle navigation, and autonomous robot navigation. In video surveillance, it is very important to detect the moving objects. This paper summarizes different algorithms for detection of moving objects. Some methods that are explained are background subtraction, colour histogram, SSIM index, frame difference. In this paper, we have evaluated strengths and weaknesses of different algorithms.

Keywords:- Object Detection, background subtraction, frame difference, color histogram. SSIM Index

I. INTRODUCTION

The ability to detect moving objects in a video stream is a very important factor in video surveillance. Video surveillance systems depend mainly on the ability to detect moving objects. Motion detection can be defined as action of sensing physical movement in a given specified environment. Object detection is performed to check if there is any existence of objects or if there is any change in an environment. The detected objects can be human, bicycles, buildings, vehicles or it can be environmental change [6].

For every tracking, method object detection mechanism is required for every frame in order to track object. Each image is segmented i.e. divided into cells by automatic analysis techniques [6]. In most of the object detection approaches information in a single frame is used. However, some object detection algorithms make use of information, which is computed from a sequence of frames in order to reduce false detections.

Different algorithm are discussed in this paper like background subtraction, frame difference, colour histogram, SSIM index. Along with explanation of algorithms, their strengths and weaknesses are also discussed. According to the application, we can select which algorithm should be used.

II. DETAILED ANALYSIS OF ALGORITHMS

In this paper, we have compiled different object detection methods along with their explanation, strengths and weaknesses.

1. Color Histogram

A histogram is graph that represents how many pixels are present at each level of color i.e. it represents distribution of colors in an image.

Color histogram is a way to model background and foreground image of a scene. Histogram comparison techniques are used to calculate similarity values, which help us to identify regions of interests.

Each video frame is divided into smaller regions called as cells to detect foreground objects. Histograms of each cell of background and foreground images are compared to detect moving object.

Color histogram algorithm for moving object detection step-by-step process [5]:

1. Compute background model using initial frame of video sequence.

2. New frames are compared with background model.

3. Overlay 40x40 pixel grid on top of each frame.

4. Compute histogram for each cell of background model.

5. Compute histogram for each cell of current frame.

6. Compare histograms of background and current frame on cell-by-cell basis to determine which cells contain foreground objects.

To compare two histograms (H_1 and H_2), first we have to choose a *metric* ($d(H_1, H_2)$) to express how well both histograms match.

Following are different formulas to compute histogram similarity:

- Correlation

$$d(H_1, H_2) = \frac{\sum_I (H_1(I) - \bar{H_1})(H_2(I) - \bar{H_2})}{\sqrt{\sum_I (H_1(I) - \bar{H_1})^2 \sum_I (H_2(I) - \bar{H_2})^2}}$$

where

$$\bar{H_k} = \frac{1}{N} \sum_J H_k(J)$$

and N is the total number of histogram bins.

- Chi-Square (CV_COMP_CHISQR)

$$d(H_1, H_2) = \sum_I \frac{(H_1(I) - H_2(I))^2}{H_1(I)}$$

- Intersection (method=CV_COMP_INTERSECT)

$$d(H_1, H_2) = \sum_I \min(H_1(I), H_2(I))$$

- BhattacharyyaDistance (CV_COMP_BHATTACHARYYA)

$$d(H_1, H_2) = \sqrt{1 - \frac{1}{\sqrt{\bar{H_1}\bar{H_2}N^2}} \sum_I \sqrt{H_1(I) \cdot H_2(I)}}$$

1.Background Subtraction Algorithm:

A method which is simple to implement by just subtracting the current frame from previous frame and obtaining threshold value of difference between given pixel value and obtained pixel value. If threshold value is greater than the given pixel it is considered as foreground. Most of background Algorithm follow simple path given in following Fig 1. The background subtraction algorithm has four main steps. The algorithm includes pre-processing, background modelling, foreground detection and data validation as four major steps.

Preprocessing:

If we use multiple cameras at different locations then we need to align the images of the successive frames. In pre-processing the colour or intensity of pixels of foreground, detection is major issue. In Background subtraction most of methods calculates luminance intensity and chrominance. Background subtraction is mostly use in the colour image.

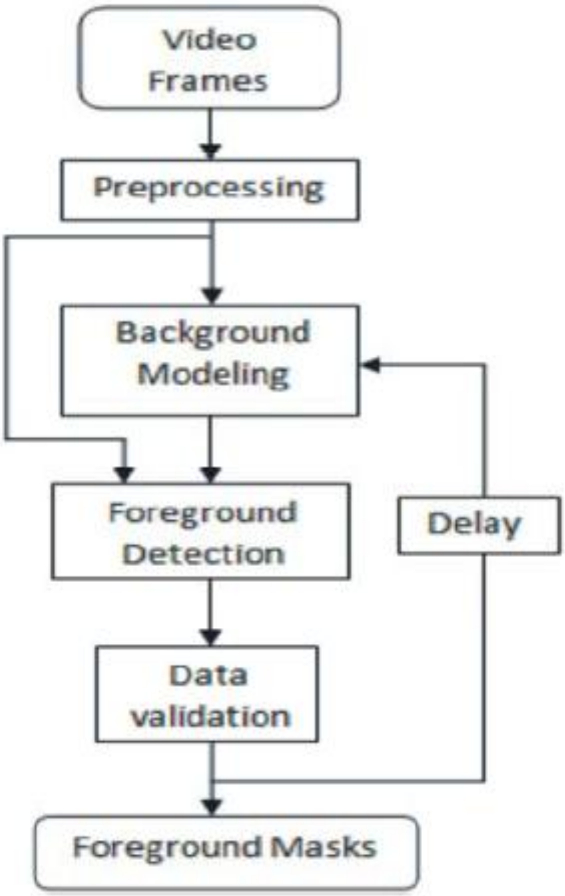

Fig 1. Flow diagram of Background subtraction algorithm.

Background Modelling:

In [2] background Modelling follow mainly two techniques are non-recursive and recursive. In non-recursive, approach of background modelling uses a sliding-window approach for background estimation. In the sliding window protocol, it stores buffer of the previous specific video frames and estimates the image based on each pixel within buffer. Frame difference method is based on non-recursive technique.

The recursive approach of background modelling does not maintain buffer instead; it recursively updates the background model for each input frame. From two techniques of background modelling recursive technique, require less storage. The examples of recursive technique of background modelling are Kalman filter and Mixture of Gaussians.

Foreground Detection:

Foreground detection compares the input video frame with background model and identifies the foreground pixel in input frame. These checks whether pixel is different from background estimation or not. Formula to be used:

$|I(x, y) - B(x, y)| > T$

Where T is the threshold value,

I (x, y) Current input video frame,

B (x, y) Background Model.

Threshold value is determined experimentally most of the times.

Data Validation:

It eliminates the pixels, which do not correspond to actual moving objects, and finally outputs foreground frame.

Pros:

A very widely used method, which is simple to implement.

Objects are allowed to become a part of the background without destroying the existing Background.

It learns itself and does not need to be reprogrammed.

It can be implemented in any applications.

Provides fast recovery.

Low memory requirement.

Cons:

Highly inaccurate.

Cannot deal with quick changes.

Not a good subtraction when shadow or any other obstacles are there.

Gives false positives

It does not survive with multimodal background

3. Frame Difference Method

In Fig 2, [3] one background image having not a single moving object is takes as the reference image. Then we calculate the pixel value for each co-ordinate (x, y) of the background image. The calculated value is subtracted from the corresponding pixel value of the input image. If the resulted value is greater than a particular threshold value, then we assume there is change in foreground pixel otherwise background.

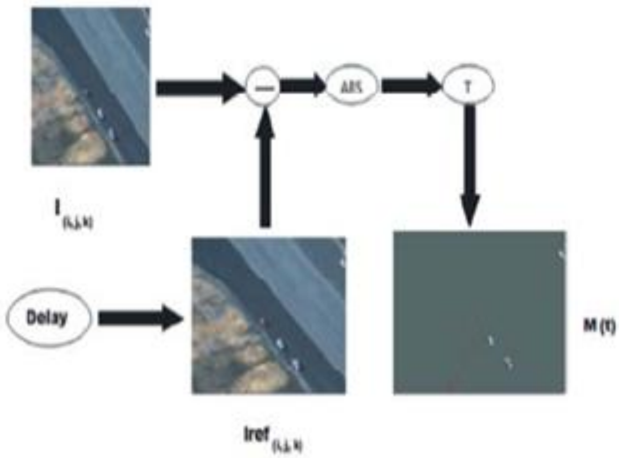

Fig 2. Frame difference method.

For each frame input I and for the reference image Ir, if for a particular pixel, | I -Ir | > T then that pixel is classified as foreground. That is I (i, j, k) -Ir (i, j, k) > T where,

I (i, j, k) is the co-ordinate of I (i, j) pixel value for kth color for the current input image I.

Ir (i, j, k) is for reference frame.

Pros:

Perform well for static background.

High accuracy.

Easiest method.

Cons:

It must require background without moving object.
Method having Computational time low to moderate.

4 SSIM index (Structure Similarity Index)

The structural similarity (SSIM) index [7] is a method for measuring the similarity between two images. The SSIM index is a full reference metric. The measuring of image quality based on an initial uncompressed or distortion-free image as reference. The traditional methods like peak signal-to-noise ratio (PSNR) and mean squared error (MSE) have proven to be inconsistent with human eye perception; SSIM is designed to improve these inconsistencies.

Here SSIM will be used to find the difference between two images. So that we will be able to find the moving object.

The SSIM index is calculated on various windows of an image. The measure between two windows x and y of common size *N*×*N* is:

$$\mathrm{SSIM}(x,y) = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)}$$

Where,

- μ_x the average of x;
- μ_y the average of y;
- σ_x^2 the variance of x;
- σ_y^2 the variance of y;
- σ_{xy} the covariance of x and y;
- $c_1=(k_1L)^2$, $c_2=(k_2L)^2$ two variables to stabilize the division with weak denominator;
- L the dynamic range of the pixel-values (typically this is $2^{\#bits\ per\ pixel}-1$);
- $k_1=0.01$ and $k_2=0.03$ by default.

The resultant SSIM index is a decimal value between -1 and 1, and value 1 is only reachable in the case of two identical sets of data.

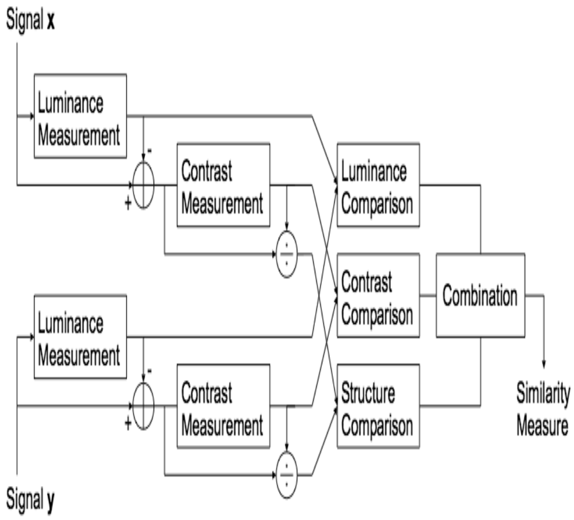

Fig 3. Logical working flow of SSIM

- The SSIM maps are generated from two adjacent frames.
- The mean value of the SSIM map gives an index value, which indicates how different the two images are

III. CONCLUSION

In this paper, we have studied various object detection methods. Available methods for these phases have been explained in details and a number of shortcoming and limitations were highlighted Colour histogram is fast and reliable. It does not require prior training. In detection phase, it compares the histograms computed for the cells of current image with histograms of cells of background image. Background subtraction is simplest method to detect objects, requires less memory. It is fast to recover and it does not need to be reprogramed. However, it does not support quick changes in background. It does not give accurate result in multimodal background and when shadows or obstacles are present in background image. Frame difference is easy and work well for static background. It is highly accurate. However, it requires background without moving objects. SSIM index is used to detect moving object by calculating the similarities between two adjacent images.

REFERENCES

.

[1] Yoshida, T., 2004. Background differencing technique for image segmentation based on the status of reference pixels." International Conference on Image Processing ICIP'04, pages 3487–3490, Singapore, October 24-27, 2004.

[2] Performance Comparison of Background Estimation Algorithms for Detecting Moving Vehicle Mahmoud Abdulwahab Alawi, *Othman O. Khalifa and M.D. Rafiqul Islam* Department of Electrical and Computer Engineering, International Islamic University Malaysia

[3] A Review on Object Detection and Tracking Methods Payal Panchal1, Gaurav Prajapati2, Savan Patel3, Hinal Shah4 and Jitendra Nasriwala5 DCST, Uka Tarsadia University, Bardoli, Gujarat, India

[4] Sen-Ching, S. Cheung and C. Kamath, 0000. Robust techniques for background subtraction in urban traffic video Center for Applied Scientific Computing

[5] Using Histogram to Detect and Track Objects in Color Video Michael Mason and Zoran Duric Department of Computer Science George Mason University VA22030

[6] Performance evaluation of object detection algorithms for video surveillance Jacinto Nascimento⋆, *Member, IEEE* Jorge Marques

[7] Registration of IR and EO Video Sequences based on Frame Difference Zheng Liu and Robert Lagani`ere School of Information Technology and Engineering University of Ottawa 800 King Edward Ave. Ottawa, ON K1N6N5 Canada E-mail: laganier@site.uotta.ca