A Survey on Moving Object Detection in Static and Dynamic Background for automated video analysis

Ms. Dipali Shahare [1] Ms. Ranjana Shende [2]
[1]M.Tech. [2]Lecturer
[1, 2]Department of Computer Science & Engineering
[1, 2] G. H. Raisoni Institute of Engineering and Technology For women's, Nagpur, India

Abstract— Detection of moving objects in a video sequence is a difficult task and robust moving object detection in video frames for video surveillance applications is a challenging problem. Object detection is a fundamental step for automated video analysis in many vision applications. Object detection in a video is frequently performed by object detectors or background subtraction techniques. Frequently, an object detector requires manual labeling, while background subtraction needs a training sequence. To automate the analysis, object detection without a separate training phase becomes a critical task. This paper presents a survey of various techniques related to moving object detection and discussed the optimization process that can lead to improved object detection and the speed of formulating the low rank model for detected object.

Key words: Object Detection, Soft Impute method, Markov Random Field, Temporal Differencing, Moving object extraction, background subtraction.

I. INTRODUCTION

Automated video analysis is important for many vision applications [11]. There are three steps for video analysis: object detection, object tracking, and actions recognition. Object detection aims to find and segment interesting objects in a video sequences. Then, such objects can be tracked from frame to frame, and then those track objects can be analyzed to distinguish object behavior. Thus, object detection acting a vital role in practical applications.

The primary goal of this paper is to critically discuss the various techniques for detecting moving objects methods in static and dynamic background in video. A second goal is to present a technique for formulating low rank model for detected object.

The paper has been structured by following manner: section 2 we discuss existing approaches for Moving Object Detection techniques, while section 3 discuss the proposed method for detecting object accurately and section 4 is summarized in the conclusions.

II. MOVING OBJECT DETECTION TECHNIQUE

Detection and extraction of moving object form a video sequences is used in various application like Video surveillance system, Traffic monitoring , Human motion capture, Situational awareness, Border protection and monitoring ,Airport safety.

Moving object can be detected from video sequences of either a fixed or a moving camera.

The main purpose of foreground detection is to distinguishing foreground objects from the stationary background. The automatic detection of moving objects in monitoring system desires efficient algorithms. The familiar technique is background subtraction i.e. to subtract current frame from background. When brightness difference between moving objects and background is small then it became difficult detect the difference between moving object.

There are several methods to detect moving objects, which are given below:

A. Optical Flow Method

Optical flow method is a complex and bad anti-noise performance, and it cannot be applied to real-time processing without special hardware device. [13] Proposes an automatic extraction technique of moving objects using x-means clustering. In this proposed method, the features are extracted from a current frame, and the clustering method, x-means classifies the feature points based on their estimated affine motion parameters. The segmented region is labeled; that region is obtained by morphological watershed algorithm. The segment labeling result characterize the moving object extraction.

B. Consecutive Frames Subtraction

Consecutive Frames Subtraction is a simple operation, realizes easily and has strong adaptability on the dynamic changes in the environment. But it cannot be completely extracted moving targets. [14] proposes a novel method for extracting moving objects from video sequences, which is based on Gaussian mixture model and watershed, is proposed where first the difference between neighboring frames is calculated and is described by a Gaussian mixture model, then divided into moving areas and background by improved Expectation-Maximization (EM) algorithm.

C. Background Subtraction

Background subtraction is a common method for detecting moving objects and it has been widely used in many surveillance systems, but it is yet a difficult problem to distinguish moving objects from backgrounds when these backgrounds change significantly. Separating foreground from background in a video sequence is one of the most fundamental tasks in many applications of computer vision. To detect moving objects, each incoming frame is compared with the background model learned from the previous frames to divide the scene into foreground and background. Therefore, background modeling has been actively investigated in the past decade. The difficulty encountered in background modeling is that the outdoor backgrounds are usually non-stationary in practice. Broadly speaking, there are two categories of online methods to model the background. The first one models the background using a single model per pixel, whereas the second one employs multiple models per pixel. Background subtraction is a widely used approach for detecting moving objects from static cameras [16].

The four most important steps in a background subtraction Algorithm are:

1) Pre-processing
2) Background Modelling
3) Foreground Detection
4) Data validation

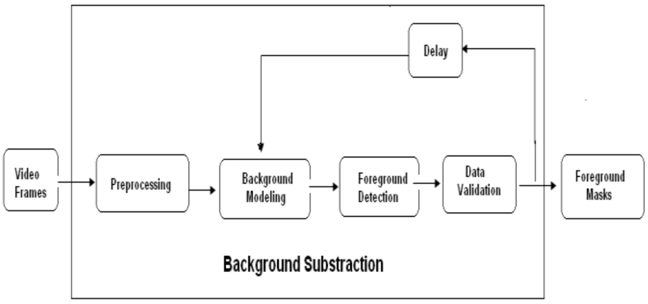

Fig. 1: Illustration of Background Subtraction

In background subtraction, background model can be obtained from a training sequence that does not contain foreground objects.

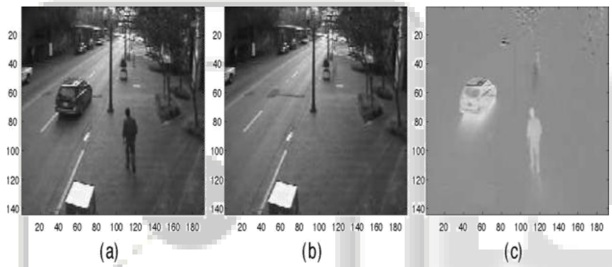

Fig. 2: Decomposition-based background subtraction: (a) an input image, (b) rebuild image after projecting input image, (c) difference image.

1) Color and Edge Information

Jabri, et al. [4] proposed an approach in which background modeling and subtraction approach are used to detect a human in the video images. This approach is used to segment the person from the background by computing the mean image for all video sequences. The incoming frame is subtracted from the mean image to identify the pixels which have changed the color. However, the problem with this approach is both the color and edge channel are subtracted separately before finding the result and, as a consequence, the computational time increases.

2) Standard Subtraction

The method developed by Davis and Taylor [5] is a motion-based method for differentiating normal walking movements at multiple speeds when atypical or non-walking locomotion is involved. Human walking movements are detected using low level regularities and constraints. The person's shape in each video frame is extracted with standard background subtraction. This approach locates the head, waist and feet using the W4 approach [6]. Standard subtraction techniques, which use RGB pixel differences, dilations and removal of small pixel region, are employed. The centroid of the outline pixel is called the head pixel, while the mean value of silhouette pixels in the torso region is called the waistline. The waistline is divided into two halves in order to find information concerning to the feet. Dynamic regularity features are calculated using cycle time, stance/swing ratio and double support time. Dynamic regularity features are independent of the camera position, but this approach uses view-based constraint of extension angle, which is suitable for non-walking locomotion and not for other regular locomotion's.

3) Object Extraction

The algorithm proposed by Yoginee, et al. [12] has moving object segmentation, blob analysis and tracking. Blob analysis is used to count the vehicle from which the speed and flow are calculated. Boundary Block Detection (BBD) algorithm is used for moving object detection by identifying the blocks which contain the moving objects boundaries. The system requires the model background with no moving objects and scene which contain moving objects. The system finds the boundary of the moving objects and the number of moving objects from a given video scene. Aviread function [13] is used to extract all frames in the video. Background subtraction extracts the object, while the pixels of the background model image are used as threshold. All images are divided into two parts, viz., background and foreground in binarization. The new video frame was subtracted from those background images, if the pixel difference is higher than the threshold, that images are foreground or object. If the pixel significantly differs from the background image, then the pixel is marked as a moving object. Each image frame must update the threshold level. To count the moving object flow, the algorithm tracks each vehicle within successive image frames. This algorithm works only for the videos obtained from fixed cameras and which has the normal background and stable videos. The algorithm can be modified to work on complex background and videos that are not stable. In addition, the performance can be improved by using optimizing algorithm such as fuzzy logic and neural network.

4) Gaussian Mixture

A Gaussian Model calculates each pixel-value from all the sample pixels' mean and variance. The model will remove all pixels that are outside of lower bound and upper bound of a norm. If a video is to run for a longer period of time, the pixels' averages will the same to the background's value unless the foreground object stays static. This is a common method for real-time segmentation of moving regions in frame sequences. Model Gaussians are updated using K-means approximation method. Each pixel is then evaluated and classified as a moving region or as a background.

Stauffer and Grimson [3] presented a novel adaptive online background mixture model that can robustly deal with illumination changes, monotonous motions, clutter, introducing or removing objects from the scene and slowly moving objects. Their inspiration was that a unimodal background model could not handle image acquisition noise, light change and some surfaces for a particular pixel at the same time. Thus, they use a mixture of Gaussian distributions to symbolize each pixel in the model.

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 an exceedingly 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 gradually. Temporal differencing method fails in detecting a change between consecutive frames and loses the object when a foreground object stops moving,

Let Frame i represent the gray-level intensity value at pixel position i 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 developed two frames temporal differencing scheme suggests that a pixel is moving if it satisfies the following:

$$| \text{Frame } i - \text{Frame } i\text{-}1| > th$$

This estimated background is just the previous frame. It evidently works only in particular condition of objects speed and frame rate and very sensitive to the threshold.

This method is computationally less complex and adaptive to dynamic changes in the video frames. In temporal difference technique, taking out 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 [2]. Temporal difference require special supportive algorithm to detect stopped objects.

Comparison of several popular methods for moving object detection:

Optical Flow method	**Consecutive Frames Subtraction**	**Background Subtraction**
Complex and bad anti-noise performance	simple operation, realizes easily	provides a moving object comprehensive and reliable Information
cannot be applied to real-time processing without special hardware device	has strong adaptability on the dynamic changes in the environment	very sensitive to the irradiation which is caused by dynamic scene changes
		Advantage of not need previous information of moving objects such as shapes or movements **Disadvantage** cannot discriminate moving objects from backgrounds when these backgrounds change significantly

III. THE PROPOSED METHODS

In this section, it integrates the object detector and background subtraction in to the single process of optimization which can work efficiently for moving object detection.

Moving Object detection is the basic step for further analysis of video. All tracking method needs object detection methods either in every frame or when the object first appears from stationary background object.

When working with video data, it can be helpful to select a representative frame from video and the methods can be applied to the processing of all the frames in the video. The method computes the estimated foreground and background model of frame specified by rank. As illustrates in figure is as follows

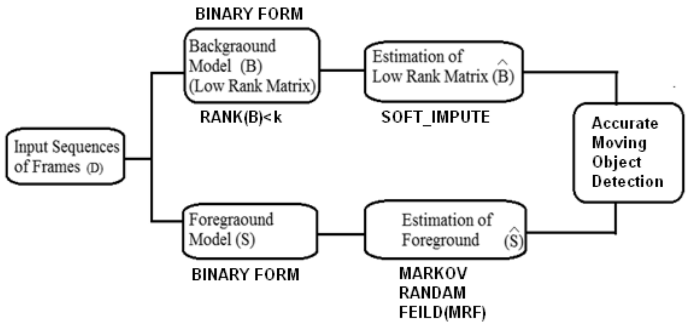

Fig. 3: Framework for Moving Object Detection in Low Rank Model

Here first consider the background model that uses the SOFT_IMPUTE algorithm to formulate the low rank model for background and secondly the foreground model uses the Markov Random Fields method to compute the Foreground.

A. Detecting moving objects from video sequences of a fixed camera

Background refers to a static scene and foreground refers to the moving objects. Objective is to estimate the foreground support as well as underlying background images. To compose the crisis well posed, we have the following models to explain the foreground and the background model.

1) Background Model

The background intensity should be unchanged over the sequence except for variations arising from illumination change or periodical motion of dynamic textures. Thus; background images are linear associate with each other, forming a low-rank matrix B. The only Constraint on B is:

$$\text{rank}(B) <= K;$$

Where, K is a constant to define the complexity of the background model.

To formulate the background model, the SOFT-IMPUTE [10] method is used which produces a sequence of solutions for which the criterion decreases to the optimal solution with every iteration and the successive iterates get closer to the optimal set of solutions of the problem. In many applications measured data can be represented in a matrix $Xm{\times}n$, for which only a relatively small number of entries are observed. The problem is to “complete” the matrix based on the survey entries, and has been dub the matrix completion problem.

2) Foreground Model

The foreground is defined as any object that moves differently from the background. Foreground motion provide intensity changes that cannot be fitted into the low-rank model of background. Generally, we have a prior that foreground objects should be contiguous pieces with relatively small size.

To compute the foreground model, Markov random field (MRFs) methods are used. Due to utilization of the relativity of every pixel of an image, the Markov Random Field (MRF) model is effective in solving the problem of detecting moving object under a complex background.

The Markov Random Fields (MRFs) [9] is statistical model, which used for restore the true image; images are often treated as realizations of a random process and MRFs to improve the accuracy of detecting contiguous outliers.

B. Detecting moving objects using moving camera

In this section we discussed the dynamic background (motion caused by moving camera). Here the 2D parametric transforms.

Method is used for moving camera; motions must be due to moving objects. The camera-induced image motion depends both on the ego motion parameters and the depth of each point in the scene. Estimating all of these physical parameters to relation for the camera-induced motion is, in general, an inherently ambiguous problem [7]. When the scene contains huge depth variations, these parameters may be improved. We refer to these scenes as *3D scenes*. However, in *2D scenes*, namely, when the depth variations are not significant, the recovery of the camera and scene parameters is usually not robust or reliable.

An effective approach to accounting for camera-induced motion in 2D scenes is to model the image motion in terms of a global 2D parametric transformation. This method is strong and consistent when applied to flat scenes, distant scenes, or when the camera is undergoing only rotations and zooms.

Therefore, 2D algorithms and 3D algorithms address the moving object-detection problem in very different types of scenarios. These are two extremes in a continuum of scenarios: flat 2D scenes (i.e., *no 3D parallax*) vs. 3D scenes with dense depth variations (i.e., *dense 3D parallax*). Both classes fail on the other extreme case or even on the intermediate case [7].

Our techniques are based on a stratification of the moving object-detection problem into scenarios which progressively rise in their complexity:

1) Situation in which the camera-induced motion can be modelled by a single 2D Parametric Transform
2) Those in which the camera-induced motion can be modeled in terms of a small number of *layers* of parametric transformations, and
3) Common 3D scenes, in which a more complete parallax motion analysis is required.

IV. CONCLUSION

In this paper, we discussed a variety of techniques to detect moving object in video frames. Amongst the methods reviewed, the background subtraction method; the subtraction of color and edge channels are performed separately before finding out the result. It is not robust against changes in illumination. It cannot detect non stationary background object such as swinging leaves, rain snow and shadow cast by moving object. Furthermore, in this paper, we have proposed a single process of optimization which integrates the object detection and background learning which can be used to detect the moving object accurately, such that the time and accuracy attributes can be improved.

REFERENCES

[1] 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.
[2] A Survey on Moving Object Detection and Tracking in Video Surveillance System Kinjal A Joshi, Darshak G. Thakore International Journal of Soft Computing and Engineering (IJSCE) ISSN: 2231-2307, Volume-2, Issue-3, July 2012.
[3] C. Stauffer and W. Grimson. Adaptive background mixture models for realtime tracking. In Proc. of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, page 246252, 1999.
[4] Sumer Jabri, Zoran Duric, Harry Wechsler, Azriel Rosenfeld, “Detection and Location of People in Video Images Using Adaptive Fusion of Color and Edge Information,” In Proc. 15th Int’l Conf. on Pattern Reg, 2000,vol. 4,pp. 627 – 630.
[5] James W. Davis, Stephanie R. Taylor, “Analysis and Recognition of Walking Movements,” In Proc.16th Int’l Conf. on pattern Recognition, 2002, vol.1, pp. 315 – 318.
[6] I. Haritaoglu, D. Harwood, and L. Davis, “W4: Who? When? Where? What? A real time system for detecting and tracking people”. In Proc. Int. Conf. Auto. Face and Gesture Recog., 1998, pages 222– 227.
[7] A Unified Approach to Moving Object Detection in 2D and 3D Scenes Michal Irani and P. Anandan IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE, VOL. 20, NO. 6, JUNE 1998 577.
[8] *Shireen Y. Elhabian*, Khaled M. El-Sayed and Sumaya H. Ahmed,” Moving Object Detection in Spatial Domain using Background Removal Techniques - State-of-Art”, Recent Patents on Computer Science, 2008.
[9] Xiaowei Zhou, Can Yang, and Weichuan Yu,” Moving Object Detection by Detecting Contiguous Outliers in the Low-Rank Representation”, IEEE TRANSACTIONS ON PATTERN ANLYSIS AND MACHINE INTELLIGENCE, IEEE, March 2013.
[10] R. Mazumder, T. Hastie, and R. Tibshirani, “Spectral Regularization Algorithms for Learning Large Incomplete atrices,” J. Machine Learning Research, vol. 11, pp. 2287-2322, 2010.
[11] A. Yilmaz, O. Javed, and M. Shah, “Object Tracking: A Survey,” ACM Computing Surveys, vol. 38, no. 4, pp. 1-45, 2006.
[12] Yoginee B. Bramhe (Pethe), P.S. Kulkarni, “An Implementation of Moving Object Detection, Tracking and Counting Objects for Traffic Surveillance System,” Int’l Conf. on Computational Intelligence and Comm. Networks (CICN), 2011, pp. 143 – 148.

[13] Imamura.K, Kubo.N, Hashimoto.H, "Automatic moving object extraction using x-means clustering Picture Coding Symposium (PCS),pp246 - 249 , Dec 2010.
[14] R. Li, S. Yu, and X. Yang, "Efficient spatio-temporal segmentation for extracting moving objects in video sequences," IEEE Transactions on Consumer Electronics, vol. 54, pp. 1161-1 167, Mar 2007 .
[15] R. Li, S. Yu, and X. Yang, "Efficient spatio-temporal segmentation for extracting moving objects in video sequences," IEEE Transactions on Consumer Electronics, vol. 54, pp. 1161-1 167, Mar 2007 .
[16] A. M. McIvor. “Background subtraction techniques”, In Proc. of Image and Vision Computing, Auckland, New Zealand, 2000.

