ISSN:2321-0613

Detection and Tracking of Moving Object in Video - A Survey

Dhaval Deshpande[1] Nikhil Aatkare[2] Prof.Reena Somani[3]
[1,2,3]Department of Information Technology Engineering
[1,2,3]Atharva College Of Engineering,Mumbai - 400095, Maharashtra, India,University of Mumbai.

Abstract— Object detection and object tracking in video sequence is very important concept and has various applications like security and surveillance, traffic control, medical imaging or wild life observation. The objective of this paper focuses on review of various methods for detection of moving object in video and then tracking of that object in video sequence. Object detection is very challenging task. It is very important detect object from video sequence before you can track it. Object tracking required in higher level applications that require the location and shape of object in every frame. In this paper, we have reviewed Background subtraction, statistical method and Temporal frame differencing to detect moving object. We have also reviewed tracking method based on point tracking, kernel tracking and silhouette tracking.

Key words: Object detection, object tracking, video sequence and Background subtraction

I. INTRODUCTION

Surveillance of large “structures” is a major modern concern for governments and companies. Energy production centers, transportation infrastructures, food and water supply centers or storage facilities of sensitive classified materials are examples of critical infrastructures (CIs) where a robotic security solution can be applied. In this and other related areas, detection and tracking of dynamic objects has become an emerging research field in which solutions are required for the correct development of multidisciplinary applications, such as traffic supervision, autonomous navigation of robots and other vehicles or autonomous surveillance of large facilities . Due to this increasing interest, robotic solutions for security and surveillance are currently being developed by different research laboratories and commercial enterprises such as MoviRobotics or Robowatch. Although the commercial products have experienced significant improvements during the past years,there are still problems not fully resolved in areas, such as robot positioning, or by the detection and tracking of mobile objects. This paper focuses on this latter subject. In conventional security and surveillance applications, automatic systems are capable of detecting movement within a surveillance zone, leaving to the human operator the definition of the risk level [6, 8]. Emerging new applications require autonomous surveillance systems capable of both detecting moving objects simultaneously and tracking their trajectories within large security zones [4]. Different sensors, such as laser systems, visual and infrared cameras or ultrasound systems, can be used to detect dynamic objects within a security perimeter. It is the aim of the present work to develop a series of algorithms capable of handling several detected parameters to enable an autonomous decision made by surveillance robots operating in real scenarios. This requires the implementation of accurate methods of detecting and tracking dynamic objects at long distances.

II. MOVING OBJECT DETECTION

The basic step for analysis of video and tracking of moving object is object detection. For every tracking method it is important to have an object detection mechanism either in every frame or when the object first appears in the video [7]. Object detection mechanism handles separation of moving objects from stationary background [4]. This focuses on higher level processing. It also decreases computation time. This type of object detection mechanism faces various problems due to various environmental conditions like illumination changes shadow etc. Most of the object detection techniques uses information obtained from a single frame. However, some object detection methods make use of the temporal information. Temporal information can be described as information computed from a sequence of frames to reduce the number of false detections. This temporal information is usually in the form of frame differencing. It highlights regions that changes dynamically in consecutive frames. This section reviews three moving object detection methods that are background subtraction, temporal difference and statistical methods [4].

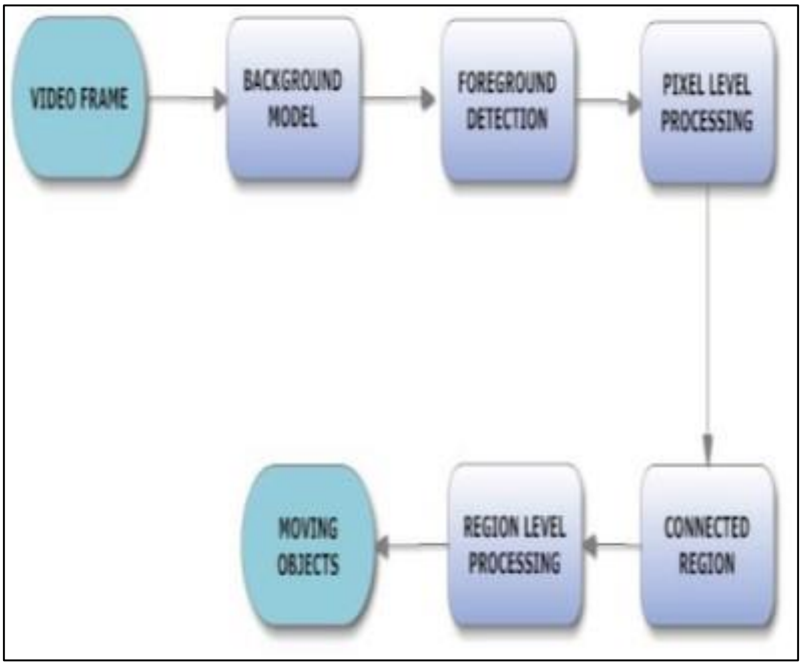

Fig. 1: Framework of Moving Object Detection System

A. Background Subtraction:

The background has to be model first in order to obtain background subtraction. Then the incoming frame is obtained, and subtract out from the background model. A moving object can be detected with reference to background model. This algorithm is called as "Background Subtraction" [8]. The efficiency of a background subtraction technique correlates with three important steps: modelling, noise removal and data validation. Background modelling, is the basic step of the Background Subtraction algorithm. Background model defines the type of model selected to represent the background, and the model representation can simply be a frame at time formula such as the median model. Model Adaption is the procedure used for adjusting the background changes that may occur in a scene. Noise removal is a procedure that eliminates noise in the scene. Data validation is involved with the collection of techniques to reduce the misclassification of pixels. A reliable and robust background subtraction algorithm should handle Sudden or gradual illumination changes, High frequency, repetitive motion in the background and Long-term scene changes [8]. The advantages of background subtraction models are extremely easy implementation and use, pretty fast, corresponding background models are not constant, they Change over time though there are few shortcomings also which includes that accuracy of frame diffrencing depends on object speed and frame rate and Mean and median background models have relatively high memory requirements [3].

B. Statistical Methods:

Statistical Methods are used to overcome the shortcoming of the basic background methods. Statistical methods are used to extract change regions from background. These statistical methods are mainly inspired by the backgroundsubtraction methods. It uses characteristics of individual pixels of group of pixels to construct advance background model. That statistics of background are updated dynamically during processing. At each frame this method keeps and updates dynamic statistics of pixels that belongs to background image process. Foreground pixels are identified by comparing each pixel's statistics with that of the background model. This approach is becoming more popular due to its reliability in scenes that contain noise, illumination changes and shadow. [2] One of the example of statistical methods, Stauffer and Grimson described an adaptive background mixture modeled by a mixture of Gaussians which are updated on-line by incoming image data. In order to detect whether a pixel belongs to a foreground or background process, the Gaussian distributions of the mixture model for that pixel are evaluated. [1, 9]

C. Temporal Differencing:

Temporal differencing technique uses the pixel wise difference between two or three frames in frame sequences to extract a moving object. Temporal differencing technique is computationally simple and fast as well as is adaptive to dynamic environment. In temporal difference technique, extraction of moving pixel is simple and fast. It takes difference of the current and previous frames as shown in equation

$$| I (x, y) - I -1 (x, y) |>t\ i\ i \quad (3)$$

Where 't' is predefined threshold.

This method is computationally less complex and adaptive to dynamic changes in the video frames. In temporal difference technique, extraction 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. Temporal difference require special supportive algorithm to detect stopped objects. [1, 5]

III. OBJECT TRACKING

Object tracking is higher level computer vision problem. It is the important issue in human motion analysis [7]. In its simplest form, tracking can be defined as the problem of estimating the trajectory of an object in the image plane as it moves around a scene. Object tracking refers to the process of matching detected foreground objects between consecutive frames using different features of object like motion, velocity, color, texture. [1] The problem of motion-based object tracking can be divided into two parts: detecting moving objects in each frame, associating the detections corresponding to the same object over time. In order to decide which approach to be used for object tracking one should try to find answer to these questions: Which object representation is suitable for tracking? Which image features should be used? How should the motion, appearance, and shape of the object be modelled? A large number of object tracking methods have been proposed which attempt to answer these questions for a variety of scenarios. [4] Different object tracking methods are described as follows.

A. Point Tracking:

Point tracking is robust, reliable and accurate tracking method developed by Veenman. This method is generally is used for to track the vehicles. This method require deterministic or probabilistic methods. object is tracked is based on point which is represented in detected object in consecutive frames and association of the points is based on the previous object state which can include object position and motion. This approach requires an external mechanism to detect the objects in every frame. In this objects detected in consecutive frames are represented by points, and the current state of points is compared to the previous object state which can include object position and motion. This approach requires an external mechanism to detect the objects in every frame. This method requires deterministic or probabilistic methods Point correspondence methods can be divided into two broad categories, namely, deterministic which uses qualitative motion heuristic to constrain the correspondence problem and statistical or probabilistic methods which explicitly take the object measurement and take uncertainties into account to establish correspondence. [1, 7]

B. Kernel Tracking:

In this approach kernel works according to shape and appearance of the object. In kernel tracking any feature of object can be used to track object like rectangular template or an elliptic shape with an associated histogram. After computing the motion of the kernel between consecutive frames object can be tracked. The computed motion of object is usually represented as parametric transformation such as translation, rotation and affine.[1] These algorithms depends on following terms: appearance representation used, the number of objects tracked, and the method used to estimate the object motion.These methods can be broadly divided into into two subcategories namely, templates and density-based appearance models, and multiview appearance models. [4] An example of kernel tracking method is mean-shift tracking which uses E-kernel is used. It contains histogram feature based by spatial masking with an isotropic kernel.

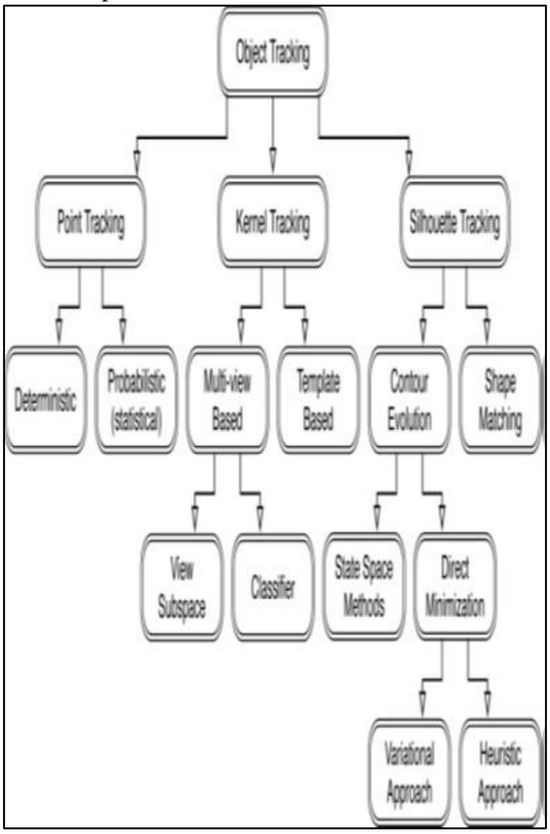

Fig. 2: Taxonomy of tracking methods

C. Silhouette Tracking:

In this approach tracking is performed by Silhouette or object region is extracted from detected object in each frame. This approach uses information encoded object region in each frame. By shape matching or contour evolution silhouettes are tracked either by calculating object region in consecutive frame tracking is done. [4] Both of these methods can essentially be considered as object segmentation applied in the temporal domain using the priors generated from the previous frames. We can broadly divide silhouette trackers into to categories, namely, shape matching which search for the object silhouette in the current frame and contour tracking, on the other hand, evolve an initial contour to its new position in the current frame by either using the state space models or direct minimization of some energy functional. Silhouette based methods provide an accurate shape description for objects that may have complex shapes, for example, hands, head, and shoulders that cannot be well described by simple geometric shapes. Object tracking depends on the features, it also requires selecting the right features, which plays a critical role in tracking. [1] In general, the features used for tracking must be unique so that the objects can be easily distinguished in the feature space. Following various features are used for object tracking:

D. Color

The apparent color of an object is decided by two physical factors, namely spectral power distribution of the illuminant and second is the surface reflectance properties of the object. The RGB (red, green, blue) color space is usually used to represent color in image processing,

E. Edges:

Object boundaries usually generate strong changes in image intensities. Edge detection is used to identify these changes. An important property of edges is that they are less sensitive to illumination changes compared to color features.

F. Centroid:

The Center of mass (centroid) is vector of 1-by-n dimensions in length that specifies the center point of a region. For each point it is worth mentioning that the first element of the centroid is the horizontal coordinate (or x-coordinate) of the center of mass, and the second element is the vertical coordinate (or y-coordinate).

G. Texture:

Texture is used for classification as well as tracking purpose. This feature is used to identify region or object in which we are interested. It is a measurement of the intensity variation of a surface which quantifies properties such as smoothness and regularity [20]. Compared to color, texture requires a processing step to generate the descriptors. Among all features color and texture features are widely used to track the object. Color bands are sensitive to illumination variation.

H. Correspondence Based Matching Algorithm:

Correspondence based object matching algorithm refers to taking the objects of the previous frame and the objects of the current frame, and match the pairs which are close. In this method we compute the distance between the centroid that is smaller than a pre-defined threshold T [7]. For example, suppose two objects (c and p, c for current frame and p for previous frame) with center of mass (xc, yc) and (xp, yp) respectively, then the Euclidian distance between centers expressed as shown in equation < T (4) There are varies number of objects (blobs) in the current and previous frame In and In-1. Let Ln−1 and Ln be the number of objects (blobs) in these frames, respectively. There are three possible cases: Case I: Ln > Ln−1 Case II: Ln < Ln−1 Case III: Ln = Ln−1 Case I: In this case the numbers of objects in the current frame are more than the number of object in the previous frame [4]. In this case we find correspondence of the objects in the current frame that have correspondence with the previous frame rest of the objects in the current frame not tracked. Here, numbers of not tracked objects are (Ln – Ln-1). Case II: In this case the numbers of objects in the current frame are same as the number of object in the previous frame. In this case we find correspondence of the all objects in the current frame with all objects in the previous frame. In this case all objects are tracked. Case III: In this case the numbers of objects in the current frame are less than the number of object in the previous frame. In this case we find correspondence of the all objects in the current frame that have correspondence with the previous frame. [1]

IV. CONCLUSION

In this paper, we have reviewed and presented different methods for moving object detection and Moving object tracking, and also various other factors. In order to analyze different images and extracting higher level information from it, image enrichment, motion detection, object tracking and performance understanding researches have been studied. We have described Three different methods for object detection, namely, background subtraction, temporal differencing, statistical methods. We have also discussed various related issues, advantages and shortcomings of those methods. Temporal differencing method is computationally less complex and adaptive to dynamic changes in the video frames. Though in temporal difference technique, extraction 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. Temporal difference require special supportive algorithm to detect stopped objects. We presented detail of background subtraction method in deep because of its computational effectiveness and accuracy. Though background subtraction technique also has few shortcomings. Statistical method can be used to overcome the shortcomings of background subtraction and temporal differencing techniques. In this paper we have discussed three techniques for object tracking, namely, point tracking, kernel tracking and contour tracking according to the representation method of a target object. In point tracking approach, is tracked is based on point which is represented in detected object in consecutive frames and association of the points is based on the previous object state which can include object position and motion. In kernel tracking approach, any feature of object can be used to track object like rectangular template or an elliptic shape with an associated histogram. After computing the motion of the kernel between consecutive frames object can be tracked. Now a day, the most preferred and popular kernel tracking techniques are based on Mean-shift tracking and particle filter. In silhouette contour tracking, tracking is performed by Silhouette or object region is extracted from detected object in each frame. This approach uses information encoded object region in each frame.

REFERENCES

[1] Yilmaz, A., Javed, O., and Shah, M. 2006. Object tracking: A survey. ACM Computing Survey. 38, 4, Article 13, (Dec. 2006), 45 pages. DOI = 10.1145/1177352.1177355 http://doi.acm.org/10.1145/1177352.1177355
[2] Safvan Vahora, Narendra Chauhan and Nilesh Prajapati, "A Robust Method for Moving Object Detection Using Modified Statistical Mean Method", International Journal of Advanced Information Technology (IJAIT) Vol. 2, No.1, February 2012.
[3] Birgi Tamersoy, "Background Subtraction", http://www.cs.utexas.edu/~grauman/courses/fall2009/slides/lecture9_background.pdf
[4] Kinjal A Joshi, Darshak G. Thakore, "A Survey on Moving Object Detection and Tracking in Video Surveillance System", International Journal of Soft Computing and Engineering (IJSCE) ISSN: 2231-2307, Volume-2, Issue-3, July 2012.
[5] Raquib Buksh, Soumyajit Routh, Parthib Mitra, Subhajit Banik, Abhishek Mallik, Sauvik Das Gupta, "Implementationof MATLAB based objectdetection technique on Arduino Board and iROBOTCREATE, International Journal of Scientific and Research Publications, Volume 4, Issue 1, January 2014 1 ISSN 2250-3153.
[6] Ms Jyoti J. Jadhav, "Moving Object Detection and Tracking for Video Survelliance", International Journal of Engineering Research and General Science Volume 2, Issue 4, June-July, 2014, ISSN 2091-2730.
[7] Yiwei Wang, Robert E. Van Dyck, and John F. Doherty, "Tracking Moving Objects In Video Sequences", Department of Electrical Engineering, The Pennsylvania State University, University Park, PA16802.
[8] Vishwadeep Uttamrao Landge, "Object Detection and Object Tracking Using Background Subtraction for Surveillance Application", International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering, (An ISO 3297: 2007 Certified Organization) Vol. 3, Issue 7, July 2014.
[9] William M. Wells III, "Statistical Object Recognition", Submitted to the Department of Electrical Engineering and Computer Science on November 24, 1992, inpartial fulfillment of the requirements for the degree of Doctor of Philosophy.

: