International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor : 6.887
Volume 6 Issue II, February 2018- Available at www.ijraset.com

Multiple Object Detection and Tracking: A Survey

Dhenuka M. Patel[1], Udesang K. Jaliya[2], Hemant D. Vasava[3]
[1]*P.G.Student, Department of Computer Engineering, B.V.M. Engineering College, V. V. Nagar, India*
[2,3]*Assistant Professor, Department of Computer Engineering, B.V.M. Engineering College, V. V. Nagar, India*

Abstract*: Multiple Object Tracking is the process of locating multiple objects over time in a video stream. Object detection and classification are two prior steps before performing tracking over video scene. Object detection is the process of locating an object of interest in a single frame. So, in other words we can say that multiple object tracking is the process of associating detected objects in consecutive video frames. The detected objects may belong to various categories such as vehicles, humans, swaying trees or other moving objects. So, object classification is the process to classify these objects using different approaches. However, some object tracking applications may not need to classify detected objects. In this paper, we had discussed various object detection and tracking methods, which are available in the literature.

Keywords: Object Detection, Object Tracking, Video Surveillance, Background Modelling, Motion Estimation.

I. INTRODUCTION

Object tracking is a very important aspect of artificial intelligence, digital image processing and computer vision [2] [4]. Object tracking is a core of several computer vision applications such as traffic monitoring, video surveillance, video animation, robot vision and many more [1]. In recent years, a number of single-object tracking algorithms were introduced and implemented robustly, but in case of multiple object tracking scenario is quite different and challenging too because in case of multiple object tracking, targets needs to be matched from frame to frame in video sequence [5]. To deal with this problem, researchers did lot of works to provide better solutions to such applications based on multiple object tracking.

The basic flow diagram for multiple object tracking is as following.

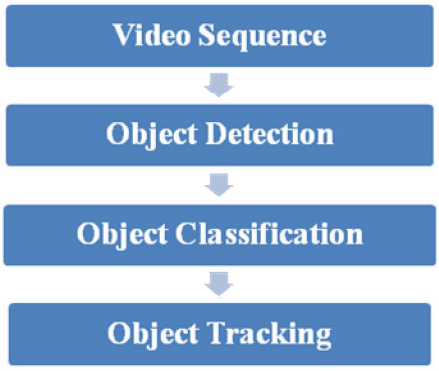

Fig 1: basic flow diagram of multiple object tracking [1]

This paper is structured in this way: section 1 gives introduction to multiple object tracking. Section 2 provides literature discussion about work related to multiple object tracking. In section 3 we discussed about object detection methods from video. Section 4 briefly describes object tracking methods. At last, section 5 concludes this paper.

II. RELATED WORK

In literature [5], feature-based algorithm using Kalman filter motion to track multiple objects is proposed by authors. They had used background subtraction method to detect and extract moving object. Algorithm is validated on human as well as vehicle image sequence and also under confusing situation; it achieves efficient tracking of objects. Literature [6] focuses on tracking players in football match. Authors have used modified Hungarian algorithm and Kalman filter. As the result concludes that for tracking multiple objects in football match, the linking process is achieved by modified Hungarian algorithm and motion model building and prediction is achieved by Kalman filter successfully. In literature [7], parallel Kalman filter is used for moving object detection and tracking. The precision and recall value proves that, proposed method is effective for detecting and tracking multiple objects. In literature [8], authors have proposed unique method named as Dual Layer Particle Filtering (DLPF), which simultaneously detect and track multiple target objects. It uses parent-particles (PP) in first layer to detect multiple objects and child-particles (CP) in

second layer to track that detected object. In literature [9], the proposed method is based on particle filter, which overcomes some challenges of object tracking. It can track the target object robustly when target object is occluded by the background object or other object. Authors had written that, particle filter has higher flexibility than Kalman filter, also they modifies particle filter to overcome different challenges. In literature [10], proposed method is based on multiple hypothesis tracking. The method is for generic object tracking, which means that there are no priori restrictions in type of objects that can be tracked. However, authors confesses that, proposed method is quite complex too. In literature [11], adaptive template matching algorithm is used for tracking human upper body. The proposed method is fast and robust, as they had added only head edge detection for better tracking. But also they assume that person's upper body and face are visible without any occlusion. In literature [12], authors had written that, in some complex situation like target object have scale changes or similar color with background, traditional meanshift algorithm cannot obtain accurate results. So they suggest new meanshift target tracking algorithm, named as DEPTH & SIFT-Meanshift algorithm. This algorithm is proposed by using a depth camera and SIFT (Scale Invariant Feature Transform) feature metric. The experimental result shows that, proposed method has abilities to overcome described challenges. Literature [13] proposes a robust approach for tracking arbitrary objects. Authors propose a new motion model based on Kernelized Harmonic Means and particle filter, and they introduce their proposed model within a SVM framework.

III.OBJECT DETECTION METHODS

Object detection from the video scene is the very first step in the process of object tracking. It is the technique of computer vision which deals with the identifying instances of objects such as vehicle, human, bird, animal or any other moving object [2]. Few approaches for object detection are described below.

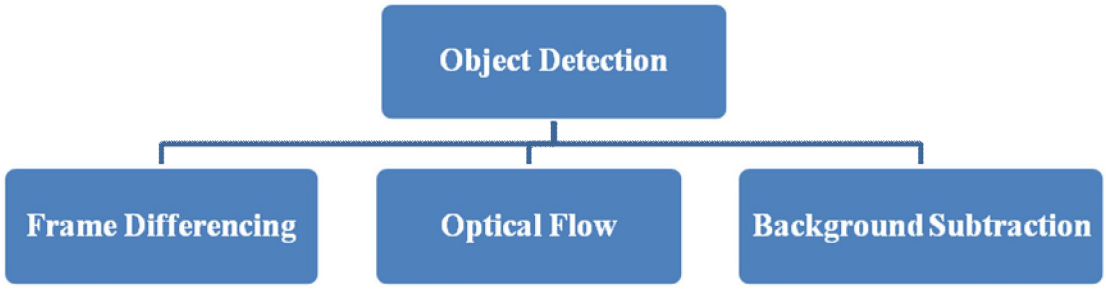

Fig-2: Types of object detection [1]

A. Frame Differencing

Frame differencing method detects the objects by calculating difference of two consecutive frames of a video sequence [1]. The calculation of this method is simple and implementation is easy. It is highly adaptive and also works well for dynamic environments. However for certain types of objects, it does poor extraction. Also there may appear small holes in detected object [2]. In this method, the current frame is simply subtracted from the previous frame and if the difference in the pixel values for a given pixel is greater than a threshold T_s, the pixel is considered as a part of foreground [4].

$$| frame_i - frame_{i-1} | > T_s \quad \text{..... (1)}$$

B. Optical Flow

Optical flow is a two dimensional vector-based method, which estimates velocities and direction of each point of an image sequence [2] [3]. It estimates the motion of object in video by matching points on objects over multiple frames [2]. It is more suitable for multiobjective moving analysis in complex scenes, because of its higher detection accuracy. However, it is less suitable when there is occlusion [3].

C. Background Subtraction

This is the technique, where each video sequence is compared with background. In this method, background modelling is the first step. Mean filter and median filter are the techniques used to build background [1]. In this method the difference between current image and background image is being calculated, and if it is above a threshold value then the pixel is considered as foreground and others are as background [4].

$$B(x, y, t) = \text{median } \{F(x, y, t-i)\} \quad \text{..... (2)}$$

$$|F(x, y, t) - \text{median}\{F(x, y, t- i)\}| > T_s \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad (3)$$

Where, $i \in \{0, 1\ldots n-1\}$

This is the best suitable method for object detection, in the case where background is known. However, it is very sensitive to changes in the external environment [1]. This method has two approaches.

1) *Recursive Algorithm*: Recursive algorithm does not store all the previous frames into buffer for background modelling; hence there is no need of buffer and require less storage. On the basis of each input frame, it recursively updates background model. The main advantage of this method is, the distant past input frames does not affect the current background model. Various methods such as approximate median, Gaussian of mixture, adaptive background are used by this algorithm [1] [4]
2) *Non- Recursive Algorithm:* This method is highly adaptive. This algorithm uses sliding- window method for background estimation. It stores all the previous frames into buffer and on the basis of variation of each pixel within the buffer, it estimates the background. The significant storage requirement is the limitation of this algorithm [1] [4].

IV. OBJECT TRACKING METHODS

Object tracking means to follow a particular object in the video scene which is moving. Once the objects are detected, the next step in the process of object tracking is to track that objects from one frame to another using tracker. Sometimes, due to illumination change, occlusion, complex object shape- tracking method becomes more complex and challenging [2]. According to literature [1[10]], object tracking methods are mainly classified into three categories: point tracking, kernel tracking and silhouette tracking, are briefly explained below.

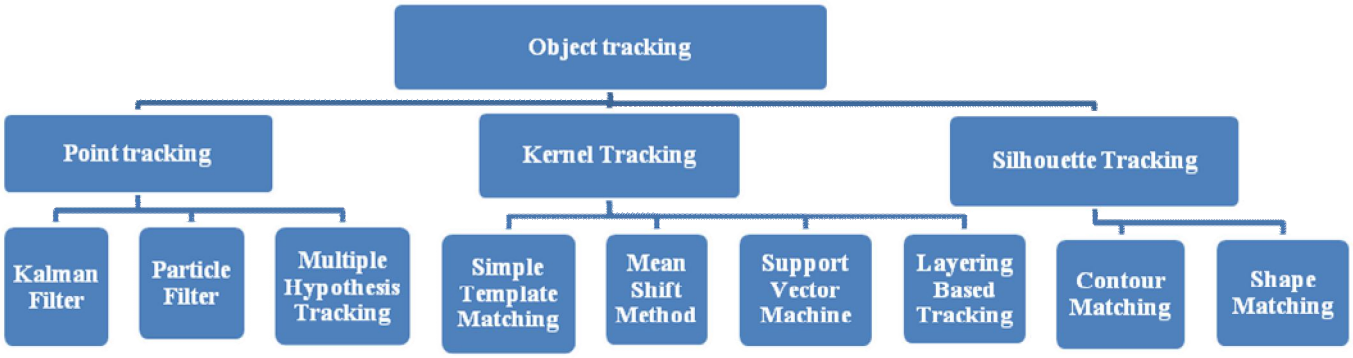

Fig-3: Types of object tracking [1]

A. Point Tracking Methods

At the time of object tracking, all the detected moving objects are represented by their feature points. Using these feature points, point trackers detects and tracks perticular object in every frame. The identification of these points are based on thresholding [1]. Point tracking methods are reliable and robust [4].

1) *Kalman Filter*: Mathematically, Kalman filter is an estimator which is used to predict and correct the states of wide range of linear processes [5]. Kalman filter is often used in computer vision applications which require tracking of moving objects. It predicts the velocity and motion gradient of an object [7]. Generally Kalman filter works in two basic equations including time update and measurement update. Time update equations predict the future state using Gaussian Probabilistic approach and after that, measurement update generates correction parameter for next prediction. Also measurement update state generates Kalman Gain, which is used to determine how much prediction influence to the measurement in the basic equation [6]. In paper [7] authors have used parallel Kalman filter which is modified version of Kalman filter. It can also accurately determine the motion of detected object better than normal Kalman filter. The multiple object tracking is a challenging and harder while the objects have a similar appearance, still Kalman filter based tracker does good job for the same [5]

2) *Particle Filter:* The particle filter is used to estimate the posterior probability distribution by many particles. The basic steps of particle filter are: 1) prediction, 2) weight calculation and 3) filtering. At the very first step, the particles are transited by the state transition function. At the second step, the weight of each particle is calculated by normalizing likelihood. It is obtained by the likelihood function in the weight calculation process. After weight calculation, the particles are filtered by the weighted resampling with replacement. Following these steps, the particle filter estimates the state variables by its weighted average [9].
3) *Kalman filter Vs Particle filter:* Kalman filter works only for linear dynamical system and it is optimal estimator when there is Gaussian noise processes. On the other hand, particle filters works for linear as well as non-linear dynamical systems and does not restrict to Gaussian noise [8]. Paper [14] compares these two position estimators. Author concludes that a particle filter may more reliable for multiple objects tracking in complex situation, and a Kalman filter should be chosen for simpler tasks, like surveillance applications having low populated areas.
4) *Multiple Hypothesis Tracking (MHT):* Multiple hypothesis tracking is an iterative algorithm. MHT can track multiple objects and is capable of handling occlusion and calculating optimal solution [1]. In literature [10], authors have performed coarse supervised segmentation to remove non-object regions such as road, buildings or vegetation. Furthermore they perform unsupervised multi-scale segmentation to extract scale-stable objects. It is possible that correct object may not be determined on a single frame basis, as it may overlap by given proposals. So authors perform multi-Hypothesis tracking at the level of object proposals.

B. Kernel Tracking Methods

Kernel tracking method tracks the object based on its shape and appearance. Basically object contains various features to represent its embryonic region, and one of those features is used to track object as kernel. The parametric motion of kernel such as translation, affine etc. is calculated, in order to track the object [1] [4].

1) *Simple Template Matching Method*: Template based matching is a simple and fast method. Normally, foreground segmentation yields in foundation of template based matching. It determines best location of object for a given model; by sliding the patch over the input frame through normalized cross correlation, and by picking the point with maximal value as the best fit position. Template matching method is efficient enough to handle frequent change in object orientation and pose [11]. Template matching can be either implemented as static or adaptive. In literature [11] authors have used adaptive template matching algorithm, which means that the model of the template are extracted from the previous frame.
2) *Mean Shift Method:* Mean shift is an adaptive and iterative algorithm. It can be applied to image segmentation, nonparametric density estimation and video tracking. First of all, it searches maximum Bhattacharya coefficient, using spatial kernel weighted color histogram as a template. Using that coefficient, algorithm achieves target tracking and positioning. However, due to instability of color characteristics and lightning changes its tracking results are unstable and scope of its practical application is also limited [12]. Additionally, when there is scale change in object occurs, tracking window remains same. Thus targets are easily lost while moving forward or backward [11]. Literature [12] proposed DEPTH & SIFT-meanshift algorithm, which overcomes these two limitations of traditional mean shift algorithm.
3) *Support Vector Machine:* Support vector machines are basically used for classification problem. In object tracking problems SVM generalized well using kernels, and it provides robust solution against noise. The main task in this method is to take set of training example as input and learn a classification function [13]. Here, classification function is used to make binary labels ±1, which means that positive samples contains interested area, i.e. objects that are being tracked and negative samples contains all the remaining things that should not be tracked [13] [1].
4) *Layering based Tracking:* In the scenario where the object is fully occluded by another, this technique is capable of tracking object. In this method, layer consists of motion such as translation and rotation, shape representation, and intensity based layer appearance. First of all object's motion is estimated from the rewarded image and by compensating background motion, layering is achieved. After that on the basis of object's foregoing motion and shape features, pixel's probability is calculated [1].

C. Silhouette Tracking Methods

Sometimes, simple geometrical shapes cannot properly describe objects with complex shapes such as hand, fingers, shoulder etc. In such a scenario, silhouette tracking method is used to accurately describe the objects. In this method, object model generated by previous frames, is used by present frame to find the object region. These methods works well to handle occlusion, split & merge and also provides flexibility to handle variety of object shapes [1].

1) *Contour Matching*: Contour based tracking method represents moving object's outline as bounding counters. It dynamically keeps updating those bounding counters in order to achieve object tracking. Due to occlusion if some object losses its outline, then this method works well in case of recovering object's shape back [2].
2) *Shape Matching:* The performance of shape matching technique is quite similar to the template based technique. The main task in this method is to find matching silhouette in two successive frames. For that, background subtraction method is used to detect silhouette. This method deals with single object only [1].

V. CONCLUSION

Two important phases of multiple object tracking process are briefly described in this paper. Methods used for multiple moving object detections such as background subtraction, optical flow and frame difference are explained briefly. Among these three methods, background subtraction provides complete information about object compared to frame difference and optical flow method. Also this paper briefly describes multiple object tracking methods such as point tracking, kernel tracking and silhouette tracking. The various sub-approaches of these tracking methods are also explained briefly. Point tracking methods involve detection of object in every frame; while kernel based tracking and silhouette based tracking require detection only when object appears first time in video scene

VI. ACKNOWLEDGMENT

I would like to thank my guide Dr. Udesang K. Jaliya and Prof. Hemant D. Vasava for their advice and continued support.

REFERENCES

[1] Himani S. Parekh, Darshak G. Thakore, Udesang K. Jaliya "A Survey on Object Detection and Tracking Methods", 2014 IJIRCCE.
[2] Shipra Ojha, Sachin Sakhare, "Image Processing Techniques for Object Tracking in Video Surveillance- A Survey", ©2015 IEEE.
[3] Sepehr Aslani, Homayoun Mahdavi-Nasab, "Optical Flow Based Moving Object Detection and Tracking for Traffic Surveillance", IJECE 2013.
[4] Pawan Kumar Mishra, G. P. Saroha, "A Study on Video Surveillance System for Object Detection and Tracking", ©2016 IEEE.
[5] Xin Li, Kejun Wang,Wei Wang and Yang Li, "A Multiple Object Tracking Method Using Kalman Filter", ©2010 IEEE.
[6] (Bima Sahbani, Widyawardana Adiprawita, "Kalman Filter and Iterative-Hungarian Algorithm Implementation for Low Complexity Point Tracking as Part of Fast Multiple Object Tracking System", ©2016 IEEE.
[7] Nilesh J. Uke (IEEE Member), Pravin R. Futane, "Efficient Method for Detecting and Tracking Moving Objects in Video", ©2016 IEEE.
[8] Kyungwon Jung, Nahyun Kim, Seungwon Lee and Joonki Paik, "Dual-layer particle filtering for simultaneous multiple objects detection and tracking", ©2013 IEEE.
[9] "Shinji Fukui, Sou Hayakawa, Yuji Iwahori, Tsuyoshi Nakamura, M. K. Bhuyan, "Particle Filter Based Tracking with Image-Based Localization", 2016 ELSEVIER.
[10] Aljosa Osep, Alexander Hermans, Francis Engelmann, Dirk Klostermann, Markus Mathias and Bastian Leibe, "Multi-Scale Object Candidates for Generic Object Tracking in Street Senes", ©2016 IEEE.
[11] Shusheng He, Alei Liang, Ling Lin, Tao Song, "A Continuously Adaptive Template Matching Algorithm for Human Tracking", © 2017 IEEE.
[12] Lu Lu, Minrui Fei, Haikuan Wang, Huosheng Hu, "A New Meanshift Target Tracking Algorithm by Combining Feature Points from Gray and Depth Images", © Springer.
[13] Kumara Ratnayake, Maria A. Amer, "Object Tracking with Adaptive Motion Modeling of Particle Filter and Support Vector Machines", ©2015 IEEE.
[14] M. Marron, J.C. Garcia, M.A. Sotelo, M. Cabello, D. Pizarro, F. Huerta, J. Cerro, "Comparing a Kalman Filter and a Particle Filter in a Multiple Objects Tracking Application", ©2007 IEEE.