A Survey on Object Detection and Tracking Methods

Himani S. Parekh[1], Darshak G. Thakore[2], Udesang K. Jaliya[3]

P.G.Student, Department of Computer Engineering, B.V.M. Engineering College, V. V. Nagar, India[1]

Associate Professor, Department of Computer Engineering, B.V.M. Engineering College, V. V. Nagar, India[2]

Assistant Professor, Department of Information Technology, B.V.M. Engineering College, V. V. Nagar, India[3]

ABSTRACT: The goal of object tracking is segmenting a region of interest from a video scene and keeping track of its motion, positioning and occlusion.The object detection and object classification are preceding steps for tracking an object in sequence of images. Object detection is performed to check existence of objects in video and to precisely locate that object. Then detected object can be classified in various categories such as humans, vehicles, birds, floating clouds, swaying tree and other moving objects. Object tracking is performed using monitoring objects' spatial and temporal changes during a video sequence, including its presence, position, size, shape, etc.Object tracking is used in several applications such as video surveillance, robot vision, traffic monitoring, Video inpainting and Animation. This paper presents a brief survey of different object detection, object classification and object tracking algorithms available in the literature including analysis and comparative study of different techniques used for various stages of tracking.

Keywords: Object Detection, Object Tracking, Object Classification, Video Surveillance, Background Modelling

I. INTRODUCTION

Videos are actually sequences of images, each of which called a frame, displayed in fast enough frequency so that human eyes can percept the continuity of its content. It is obvious that all image processing techniques can be applied to individual frames. Besides, the contents of two consecutive frames are usually closely related [1].

The identification of regions of interest is typically the first step in many computer vision applications including event detection, video surveillance, and robotics. A general object detection algorithm may be desirable, but it is extremely difficult to properly handle unknown objects or objects with significant variations in color, shape and texture. Therefore, many practical computer vision systems assume a fixed camera environment, which makes the object detection process much more straightforward [3]. An image, usually from a video sequence, is divided into two complimentary sets of pixels. The first set contains the pixels which correspond to foreground objects while the second and complimentary set contains the background pixels. This output or result is often represented as a binary image or as a mask. It is difficult to specify an absolute standard with respect to what should be identified as foreground and what should be marked as background because this definition is somewhat application specific. Generally, foreground objects are moving objects like people, boats and cars and everything else is background [11]. Many a times shadow is classified as foreground object which gives improper output.

Following are the basic steps for tacking an object, as describe in many literature.

1) *Object Detection*

Object Detection is to identify objects of interest in the video sequence and to cluster pixels of these objects. Object detection can be done by various techniques such as frame differencing, Optical flow and Background subtraction**.**

2) *Object Classification*

Object can be classified as vehicles, birds, floating clouds, swaying tree and other moving objects. The approaches to classify the objects are Shape-based classification, Motion-based classification, Color based classification and texture based classification.

3) Object Tracking

Tracking can be defined as the problem of approximating the path of an object in the image plane as it moves around a scene. The approaches to track the objects are point tracking, kernel tracking and silhouette.
Following are some of the challenges that should be taken care in object tracking as described in [10]:

1. Loss of evidence caused by estimate of the 3D realm on a 2D image,
2. Noise in an image,
3. Difficult object motion,
4. Imperfect and entire object occlusions,
5. Complex objects structures.

This paper is structured in the following way: Section 1 gives introduction to object tracking. Section 2 provides the related work. Section 3 deals with brief explanation on several object detection methods. Section 4 consists of detailed study on object classification methods and Section 5 describes object tracking methods. Section 6 provides conclusions.
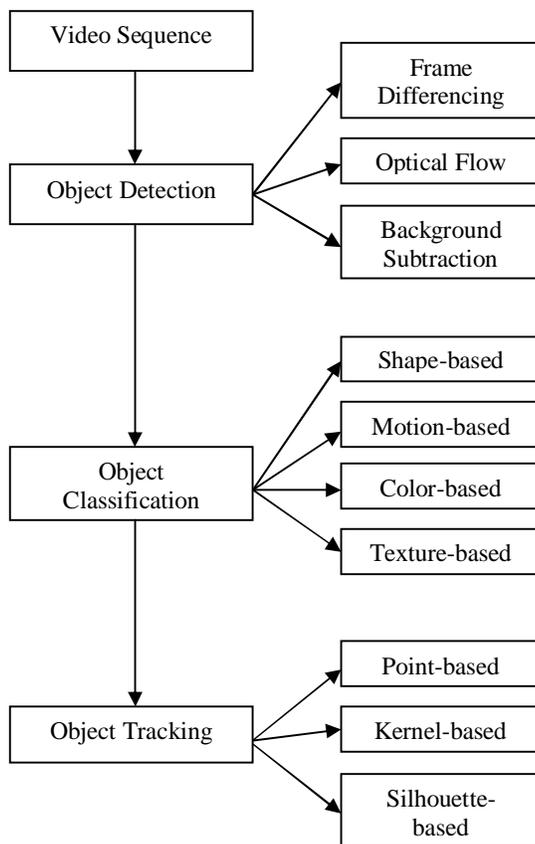

Fig 1: Basic steps for tracking an object [8]

II. RELATED WORK

In paper [4] multiple human object tracking approach is used which based on motion estimation and detection, background subtraction, shadow removal and occlusion detection. Video sequences have been captured in the laboratory and tested with the proposed algorithm. The algorithm works efficiently in the event of occlusion in the video sequences. In paper [5] A tracking algorithm based on adaptive background subtraction about the video detecting

and tracking moving objects is presented in this paper. Firstly, median filter is used to achieve the background image of the video and denoise the sequence of video. Then adaptive background subtraction algorithm is used to detect and track the moving objects. The simulation results by MATLAB show that the adaptive background subtraction is useful in both detecting and tracking moving objects, and background subtraction algorithm runs more quickly. Paper [6] attempts to find moving objects by subtracting the background images from static single camera video sequences in security systems. It aims to improve the background subtraction techniques for indoor video surveillance applications. The novel automatic threshold updating (ATU) algorithm is also developed and tested for various indoor video sequences which give better efficiency. The statistical and temporal differencing methods are also presented. Finally, novel approach is compared with the existing methods. Paper [7] presents a new algorithm for detecting moving objects from a static background scene to detect moving object based on background subtraction. Reliable background updating model is set up based on statistical. After that, morphological filtering is initiated to remove the noise and solve the background interruption difficulty. At last, contour projection analysis is combined with the shape analysis to remove the effect of shadow; the moving human bodies are accurately and reliably detected. The experiment results show that the proposed method runs rapidly, exactly and fits for the concurrent detection.

III. OBJECT DETECTION METHODS

First step in the process of object tracking is to identify objects of interest in the video sequence and to cluster pixels of these objects. Since moving objects are typically the primary source of information, most methods focus on the detection of such objects. Detailed explanation for various methods is given below.

A. Frame differencing

The presence of moving objects is determined by calculating the difference between two consecutive images. Its calculation is simple and easy to implement. For a variety of dynamic environments, it has a strong adaptability, but it is generally difficult to obtain complete outline of moving object, responsible to appear the empty phenomenon, as a result the detection of moving object is not accurate [7].

B. Optical Flow

Optical flow method [1] is to calculate the image optical flow field, and do clustering processing according to the optical flow distribution characteristics of image. This method can get the complete movement information and detect the moving object from the background better, however, a large quantity of calculation, sensitivity to noise, poor anti-noise performance, make it not suitable for real-time demanding occasions.

C. Background subtraction

First step for background subtraction is background modelling. It is the core of background subtraction algorithm. Background Modeling must sensitive enough to recognize moving objects [10]. Background Modeling is to yield reference model. This reference model is used in background subtraction in which each video sequence is compared against the reference model to determine possible Variation. The variations between current video frames to that of the reference frame in terms of pixels signify existence of moving objects [10]. Currently, mean filter and median filter [2] are widely used to realize background modeling. The background subtraction method is to use the difference method of the current image and background image to detect moving objects, with simple algorithm, but very sensitive to the changes in the external environment and has poor anti- interference ability. However, it can provide the most complete object information in the case background is known. As describe in [11], background subtraction has mainly two approaches:

1. Recursive algorithm

Recursive techniques [11] [6] do not maintain a buffer for background estimation. Instead, they recursively update a single background model based on each input frame. As a result, input frames from distant past could have an effect

on the current background model. Compared with non-recursive techniques, recursive techniques require less storage, but any error in the background model can linger for a much longer period of time. This technique includes various methods such as approximate median, adaptive background, Gaussian of mixture

2. *Non-Recursive Algorithm*

A non-recursive technique [6] [11] uses a sliding-window approach for background estimation. It stores a buffer of the previous L video frames, and estimates the background image based on the temporal variation of each pixel within the buffer. Non-recursive techniques are highly adaptive as they do not depend on the history beyond those frames stored in the buffer. On the other hand, the storage requirement can be significant if a large buffer is needed to cope with slow-moving traffic.

TABLE 1
COMPARATIVE STUDY OF OBJECT DETECTION METHODS [8]

Methods		Accuracy	Computational Time	Comments	
Background Subtraction	Gaussian Of Mixture	Moderate	Moderate	+	Low memory requirement
				-	It does not cope with multimodal background
	Approximate Median	Low to Moderate	Moderate	+	It does not require sub sampling of frames for creating an adequate background model.
				-	It computation requires a buffer with the recent pixel values
Optical Flow		Moderate	High	+	It can produce the complete movement information
				-	Require Large amount of calculation
Frame Differencing		High	Low to Moderate	+	Easiest Method. Perform well for static background.
				-	It requires a background without moving objects

IV. OBJECT CLASSIFICATION METHODS

The extracted moving region may be different objects such as humans, vehicles, birds, floating clouds, swaying tree and other moving objects. Hence we use the shape features of motion regions [7]. As per literatures, approaches to classify the objects are as follows:

A. *Shape-based classification:*
Different descriptions of shape information of motion regions such as representations of points, box and blob are available for classifying moving objects. Input features to the network is mixture of image-based and scene-based object parameters such as image blob area, apparent aspect ratio of blob bounding box and camera zoom. Classification is performed on each blob at every frame and results are kept in histogram [14].

B. *Motion-based classification:*
Non-rigid articulated object motion shows a periodic property, so this has been used as a strong cue for moving object classification. Optical flow is also very useful for object classification. Residual flow can be used to analyze rigidity and periodicity of moving entities. It is expected that rigid objects would present little residual flow where as a non rigid moving object such as human being had higher average residual flow and even displayed a periodic component [14].

C. Color-based classification

Unlike many other image features (e.g. shape) color is relatively constant under viewpoint changes and it is easy to be acquired. Although color is not always appropriate as the sole means of detecting and tracking objects, but the low computational cost of the algorithms proposed makes color a desirable feature to exploit when appropriate. To detect and track vehicles or pedestrians in real-time color histogram based technique is used. According to [2] a Gaussian Mixture Model is created to describe the color distribution within the sequence of images and to segment the image into background and objects. Object occlusion was handled using an occlusion buffer.

D. Texture-based classification

Texture based technique [8] counts the occurrences of gradient orientation in localized portions of an image, is computed on a dense grid of uniformly spaced cells and uses overlapping local contrast normalization for improved accuracy.

According to paper [8], table 2 describes comparative study of classification methods using accuracy and computational time. Advantages and limitations of various techniques are also described in table 2.

TABLE 2
COMPARATIVE STUDY OF OBJECT CLASSIFICATION METHODS [8]

Methods	Accuracy	Computational Time	Comments
Shape-Based	Moderate	Low	Simple pattern-matching approach can be applied with appropriate templates. It does not work well in dynamic situations and is unable to determine internal movements well.
Motion-based	Moderate	High	Does not require predefined pattern templates but struggles to identify a non-moving human.
Texture-based	High	High	Provides improved quality with the expense of additional computation time.
Color-based	High	High	It creates a Gaussian Mixture Model to describe the color distribution within the sequence of images and to segment the image into background and objects

V. OBJECT TRACKING METHODS

Tracking can be defined as the problem of approximating the path of an object in the image plane as it moves around a scene. The purpose of an object tracking is to generate the route for an object above time by finding its position in every single frame of the video [5]. Object is tracked for object extraction, object recognition and tracking, and decisions about activities. According to paper [10], Object tracking can be classified as point tracking, kernel based tracking and silhouette based tracking. For illustration, the point trackers involve detection in every frame; while geometric area or kernel based tracking or contours-based tracking require detection only when the object first appears in the scene. As described in [10], tracking methods can be divided into following categories:
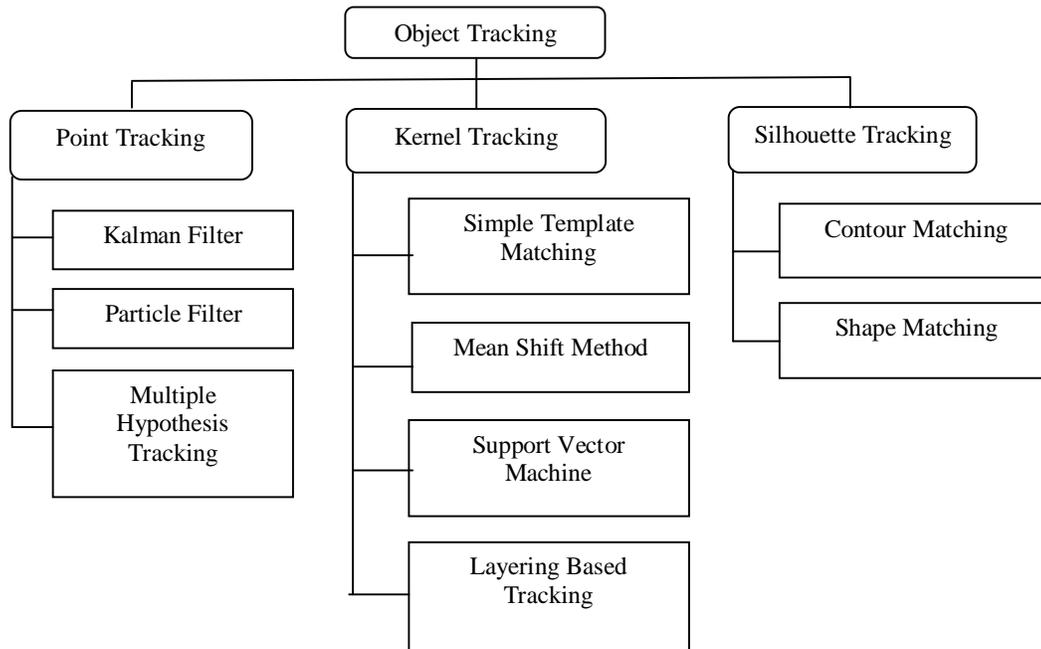
Fig 2: Types of Object Tracking Methods [10]

A. *Point Tracking*

In an image structure, moving objects are represented by their feature points during tracking. Point tracking [10] is a complex problem particularly in the incidence of occlusions, false detections of object. Recognition can be done relatively simple, by thresholding, at of identification of these points.

1. *Kalman Filter*

They are based on Optimal Recursive Data Processing Algorithm. The Kalman Filter performs the restrictive probability density propagation. Kalman filter [12] is a set of mathematical equations that provides an efficient computational (recursive) means to estimate the state of a process in several aspects: it supports estimations of past, present, and even future states, and it can do the same even when the precise nature of the modelled system is unknown. The Kalman filter estimates a process by using a form of feedback control. The filter estimates the process state at some time and then obtains feedback in the form of noisy measurements. The equations for Kalman filters fall in two groups: time update equations and measurement update equations. The time update equations are responsible for projecting forward (in time) the current state and error covariance estimates to obtain the priori estimate for the next time step. The measurement update equations are responsible for the feedback. Kalman filters always give optimal solutions.

2. *Particle Filtering*

The particle filtering [10] generates all the models for one variable before moving to the next variable. Algorithm has an advantage when variables are generated dynamically and there can be unboundedly numerous variables. It also allows for new operation of resampling. One restriction of the Kalman filter is the assumption of state variables are normally distributed (Gaussian). Thus, the Kalman filter is poor approximations of state variables which do not Gaussian distribution. This restriction can be overwhelmed by using particle filtering.

This algorithm usually uses contours, color features, or texture mapping. The particle filter [10] is a Bayesian sequential importance Sample technique, which recursively approaches the later distribution using a finite set of weighted trials. It also consists of fundamentally two phases: prediction and update as same as Kalman Filtering. It was

developing area in the field of computer vision communal and applied to tracking problematic and is also known as the Condensation algorithm

3. Multiple Hypothesis Tracking (MHT):

In MHT algorithm [10], several frames have been observed for better tracking outcomes MHT is an iterative algorithm. Iteration begins with a set of existing track hypotheses. Each hypothesis is a crew of disconnect tracks. For each hypothesis, a prediction of object's position in the succeeding frame is made. The predictions are then compared by calculating a distance measure. MHT is capable of tracking multiple object, handles occlusions and Calculating of Optimal solutions.

B. Kernel Based Tracking

Kernel tracking [9] is usually performed by computing the moving object, which is represented by a embryonic object region, from one frame to the next. The object motion is usually in the form of parametric motion such as translation, conformal, affine, etc.

These algorithms diverge in terms of the presence representation used, the number of objects tracked, and the method used for approximation the object motion. In real-time, illustration of object using geometric shape is common. But one of the restrictions is that parts of the objects may be left outside of the defined shape while portions of the background may exist inside. This can be detected in rigid and non-rigid objects .They are large tracking techniques based on representation of object, object features ,appearance and shape of the object.

1. Simple Template Matching

Template matching [9][4] is a brute force method of examining the Region of Interest in the video. In template matching, a reference image is verified with the frame that is separated from the video. Tracking can be done for single object in the video and overlapping of object is done partially. Template Matching is a technique for processing digital images to find small parts of an image that matches, or equivalent model with an image (template) in each frame. The matching procedure contains the image template for all possible positions in the source image and calculates a numerical index that specifies how well the model fits the picture that position. It can capable of dealing with tracking single image and partial occlusion of object.

2. Mean Shift Method

Mean-shift tracking tries to find the area of a video frame that is locally most similar to a previously initialized model. The image region to be tracked is represented by a histogram. A gradient ascent procedure is used to move the tracker to the location that maximizes a similarity score between the model and the current image region. In object tracking algorithms target representation is mainly rectangular or elliptical region. It contain target model and target candidate. To characterize the target color histogram is chosen. Target model is generally represented by its probability density function (pdf). Target model is regularized by spatial masking with an asymmetric kernel.

3. Support Vector Machine (SVM)

SVM [13] is a broad classification method which gives a set of positive and negative training values. For SVM, the positive samples contain tracked image object, and the negative samples consist of all remaining things that are not tracked. It can handle single image, partial occlusion of object but necessity of a physical initialization and necessity of training.

4. Layering based tracking

This is another method of kernel based tracking where multiple objects are tracked. Each layer consists of shape representation (ellipse), motion such as translation and rotation, and layer appearance, based on intensity. Layering is achieved by first compensating the background motion such that the object's motion can be estimated from the rewarded image by means of 2D parametric motion. Every pixel's probability of calculated based on the object's foregoing motion and shape features [13]. It can capable of tracking multiple images and fully occlusion of object.

C. Silhouette Based Tracking Approach

Some object will have complex shape such as hand, fingers, shoulders that cannot be well defined by simple geometric shapes. Silhouette based methods [9] afford an accurate shape description for the objects. The aim of a silhouette-based object tracking is to find the object region in every frame by means of an object model generated by the previous frames. Capable of dealing with variety of object shapes, Occlusion and object split and merge.

1. Contour Tracking

Contour tracking methods [9], iteratively progress a primary contour in the previous frame to its new position in the current frame. This contour progress requires that certain amount of the object in the current frame overlay with the object region in the previous frame. Contour Tracking can be performed using two different approaches. The first approach uses state space models to model the contour shape and motion. The second approach directly evolves the contour by minimizing the contour energy using direct minimization techniques such as gradient descent. The most significant advantage of silhouettes tracking is their flexibility to handle a large variety of object shapes.

2. Shape Matching

These approaches examine for the object model in the existing frame. Shape matching performance is similar to the template based tracking in kernel approach.

Another approach to Shape matching [10] is to find matching silhouettes detected in two successive frames. Silhouette matching, can be considered similar to point matching. Detection based on Silhouette is carried out by background subtraction. Models object are in the form of density functions, silhouette boundary, object edges. Capable of dealing with single object and Occlusion handling will be performed in with Hough transform techniques.

V. CONCLUSION

In this paper various phases of object tracking system viz. object detection, object classification and object tracking has been studied. Available methods for these phases have been explained in details and a number of shortcoming and limitations were highlighted in each and every technique. Different methods for object detection are frame difference, optical flow and background subtraction. Object tracking can be performed using various methods like kalman filter, particle filter and multiple hypothesis tracking. It can be summarized background subtraction is a simplest method providing complete information about object compared to optical flow and frame difference for detecting objects. Advance study may be carried out to include find efficient algorithm to reduce computational cost and to decrease the time required for tracking the object for variety of videos containing diversified characteristics.

ACKNOWLEDGMENT

REFERENCES.

1. Abhishek Kumar Chauhan, Prashant Krishan, "Moving Object Tracking Using Gaussian Mixture Model And Optical Flow", International Journal of Advanced Research in Computer Science and Software Engineering, April 2013
2. M.Sankari, C. Meena, "Estimation of Dynamic Background and Object Detection in Noisy Visual Surveillance", International Journal of Advanced Computer Science and Applications, 2011, 77-83
3. Cheng-Laing Lai; Kai-Wei Lin, "Automatic path modeling by image processing techniques," Machine Learning and Cybernetics (ICMLC), 2010 International Conference on , vol.5, no., pp.2589,2594, 11-14 July 2010
4. Saravanakumar, S.; Vadivel, A.; Saneem Ahmed, C.G., "Multiple human object tracking using background subtraction and shadow removal techniques," Signal and Image Processing (ICSIP), 2010 International Conference on , vol., no., pp.79,84, 15-17 Dec. 2010
5. Ruolin Zhang, Jian Ding, "Object Tracking and Detecting Based on Adaptive Background Subtraction", International Workshop on Information and Electronics Engineering, 2012, 1351-1355.

6. K.Srinivasan, K.Porkumaran, G.Sainarayanan, "Improved Background Subtraction Techniques For Security In Video Applications"

7. Rupali S.Rakibe, Bharati D.Patil, "Background Subtraction Algorithm Based Human Motion Detection",International Journal of Scientific and Research Publications, May 2013

8. Jae-Yeong Lee; Wonpil Yu, "Visual tracking by partition-based histogram backprojection and maximum support criteria," Robotics and Biomimetics (ROBIO), 2011 IEEE International Conference on , vol., no., pp.2860,2865, 7-11 Dec. 2011

9. Mr. Joshan Athanesious J; Mr. Suresh P, "Implementation and Comparison of Kernel and Silhouette Based Object Tracking", International Journal of Advanced Research in Computer Engineering & Technology, March 2013, pp 1298-1303.

10. J.Joshan Athanesious, P.Suresh, "Systematic Survey on Object Tracking Methods in Video",International Journal of Advanced Research in Computer Engineering & Technology (IJARCET) October 2012, 242-247.

11. Sen-Ching S. Cheung and Chandrika Kamath, "Robust techniques for background subtraction in urban traffic video".

12. Greg Welch, Gary Bishop**,"** An introduction to the Kalman Filter", In University of North Carolina at Chapel Hill, Department of Computer Science. Tech. Rep. 95-041, July-2006.

13. Rahul Mishra, Mahesh K. Chouhan, Dr. Dhiiraj Nitnawwre, "Multiple Object Tracking by Kernel Based Centroid Method for Improve Localization", International Journal of Advanced Research in Computer Science and Software Engineering**,** July-2012, pp 137-140.

14. Hitesh A Patel, Darshak G Thakore,"Moving Object Tracking Using Kalman Filter", International Journal of Computer Science and Mobile Computing,April 2013, pg.326 – 332.