Scientific Journal of Impact Factor (SJIF): 3.134

e-ISSN (O): 2348-4470
p-ISSN (P): 2348-6406

International Journal of Advance Engineering and Research Development

Volume 3, Issue 1, January -2016

A Survey on object detection and tracking

Ms. Kruti Patel[1], Mr. Yask Patel[2]

[1] *Department of Computer Engineering, PIETVadodara,*
[2]*Department of Information Technology, PIET Vadodara,*

Abstract *- In Computer vision, the most active research topic is visual surveillance in dynamic scene especially for humans and vehicles. Locating a moving object or multiple objects over a time using a camera is a process called as video tracking. Video surveillance can be used for various purposes like human computer interaction, security and surveillance, video communication and compression, augmented reality, medical imaging and video editing. This paper presents the overview of various techniques that are used for tracking and detection methods and Comparison of the same.*

Keywords *– Object Tracking, Object Detection, Object Representation, Image processing and Feature Selection.*

I. INTRODUCTION

Image Processing is a form of processing with input as an image such as photograph or video frame and output can be characteristics or parameter related to image [1]. Videos are sequences of images which are actually called as frame. Object tracking can be defined as process of segmenting the part of interest from the image sequence and then keep track of its motion and extract some useful information from it. Object tracking is useful in several applications like video surveillance, video impainting, video editing, robot vision etc. Tracking object can be complex due to [2].

- Loss of information caused by projection of the 3D world on a 2D image.
- Noise in image
- Complex object motion
- Non rigid or articulated nature of object.
- Partial and full occlusion.
- Complex object shapes.
- Scene illumination changes and real time processing requirement.

In tracking, the trajectory of an object in the image plane is estimated as it moves around a scene. Basically, there are three steps in video analysis: detection of interesting moving objects, tracking of such object from frame. Some issues that need to be addressed when one sets out to develop an object tracker. The first issue is defining a suitable representation of object which will be described in section II. In section III, the next is discussed, that describe selection of image features used as an input for the tracker. In section IV, general object detection strategies are discussed and then we categorize some existing tracking methods in section V and finally we end up with the conclusion.

II.OBJECT REPRESENTATION

An Object can be anything that is of interest that we have to track like animals, vehicles for traffic monitoring, aeroplanes or humans etc. Object can be represented by their shapes and size [1]. There are number of ways to represent object by shapes, some are described below:

A. Points

An Object can be represented by a set of points or by a single point called centroid. The set of points occupies small region of interest that can be used for small area tracking purpose.

B. Primitive geometric shapes

Object shapes can be characterized by a rectangle, ellipse etc. Such representations for object motion are modeled by affine, translation or projective transformation. These primitive geometric shapes are used for characterizing simple rigid objects and they are also used for non-rigid objects.

C. Object silhouette and contour

The boundary of an object is represented as contour and the object silhouette is defined as the region inside the contour.

D. Skeletal Model

Skeletal model is used as shape representation for recognizing objects. This representation is used for both articulated object and rigid object.

E. Articulated shape model

Articulated objects are made of body parts that are made together with joints. For example, the human body is an articulated object with torso, legs, hands, head and feet connected by joints [1].

F. Skeletal models

Object skeleton can be taken by applying medial axis converted to the object silhouette. For recognizing objects, this model is mainly used as a shape representation. There are many ways to represents the appearance feature of objects. The shape and appearance both representations can be combined for tracking. Some appearance representations commonly employed in object tracking are:

G. Probability densities of object appearance:

The probability densities of object appearance features like color, texture etc. can be calculated from the image region specified by the shape models.

H. Templates

Geometric shapes or silhouettes represent a template which carries both spatial and appearance information. It is suitable only for tracking objects whose poses do not vary.

I. Multiple-view object recognition

This approach models an object with multiple views which contains information about less range of viewing condition. This representation can be done in two ways: First is sample image should be grouped that represents different view of an object and Second is to form model-view characteristics group members are widespread.

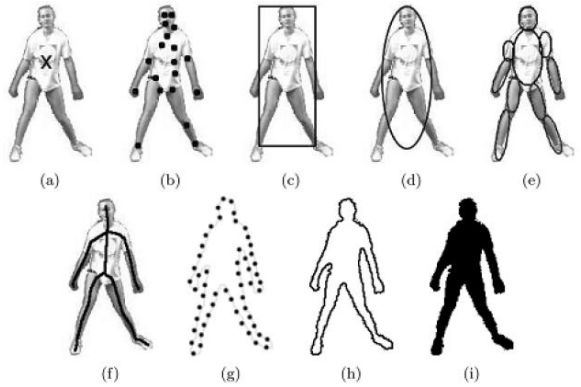

Figure1.Object representations [2] (a) Centroid, (b) multiple points, (c) rectangular patch, (d) elliptical patch, (e) part-based multiple patches, (f) object skeleton, (g) complete object contour, (h) control points on object contour, (i) object silhouette

III. FEATURE SELECTION

Feature selection plays an important role in tracking. Common features are as follows.

A. Color

In image processing, the RGB (red, green, blue) color space is usually used to represent color. There are lot of color spaces that have been used in tracking.

B. Edges

Edges are points where there is a boundary (or an edge) between two image regions [4].The main property of edges is that they are less sensitive to illumination changes than compared to color features.

C. Optical Flow

Optical flow as a feature selector is used mainly in motion based segmentation and tracking applications. It is computed using the brightness constraint which assumes brightness constancy of corresponding pixels in consecutive frames.

D. Texture [1]

Texture is a measure of the intensity variation of a surface which quantifies properties such as smoothness and regularity. Compared to color, texture requires a processing step to generate the descriptors.

IV. OBJECT DETECTION

Every tracking method requires an object detection mechanism either in every frame or when the object first appears in the video. A common approach for object detection is to use information in a single frame. A common approach for object detection is to use information in a single frame. However, some object detection methods make use of the temporal 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, which highlights changing regions in consecutive frames. Given the object regions in the image, it is then the tracker's task to perform object correspondence from one frame to the next to generate the tracks. Various object detection methods are:

A. Frame Differencing

The difference between two consecutive images is calculated by determining the presence of moving objects. 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 [4].

B. Optical Flow

Optical flow method [5] 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 antinoise performance, make it not suitable for real-time demanding occasions [6].

C. Background Subtraction

Background modeling is the first step of background subtraction. Background Modeling must sensitive enough to recognize moving objects [4].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 [6]. Background subtraction has mainly two approaches [7]:

1. **Recursive algorithm**

   Recursive techniques do not maintain a buffer for background estimation [7] [8]. On other hand, they recursively update a single background model based on each input frame. So that, the input frames from distant past could have an effect on the current background model. By Comparing 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 [6].

2. **Non-recursive algorithm**

   A non-recursive technique [7] [8] 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 [6].

Table 1. Comparative Study of Object Detection Methods [9]

Methods		Accuracy	Computational Time	Comments
Background subtraction	Gaussain 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

V. OBJECT TRACKING [1]

The aim of an object tracker is to generate the trajectory of an object over time by locating its position in every frame of the video.

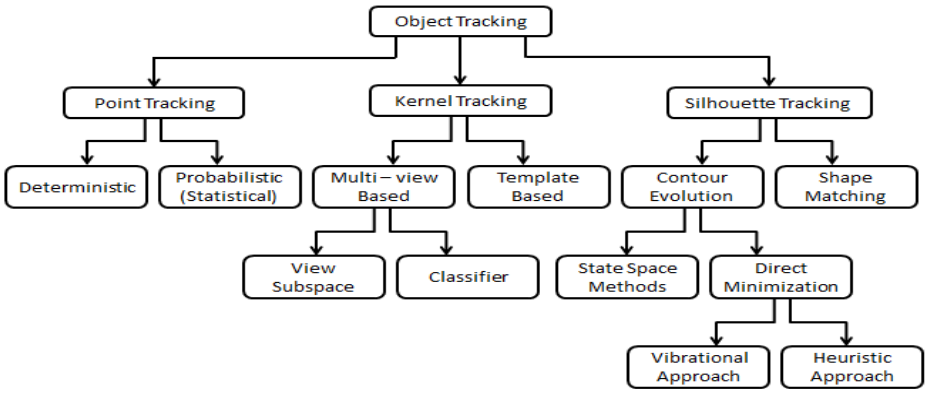

Figure2. Taxonomy of tracking methods [1]

A. Point Tracking

Objects detected in consecutive frames are represented by points, and the 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. An example of object correspondence is shown in Figure 3(a).

B. Kernel Tracking

Kernel refers to the object shape and appearance. For example, the kernel can be a rectangular template or an elliptical shape with an associated histogram. Objects are tracked by computing the motion of the kernel in consecutive frames (Figure 3(b)).The motion is usually in the form of a parametric transformation such as translation, rotation, and affine.

C. Silhouette Tracking

Tracking is performed by estimating the object region in each frame. Silhouette tracking methods use the information encoded inside the object region. This information can be in the form of appearance density and shape models which are usually in the form of edge maps. Given the object models, silhouettes are tracked by either shape matching or contour evolution (see Figure 3(c), (d)). Both of these methods can essentially be considered as object segmentation applied in the temporal domain using the priors generated from the previous frames.

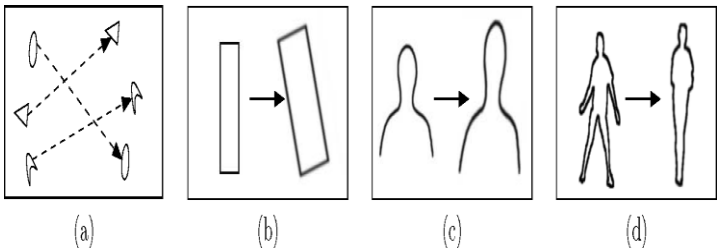

Figure 3. (a) Different tracking approaches. Multipoint correspondence, (b) parametric transformation of a rectangular patch, (c, d) two examples of contour evolution [1]

Table 2.Tracking Categories [1]

Categories	**Representative Work**
Point Tracking	
• Deterministic methods • Statistical methods	MGE tracker, GOA tracker, Kalman filter, JPDAF PMHT
Kernel Tracking	
Template and density based appearance models	Mean-shift, KLT , Layering
Multi-view appearance models	Eigen tracking , SVM tracker.

Silhouette Tracking	
Contour evolution	State space models , Variational methods , Heuristic methods.
Matching shapes	Hausdorff, Hough transform , Histogram

VI. CONCLUSION

Object tracking is an important task within the field of computer vision. It deals with estimating the trajectory of an object in the image plane as it moves around the scene. In this paper various object representation and feature selectors are briefly explained and the various object detection and tracking techniques are discussed in the later section. By using various tracking and detection techniques, objects can be detected and tracked and from that some useful information can be extracted.

VII. REFERENCES

[1] Alper Yilmaz., Omar Javed., Mubarak Shah., "Object Tracking : A Survey", In ACM Computing Surveys (2006).
[2] Shipra Ojha., Sachin Sakhare., "Image processing Techniques for Object Tracking in Video Surveillance- A Survey", In IEEE ICPC (2015).
[3] A seminar on Object Tracking in video scene by Alok K. Watve.
[4]Rupali S.Rakibe, Bharati D. Patil, "Background Subtraction Algorithm Based Human Motion Detection", International Journal of Scientific and Research Publications, May 2013
[5] 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
[6]Himani S. Parekh, Darshak G. Thakore ,Udesang K. Jaliya., "A Survey on Object Detection and Tracking Methods", In IJIRCCE Vol. 2,Issue 2, February 2014.
[7]Sen-Ching S. Cheung and Chandrika Kamath, "Robust techniques for background subtraction in urban traffic video".
[8]K. Srinivasan, K. Porkumaran, G. Sainarayanan, "Improved Background Subtraction Techniques For Security In Video Applications"
[9] 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

