International Journal of Engineering Technology Science and Research
IJETSR
www.ijetsr.com
ISSN 2394 – 3386
Volume 4, Issue 4
April 2017

Object Tracking Methods – A Survey

Pradnya Soman[1], Supragya Sonkar[1], Rishikesh Yadav[1], Jyoti Wadmare[2]
[1]Student, BE, K. J. Somaiya Institute of Engineering and Information Technology, Sion, Mumbai
[2]Dept.of Computer Science, K.J. Somaiya Institute of Engineering and Information Technology,Sion, Mumbai

Abstract—Object tracking is the process of locating moving objects in video sequences. Object tracking is an essential task in many application of computer vision such as vehicle navigation, autonomous robot navigation, human-computer interaction, security and video surveillance, augmented reality, traffic control etc. To perform video tracking an algorithm analyzes sequential video frames and outputs the movement of targets between the frames. Its main task is to find and follow a moving object or multiple objects in image sequences. Tracking a target is still one of the challenging problems of video surveillance. The process of object tracking consists of a number of stages, some of them being feature detection, object detection, its classification and tracking. This paper presents a brief overview of various video object tracking stages and techniques and examines their pros and cons.

Keywords—object tracking; image; video surveillance; object detection

I. INTRODUCTION

In image processing, videos are represented as some hierarchical structure units like scene, shot and frame. In video retrieval, generally, video application must partition a given video sequences into video frames. The complete video sequence is formed by joining two or more frames which are considered as input to the object tracking Real time object tracking is defined as the process of estimating trajectory or path of an object in consecutive frames. The main goal of object tracking is the process of segmenting the interest of an object or multiple objects from the video scene and keeping track of its motion, occlusion and orientation.

Object tracking is an important and challenging task in the field of computer vision. The rise of high-powered computers, the availability of high quality and inexpensive video cameras, and the rising need for automated video analysis has origin a great deal of interest in object tracking algorithms. Moreover, depending on the tracking domain, a tracker can also give object-centric information, such as orientation, region, or silhouette of an object.

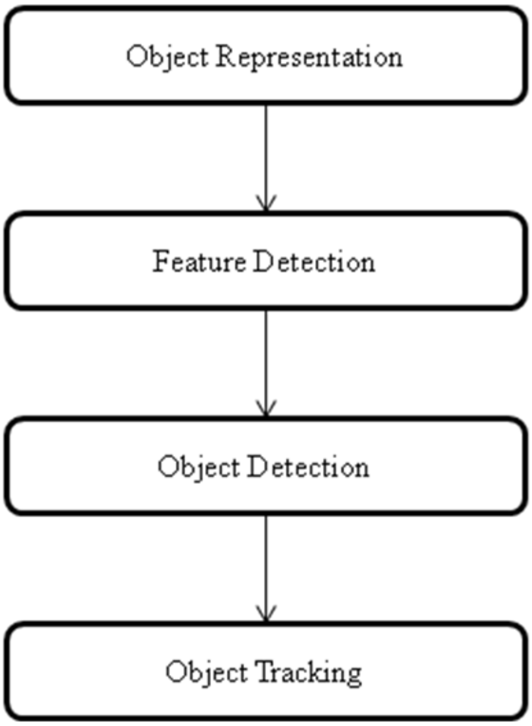

Fig 1: Flow of object tracking stages

Video object detecting and tracking has many applications in video processing such as video compression, video surveillance, vision-based control, human-computer interfaces, medical imaging, augmented reality, and robotics.

It is challenging and interesting task to track the object of interest and some difficulties or challenges appear during the various stage of object tracking. Tracking objects can be complex due to loss of information caused by projection of the 3D world on a 2D image, noise in images, complex object motion, partial and full object occlusions, complex object shapes, scene illumination changes, and real-time processing requirements. In some algorithms, the moving objects may become part of the scene when they come to a stop.

Prior knowledge about the number and the size of objects, or the object appearance and shape, can be used to simplify the problem. Many techniques are available in image processing to overcome different type of difficulties.

The process of object tracking mainly involves steps like suitable representation of the object, feature detection-selection of image features used as an input for the tracker, object detection and lastly object tracking.

II. Object Representation

In a tracking scenario, an object can be defined as anything that is of interest for further analysis. For instance, boats on the sea, fish inside an aquarium, vehicles on a road, planes in the air, people walking on a road, or bubbles in the water are a set of objects that may be important to track in a specific domain. Objects can be represented by their shape and size. Following are some object shape representations commonly employed for tracking

A. Points

An object can be characterized in terms of set of points that occupy small region of interest used for small area tracking purpose. Object may also be represented as a single point called as centroid of a person. In general, the point representation is suitable for tracking objects that occupy small regions in an image.

B. Geometric shapes

This is most suitable for simple rigid-objects. The shape of object can be represented as rectangle or ellipse, also used for non-rigid object tracking. Though primitive geometric shapes are more suitable for representing simple rigid objects, they are also used for tracking non-rigid objects.

C. Object silhouette

The area inside the contour represents the silhouette of an object. Contour represents borderline of an object. Both contour and silhouette are used for complex non-rigid shape tracking.

D. Skeletal Model

Object skeletal can be extracted by medial axis transform, to object silhouette. Medial axis represents the distance to the boundary. Every point P in a region R finds the nearest neighbour in the boundary, and its point belong to medial axis of the region. This model is commonly used as a shape representation for recognizing objects. Skeleton representation can be used to model both articulated and rigid objects.

E. Probability density appearance model

This model is used in Gaussian mixture models and histogram. The probability densities of object appearance features (color, texture) can be computed from the image regions specified by the shape models (interior region of an ellipse or a contour).

F. Multiview appearance models.

These models encode different views of an object. Each view contains information about small range of viewing condition. One approach to represent the different object views is to generate sample images which are clustered into group that represents different view of an object. Another approach to learn the different views of an object is by training a set of classifiers. Group members are widespread to form model-view characteristics. One limitation of multiview appearance models is that the appearances in all views are required ahead of time

G. Templates

Geometric shape represents a template, which contains both spatial and appearance information. It is mainly used for object tracking purpose. An advantage of a template is that it carries both spatial and appearance information. Templates, however, only encode the object appearance generated from a single view. Thus, they are only suitable for tracking objects whose poses do not vary considerably during the course of tracking.

H. Articulated shape models. Articulated objects are composed of body parts that are held together with joints. For example, the human body is an articulated object with torso, legs, hands, head, and feet connected by joints. The relationships between the parts are governed by kinematic motion models, for example, joint angle, etc. In order to represent an articulated object, one can model the constituent parts using cylinders or ellipses.

I. Active appearance models. Active appearance models are generated by simultaneously modeling the object shape and appearance. In general, the object shape is defined by a set of landmarks.

Similar to the contour-based representation, the landmarks can reside on the object boundary or, alternatively, they can reside inside the object region. Active appearance models require a training phase where both the shape and its associated appearance are learned from a set of samples using, for instance, the principal component analysis.

In general, there is a strong relationship between the object representations and the tracking algorithms. Object representations are usually chosen according to the application domain. For tracking objects, which appear very small in an image, point representation is usually appropriate. For tracking objects with complex shapes, for example, humans, a contour or a silhouette based representation is appropriate.

III. Feature Detection

In image processing, a feature is in simple terms, a piece of information relevant for sorting out the computational task related to a specific application. Features may be specific structures related to the image like edges, points, objects, texture, etc. In complex scenarios sometimes only one type of feature may not give sufficient information of the image data, which results in two or more features to be extracted. Selecting the right features plays a critical role in tracking. For example, color is used as a feature for histogram-based appearance representations, while for contour-based representation, object edges are usually used as features. In general, many tracking algorithms use a combination of these features. The details of common visual features are as follows.

A. Color

The apparent color of an object is influenced primarily by two physical factors, 1) the spectral power distribution of the illuminant and 2) the surface reflectance properties of the object. In image processing, the RGB (red, green, blue) color space is usually used to represent color. The RGB space uses combination of color and contour feature in particle filter based object tracking. Hue, saturation (lightness) and value (brightness) (HSV) are the three components of HSV space is also one of the uniform color space but it is sensitive to noise.

B. 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. Algorithms that track the boundary of the objects usually use edges as the representative feature. Because of its simplicity and accuracy, the most popular edge detection approach is the Canny Edge detector.

C. Texture

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. Similar to edge features, the texture features are less sensitive to illumination changes compared to color.

D. Moment

A moment based region feature is used for object recognition and determine the gender of a strange person by walking appearance.

E. Optical Flow.

Optical flow is a dense field of displacement vectors which defines the translation of each pixel in a region. It is computed using the brightness constraint, which assumes brightness constancy of corresponding pixels in consecutive frames. Optical flow is commonly used as a feature in motion-based segmentation and tracking applications.

Mostly features are chosen manually by the user depending on the application domain.

IV. Object Detection

Every object tracking must requires object detection as first step to identify objects of interest in the video sequences. Object detection is the locating moving object in the consecutive video frames. It can be done using different well known techniques. 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. 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.

A. Frame Differencing

It is a technique to check the difference between two consecutive frames. Frame differencing employs the input as two image frames of video and produces the output such as the new pixel value or difference of the pixel values that can be obtained by subtracting pixels value of second frame image from the first frame image. Method is easy and simple to implement but detection of object cannot give accurate result.

B. Optical flow

Optical flow is a technique that presents an apparent change in the moving object's location between frames of given video. It employs the motion field that represents the directions and velocity of each point in every frame. It takes more time to detect complex object motion and this technique is more suitable for multiple moving object detection.

C. Background Subtraction

Background subtraction method extracts the moving objects or foreground object. For that compare the reference background image to the current image and find the difference in pixel values between consecutive frames. When the difference has detected, classify that object as moving object. Performance of the background subtraction is well for static background and deals with the multiple moving objects.

V. Object Tracking

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. Object tracker may also provide the complete region in the image that is occupied by the object at every time instant. The tasks of detecting the object and establishing correspondence between the object instances across frames can either be performed separately or jointly. . The model selected to represent object shape limits the type of motion or deformation it can undergo. For example, if an object is represented as a point, then only a translational model can be used. In the case where a geometric shape representation like an ellipse is used for the object, parametric motion models like affine or projective transformations are appropriate. These representations can approximate the motion of rigid objects in the scene. For a non-rigid object, silhouette or contour is the most descriptive representation and both parametric and nonparametric models can be used to specify their motion.

A. Point based tracking

It is a complex problem to track an object according to the feature points of that objects. It causes the false detection of object and occlusion of object in frames. There are various methods that use the point based tracking as below:

Point Tracking Methods

1) Kalman Filter: Kalman filter is based on probability density function. It is a set of mathematical equations. So it is complex method but it gives always optimal solution. Feedback control can be estimated by using kalman filter.

2) Particle Filter: Particle filtering generates all models for one state variable before moves to next variable. Restriction of Kalman filter is the assumption of state variable can be overwhelmed by using the particle filtering.

3) Multiple Hypothesis Tracking: Multiple Hypothesis tracking is also known as iterative algorithm that begins with the set of existing track hypothesis. Object's position in every frame is made for each hypothesis. This method has ability to track multiple objects and handle occlusion.

B. Kernel based tracking

Kernel based tracking refers to the appearance of object and object shape. Kernel can be elliptical or rectangular shape. Object motion is the main feature of kernel based tracking method. The motion of object is may be in form of parametric transformation either translation, affine or rotation. Various kernel based tracking techniques are described:

Kernal Tracking Methods

1) Simple Template Matching: Simple template matching is the technique to find the small region of the image that matches with the template image in each frame of given input video. This template image is considered as reference image for the next frame. After that, all possible scenarios can be calculated to know that how well the model fits the position of picture in the frames. This method has ability to track single image and handles partial occlusion.

2) Mean Shift Method: Mean shift method is to finding local maxima of density function from the given discrete data samples. It is a non parametric feature space analysis method. Mean shift estimates the positions of the region in the current frame from the previous frame. It is gradient ascent strategy that tracked the object by use of histogram.

3) Support Vector Matching: Support vector matching is the technique to give a set of positive and negative training values. In that, positive set of values are considered as tracked object and other remaining objects are considered as negative training values that are not tracked image object. It can capable of handling partial occlusion but with the necessity of training and deal with the single image.

4) Layering Based Tracking: Multiple objects can be tracked by using the layering based tracking. Each layer consists of motion like translation and rotation, shape such as ellipse and layer appearance that is 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. Layering based tracking deals with the fully occlusion of multiple objects and handles multiple image tracking.

C. Silhouette based tracking

Silhouette based tracking is the method that use the information encoded inside object region for tracking. Objects may have complex shape like as shoulders, finger and hand that cannot be described properly by simple geometric shapes.

For that silhouette based tracking is used to define accurate shape of object. It has ability to handle occlusion, object split and merge and deal with the various shapes of objects.

Silhouette Tracking Methods :

1) Contour Matching: Contour matching select points on the boundary of object in every frame. Calculate the contour in previous frame and it's new position in current frame. Two different approaches are used for tracking. First is state space model that specify shape and motion of contour. Second approach minimizes the contour energy using direct minimization techniques like gradient descent. It is capable of dealing large variety of object shapes for tracking.

2) Shape Matching: Shape matching is the similar technique as template matching. Detected silhouettes in every frame can be found and after that, shape matching is applied to matches the shape or silhouettes detected in two consecutive frames. This technique is capable of handling occlusion using Hough transform techniques and tracking single object from the given video sequences.

Objects may have complex shapes, for example, hands, head, and shoulders that cannot be well described by simple geometric shapes. Silhouette based methods provide an accurate shape description for these objects. The goal of a silhouette-based object tracker is to find the object region in each frame by means of an object model generated using the previous frames. This model can be in the form of a color histogram, object edges or the object contour.

VI. CONCLUSION

In this paper, we started with how objects can be represented with important feature descriptors. We explained tracking process in detail from detection and recognition with different approaches including background subtraction, temporal differencing and optical flow, till object tracking. In the later sections, we described object tracking types such as point tracking, kernel tracking and silhouette tracking. Our further research will focus on to develop a silhouette based tracking method which can deal with the moving objects in a video frame and can use the color and shape feature to detect objects.

REFERENCES

[1] Shraddha V. Kothiya, Kinjal B. Mistree, "A Review on Real Time Object Tracking in Video Sequences", *IEEE* 2015

[2] Shipra Ojha, Sachin Sakhare, "Image Processing Techniques for Object Tracking in Video Surveillance- A Survey", *IEEE* 2015

[3] Manorama Pokheriya, Debashish Pradhan, "Object Detection and Tracking based on Silhouette Based trained shape model and Kalman Filter, *IEEE* 2014.

[4] Li Guan, Jean-Sebastian Franco, Marc Pollefeys, "Multi-Object Shape Estimation and Tracking from Silhouette Cues", *IEEE* 2008.

[5] Alper Yilmaz, Omar Javed, and Mubarak Shah, “Object tracking: Asurvey”, *ACM Computing Surveys* Vol 38, No. 4, Article 13, Publication date: December 2006

[6] Meha J. Patel, Bhumika Bhatt, “A Comparative Study of Object Tracking Techniques”, *IJIRSET 2015*

[7] S. R. Balaji, Dr. S. Karthikeyan, “A Survey On Moving Object Tracking Using Image Processing”, *IEEE 2017*