A Survey on Approaches for Object Tracking

Nandita Singh¹, Manoj Sabnis²
¹,²(Department of Information Technology, VESIT, INDIA)

ABSTRACT: Object detection and tracking has been a widely studied research problem in recent years. Currently system architectures are service oriented i.e. they offer number of services. All such common services are grouped together and are available as a domain called as service domain. One such service domain of our interest is LBS (location based service). The service of our interest is tracking. Tracking of moving objects is done in applications like surveillance systems, human computer interactions, object recognition, navigation systems etc. In real world, 3D, the object which we want to track is called as object of interest (OOI). Tracking has been a difficult task to apply in congested situations due to inaccurate segmentation of objects. Common problems of erroneous segmentation are long shadows, partial and full occlusion of objects with each other and with stationary items in the scene. Difficulties in tracking objects can arise due to abrupt object motion, changing appearance patterns of both the object and the scene, nonrigid object structures, object-to-object and object-to-scene occlusions, and camera motion. In this paper we analyze different approaches for moving object tracking and detection.

Keywords: Multiple moving object tracking, background modeling, morphology, target localization and representation, visual surveillance.

I. Introduction

Object tracking is the process of establishing correspondence between objects and object parts given consecutive frames of video locating a moving object in time using a camera. There are three key steps in video analysis: detection of object of interest, tracking of such objects from frame to frame, and analysis of object tracks to recognize their behaviour. To track an object in a video the object is segmented into sequences of video scenes. The algorithm analyses the video frames and outputs the location of moving targets within the video frame. Object tracking has several applications like it is used in video surveillance, robot vision, traffic monitoring, Video in painting and Animation. The basic steps for object tracking are mentioned below:

Object Detection: Object detection deals with the detection of Object of Interest in the video sequence. The basic techniques used for object detection are Background subtraction, Optical flow, Temporal differencing and frame differencing.

Object Representation: OOI can be represented by their shapes and appearances. The various methods used in practice are Shape-based representation, Motion-based representation, Color based representation and texture based representation.

Object Tracking: Tracking can be defined as the problem of estimating the trajectory of an object in the image plane as it moves around a scene. There are two major components of a visual tracking system: target representation and localization, as well as filtering and data association

II. Object detection

Every tracking method requires object detection mechanism. A common approach for object detection is to use information in a single frame, whereas some object detection methods make use of the temporal information computed from a sequence of frames to reduce the number of false detection. This temporal information is usually in the form of frame differencing, which highlights changing regions in consecutive frames.

Some of the common object detection methods are as follows:

- Temporal differencing
- Optical Flow
- Background subtraction

A. Temporal Differencing: The frame differencing can be considered as the simplest form of background subtraction. In this the current frame is simply subtracted from the previous frame and if a pixel value differs greater than the predefined threshold then the given pixel is considered part of the foreground [9].

Pros: strongly adaptable for dynamic environment.
Cons: inaccuracy due to difficulty in obtaining a complete outline of moving object.
B. Optical Flow: In Optical Flow each point in the image is assigned a motion vector. It is a pixel-level representation model for motion patterns. Optical Flow can be classified into three major categories, which are the phase correlation, block-based methods, and gradient-based estimation, respectively. According to the study of Zheng and Xue, the phase correlation is a fast frequency-domain approach to estimating the relative movement between two images. Therefore, the block-based methods minimize the sum of squared differences or the sum of absolute differences, or maximize the normalized cross-correlation. It has been frequently used in video compression standards. Moreover, the gradient based estimation is a function of the partial derivatives of the image signal and/or the desired flow field and higher order partial derivatives.

Pros: capable of getting the complete movement information and detect the moving object from the background.
Cons: involves a large quantity of calculation, sensitivity to noise, poor anti-noise performance.

C. Background Subtraction: The idea is to build background model first and then finding deviations from the model for each incoming frame. If the pixel difference is greater than the set threshold, then determines that the pixels appear in the moving object, otherwise, as the background pixels[10]. The pixels constituting the regions undergoing change are marked for further processing.

Pros: Objects are allowed to become a part of the background without destroying the existing background, It learns itself and does not need to be reprogrammed.
Cons: poor accuracy, cannot sustain multi-valued background.

III. Methods Of Object Representation

In a process of tracking, an object can be defined as anything that a person wants to track. In real world, 3D, the object which we want to track is called as object of interest (OOI). OOI can be represented by their shapes and appearances. The various methods of object representations are discussed below:

D. Shape-based representation: It is based on analysis of the geometric shapes to detect similarly shaped objects. The shapes of object can be determine using point method, geometric shapes, Object silhouette and contour, Articulated shape models or Skeletal models. It’s important to choose an appropriate object representation in accordance to the tracking application, as there is a close relationship between the object representations and the tracking algorithms.

E. Motion-based representation: It works on periodicity of the motion. The motion of the object is studied and based on the moving pattern object can be classified. It can represent both rigid as well as non rigid object.

F. Color based representation: The apparent color of an object is influenced primarily by two physical factors, a) the spectral power distribution of the illuminant and b) the surface reflectance properties of the object. In image processing, the RGB (red, green, blue) color space is usually used to represent color. However, the RGB space is not a perceptually uniform color space, that is, the differences between the colors in the RGB space do not correspond to the color differences perceived by humans. HSV (Hue, Saturation, and Value) is a relatively uniform colour space.

G. Texture based representation: Texture is a measure of the intensity variation of a surface which quantifies properties such as smoothness and regularity. It contains important information about the structural arrangement of surfaces and their relationship to the surrounding environment. Texture is represented by means of texture descriptors. They observe region homogeneity and histograms of region borders. Compared to color, texture requires a processing step to generate the descriptors.

IV. Approaches Of Object Tracking

There are two major components of a visual tracking system: target representation and localization, as well as filtering and data association [2].

H. Target representation and localization

Target Representation and Localization is mostly a bottom-up process. Typically the computational complexity for these algorithms is low.

Blob tracking: The Blob Tracking provides a way to track blobs (collections of pixels) from one image to the next. Often as part of any object tracking solution it is necessary to identify not only that object moved but which object was which when comparing the current frame and the last frame. The Blob Tracker will label each blob with a specific id that will be attached to the same or similar blob in the next frame. What defines a blob as being similar in two images depends on how you have configured the Blob Tracking module.
Mean shift: It is an iterative approach approach to feature space analysis. As name implies it shifts a data point to the average of data points in its neighborhood similar to clustering. It has many applications in the area of visual tracking and probability density estimation. Mean shift is an application-independent tool suitable for real data analysis. The Continuously Adaptive Mean Shift Algorithm (CAMSHIFT)[6] is an adaptation of the Mean Shift algorithm for object tracking that is intended as a step towards head and face tracking for a perceptual user interface. CAMSHIFT can handle dynamically changing color distribution by adapting the search window size and computing color distribution in a search window.

Contour Tracking:- These methods [7] 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. There are two different approaches to perform Contour Tracking. 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.

Visual Feature matching:- These approaches examine for the object model in the existing frame. its performance is similar to the template based tracking in kernel approach. Another approach 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. Filtering and Data Association

Filtering and Data Association is mostly a top-down process, which involves incorporating prior information about the scene or object, dealing with object dynamics. The computational complexity for these algorithms is usually much higher.

Kalman Filter: The Kalman filter operates recursively on streams of noisy input data to produce a statistically optimal estimate of the underlying system state. The Kalman Filter performs the restrictive probability density propagation. The Kalman filter has two distinctive features. One is that its mathematical model is described in terms of state-space concepts. Another is that the solution is computed recursively. Kalman filter 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 modeled system is unknown. The Kalman filter [8] 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.

**Particle filter:** - The particle filtering [8] 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 re-sampling. 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 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.

**VI. Comparative study**

Kalman filter are based on Optimal Recursive Data Processing Algorithm. Here Gaussian state distribution is assumed. Kalman filtering is composed of two stages, prediction and correction. Prediction of the next state using the current set of observations and update the current set of predicted measurements. The second step is gradually update the predicted values and gives a much better approximation of the next state. Kalman filter tries to find a balance between predicted values and noisy measurements. The value of the weights is decided by modeling the state equations. Analysis shows that Kalman filtering approach is capable in dealing with noise. It is applicable only for single object [1], [3]. Kalman Filter always gives optimal solution. It is used in vision community tracking. Particle filter is used to track non-linear, non-Gaussian moving objects. Algorithm based on codebook [5] and particle filter is used to detect moving objects. The algorithm uses codebook background model for detection of objects, then color histogram of every objects is obtained and particle sampling range is limited by the combination of foreground detection information, which results particle filter reflect the objects more exactly and timely. An algorithm of codebook background model is used for object detection and Particle filter algorithm is used for object tracking.

<table>
<thead>
<tr>
<th>NATURE</th>
<th>Kalman filter</th>
<th>Particle filter</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type of object tracker</td>
<td>Single object tracker</td>
<td>Multiple object tracker</td>
</tr>
<tr>
<td>Method type</td>
<td>Recursive system based on non linear.</td>
<td>Recursive implementation of monte carlo</td>
</tr>
<tr>
<td>Mean square error</td>
<td>MSE obtain through this is almost the same.</td>
<td>Decreases.</td>
</tr>
<tr>
<td>Processing</td>
<td>Based on Gaussian distributed.</td>
<td>Based on probability distribution.</td>
</tr>
<tr>
<td>Computation time</td>
<td>Takes less time on executing the state.</td>
<td>Takes more time.</td>
</tr>
</tbody>
</table>

Table1: Comparison between Kalman filter and optical filter.

Contour tracking methods develop an original contour in the foregoing frame to its new position in the present frame, overlapping of object between the current and next frame. Contour tracking is in the form of state space models [1]. State Space Models: State of the object is named by the parameters of shape and the motion of the contour. The state is updated for each time according to the maximum of probability. In [1] Author has used two types of object representation one is implicitly modelled and the other one is explicitly modelled. Performance of the technique based on contour evolution by direct minimization has been analysed. Here region statistics is calculated using grid points. Occlusion is fully handled. Visual feature matching checks for object model in the existing frame [1]. Visual feature matching performance is similar to template based tracking in kernel approach. Another approach to feature matching is to find matching silhouettes in two successive frames. Detection based on Silhouette is carried out by background subtraction. Models object are in the form of density functions, silhouette boundary, object edges [1].
Table2: Comparison between contour tracking and visual feature matching.

### VII. Conclusions

In this paper various approaches for object Tracking has been studied. Comparative study is being done for Kalman filter & particle filter, contour tracking & visual feature Matching. 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.

### References

3. Dorin Comaniciu, Senior Member, IEEE, Visvanathan Ramesh, Member, IEEE, and Peter Meer, Senior Member, IEEE “Kernel-Based Object Tracking” IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE, VOL. 25, NO. 5, MAY 2003.
7. ‘Two-Stage Object Tracking Method Based on Kernel and Active Contour’ Qiang Chen, Quan-Sen Sun, Pheng Ann Heng and De-Shen Xia, IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY, VOL. 20, NO. 4, APRIL 2010.