

## Moving Object Analysis Techniques In Videos - A Review

Ritika<sup>1</sup>, Gianetan Singh Sekhon<sup>2</sup>

<sup>1</sup>M.Tech CE Student, Ycoe, Guru Kashi Campus/ Punjabi University, Talwandi Sabo

<sup>2</sup>Assistant Professor, Ycoe, Guru Kashi Campus/ Punjabi University, Talwandi Sabo

**ABSTRACT** -- Object tracking is an important task within the field of computer vision. It is a challenging problem. Many difficulties arise in tracking the objects due to abrupt object motion, changing appearance patterns of both the object and the scene, non-rigid object structures, object-to-object and object-to-scene occlusions, and camera motion. This paper selectively reviews the research papers with regard to tracking methods on the basis of the object, their motion representations and all detailed descriptions of representative methods in each category examining their advantages/disadvantages. It also discusses the important issues related to tracking including the use of object representation, tracking, and detection.

**Keywords** – Object Representation, Object Tracking, Object Detection, Computer Vision.

### I. INTRODUCTION

To develop the real world computer vision system, tracking of moving objects is very important task. The proliferation of high-powered computers, the availability of high quality and inexpensive video cameras, and the increasing need for automated video analysis has generated a great deal of interest in object tracking algorithms [7]. Applications are like automatic video surveillance, motion-based recognition, video indexing, human-computer interaction, traffic monitoring, and vehicle navigation. The rest of the paper is organized as follows: Section 2 offers the primary concepts of object tracking. Section 3 presents the existing algorithms for object tracking, representation and detection. Section 4 concludes the paper.

### II. PRIMARY CONCEPTS OF OBJECT TRACKING

#### 2.1 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. In other words, a tracker assigns consistent labels to the tracked objects in different frames of a video. Additionally, depending on the tracking domain, a tracker can also provide object-centric information, such as orientation, area, or shape of an object [7]. Object tracking is the process of segmenting an object of interest from a video scene and keeping track of its motion, orientation, occlusion etc. in order to extract useful information.

There are numerous approaches for object tracking. These differ from each other based on the way they approach as: Which object representation is suitable for tracking? Which image features should be used? How should the motion, appearance, and shape of the object be modeled? The answers to these questions depend on the environment in which the tracking is performed and the end use for which the tracking information is being sought. A large number of tracking methods have been proposed which attempt to answer these questions for a variety of scenarios [7].

The motive of an object tracker is to generate the trajectory of an object over time by locating its position in every frame of the video. The tasks of detecting the object and establishing correspondence between the object instances across frames can either be performed separately or jointly [7]. There are some tracking methods given below:

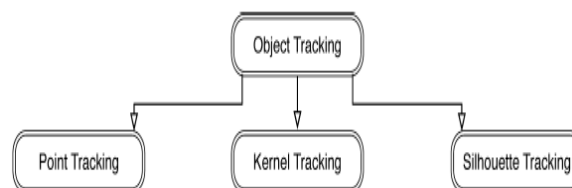


Fig.1 object tracking methods

- 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 [7].
- Kernel Tracking: Kernel tracking is based on the shape of the object and its appearance. Objects are tracked in consecutive frames. For example kernel can be rectangular or elliptical shape.
- Silhouette Tracking: Silhouette tracking methods use the information encoded inside the object region [7]. Silhouettes are also tracked by shape matching or contour evolution.

## 2.2 Feature Selection for Tracking

Feature is the uniqueness property of the object, so it can be easily distinguished in the feature space. Feature selection is closely related to the object representation [7]. Hence selecting the right feature plays the main role in tracking. The common features for tracking are as follows:

- Color: In image processing, the RGB (red, green, blue) color space is usually used to represent color. A variety of color spaces have been used in tracking.
- Edges: Boundaries of object generate strong changes in image intensities. To identify these changes edge detection is used.
- Optical Flow: Optical flow is a dense field of displacement vectors which defines the translation of each pixel in a region. Optical flow is commonly used as a feature in motion-based segmentation and tracking applications [7].
- Texture: It is a measure of the intensity variation of a surface which quantifies properties such as smoothness and regularity.

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,
- Non-rigid or articulated nature of objects,
- Complex object shapes,
- Scene illumination changes and real-time processing requirements [7].

## 2.3 Object Representation

An object can be anything like bubbles in the water, vehicles in the road, boats in the sea, people walking on the road, in the tracking scenario. Objects can be represented by their shapes and appearances. For tracking, the shape of the appearance object can be representing as below:

- Points: The object is represented by a point, that is, the centroid or by a set of points; suitable for tracking objects that occupy small regions in an image.
- Primitives Geometric Shapes: Object shape is represented by a rectangle, ellipse, etc.
- Object silhouette and contour: Contour representation defines the boundary of an object. The region inside the contour is called the silhouette of the object.
- Articulated shape models: Articulated objects are composed of body parts that are held together with joints.
- Skeletal models: Object skeleton can be extracted by applying medial axis transform to the object silhouette.

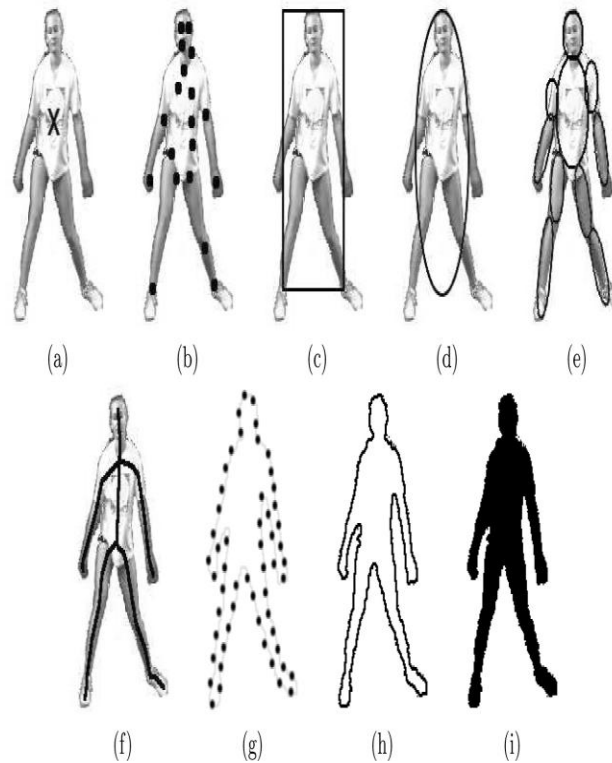


Fig. 2 object representations[7] (a) centroid, (b) multiple points, (c) rectangular patch, (d) elliptical patch, (e) part-based multiple patches, (f) object skeleton, (g) complete object contour, (h) control points on object contour, (i) object silhouette

## 2.4 Object Detection

Detection of moving objects in video images is very important. The automatic detection of moving objects in monitoring system needs efficient algorithms. The common method is simple background subtraction i.e to subtract current image from background. But it can't detect the difference when brightness difference between moving objects and background is small. The other approach is to use some algorithms such as color based subtraction technique [8]. There are several object detection techniques, which are given below:

- Point detectors: In this technique, region of interest in image is considered.
- Segmentation: In this, image is partitioned into similar regions.
- Background Subtraction: It means, to subtract current image from the background [8].

## III. RELATED WORK

There is much research work in the field of object tracking in videos over the past decades. Some of the work done has been discussed one by one below.

Alexander Toshev, Ameesh Makadia et al. [4], Presented shape based object recognition in Videos Using 3D Synthetic Object Models. This paper sorted the problem of recognition of moving objects from the videos by synthetic 3D models. At first, from the video, the silhouette images of the moving object is extracted by feature tracking, motion grouping of tracks and co-segmentation of successive frames and then matched to 3D model silhouettes. As a result, the matching of every 3D model to the video. This approach can recognize objects in videos and estimate their rough pose by using only similar but not 3D models.

Mohammed Sayed and Wael Badawy [11], Presents a novel motion estimation method for mesh-based video motion tracking. The method called mesh-based square-matching (MB-SM) motion estimation method. This method is used in terms of computational cost reduction, efficiency and image quality. It is a modified version of the hexagonal matching motion estimation method. The MB-SM motion estimation is performed in two ways. First is the rough motion estimation using block-matching algorithm. Second is the fine motion

estimation to refine the motion vectors generated from the first step. As result, MB-SM method has lower computation cost than the hexagonal matching motion estimation method while it produces almost the same PSNR values.

Minglun Gong [9], Proposed estimating 3D geometry and motion of dynamic scenes based on captured stereo sequences. A dynamic programming based technique is used for searching global optimal disparity maps and disparity flow maps under an energy minimization framework. As a result, both computations can benefit from each other and are capable of producing both 3D geometry and motion information for dynamic scenes in nearly real time.

Ming-Yu Shih et al. [5], In this paper, a method of moving object detection on moving platforms is proposed. This method composed of moving blob detection and shape refinement phases to provide robust moving object detection result. By fusing motion field's information from three consecutive frames, positions of moving blobs were precisely detected. Next, using motion compensated background models in intensity,  $r$ , and  $g$  color spaces, shapes of objects are well refined in the fused background subtraction process. By combining moving blob and contextual background information, alignment errors could be eliminated effectively to prevent foreground pixels from being adapted into background models.

Sajjad Torkan, Alireza Behrad [2], proposed a new contour based tracking method using active contour. Original greedy snake as a parametric active contour has weak performance in tracking target with high velocity and large displacement between two successive frames. This is due to concave parts of target boundary, shrinkage of contour if it is far from target boundary and the lack of target motion information. To resolve this problem a greedy snake with adaptive curvature energy and additional field energy term as an external energy. Kalman filtering is used for handling large target displacement and contour concavities. Kalman filtering is also used for the estimation of target shape and centroid in current frame. Applied on wide variety of video sequences results showed that the algorithm capable of tracking target with high speed, large aspect change and contour concavities.

Baiyang Liu, Lin Yang et al. [3], proposed an adaptive tracking algorithm for lung tumors in fluoroscopy using online learned collaborative trackers. No shape or motion priors are required for this tracking algorithm. This saves many expensive expert annotations. Accurate tracking of tumor movement in fluoroscopic video sequences is a clinically significant and challenging problem. This is due to blurred appearance, unclear deforming shape, complicate intra and inter- fractional motion, and other facts. Current offline tracking approaches are not adequate because they lack adaptivity and often require a large amount of manual labeling. Therefore this adaptive online learning algorithm is general to be extended to other medical tracking applications.

Pengwei LIU, Huiyuan WANG et al. [1], presented an approach for handling target detection and tracking in dynamic scenes, in which, motion compensation is generated by pyramidal optical flow. Because In most cases of target tracking in video sequences, the scene is dynamic due to the mobile cameras and causes problems. Main contribution in this system is that it separates the values of pyramidal optical flow into two groups: one represents background and the other foreground. Experiments show that the algorithm achieves motion compensation very effectively. It is only applied in two-dimensional motion cases such as that the camera moves only up and down.

Mark Ritch, Nishan Canagarajah [6], proposed a method to identify and track an object of interest within compressed MPEG-2 video using only motion information. The system is designed to detect interesting events taking place, such as moving object appearing and to track it without the need to decode the video. Iterative rejection is used as a basis for initial segmentation and those macro blocks whose motion vector is large than average are rejected. This is done by adopting a novel model based approach. Experimental results on a number of sequences demonstrate its effectiveness in identifying and tracking on object of interest from a compressed video stream without the need to fully decode each frame and that the system performed better than using iterative rejection alone as a segmentation method.

Huiqiong Chen, Derek Rivait and Qigang Gao [10], presents perceptual organization based method for real-time license plate identification and tracking. In this method, video content is described by Generic Edge Tokens (GETs), and an image is represented as a GET map. The structure provided by GETs allows edge

detection to be performed faster. A MGET graph representation is proposed for coding motion content, in that a license plate is a sub-MGET-graph (SMG) which satisfies license plate model. The SMG representing the license plate is identified by perceptually grouping plate shape in the MGET graph. Experiments show that the method is able to effectively segment license plates from the video sequences. Whether the video clips were taken during the day or at night did not have a large impact on the program results. The application correctly identified 66 of these license plates, giving an accuracy of approximately 98.5%.

Li-Qun Xu [12], This paper discusses an efficient and effective approach to simultaneously identifying and tracking two free moving and gesturing hands, together with the provision of accurate segmentation masks of the hands regions. The skin color model in YUV color space is created and adapted to characterize the likely appearance of hands. The initialization is performed on the first two frames of a video stream, and the probabilistic skin color map and motion mask information are used, resulting in an accurate segmentation of the two initial hands. The ensued tracking uses a set of Kalman filters to predict the centroid, orientation, and sizes of candidate hand objects, from which refinements of hands locations are made. It focuses mainly on the issue of simultaneous tracking and accurate segmentation of two gesturing hands, which are required for effective occlusion handling for 3D dynamic scene analysis and novel view synthesis in a video conferencing scenario

Beiji Zou et al. [13], This paper presents an object tracking method in diving video sequences by particle filter with multiple motion models. In this method, video paragraphing based on Hough transform technique and the knowledge of critical frame is proposed to divide a whole diving video sequences into several sub-sequences and construct a particle motion model for each sub-sequences. The particles are predicted by multiple motion models to adjust a athlete motions in diving video sequences. The experimental results have demonstrated the efficiency of the object tracking method.

Harini Veeraraghavan et al. [14], This paper present a computer vision-based approach for event detection and data collection at traffic intersections. Specifically, we make the following two contributions: (i) a robust tracking algorithm for targets through combination of multiple cues and multiple motion models, and (ii) a simple event detection system using the results of a switching Kalman filter in combination with some simple rules.

Yumi Iwashita et al. [15], proposed a algorithm of a new high speed execution technique of the level set method named the Fast Level Set Method (FLSM). The efficiency of the FLSM was verified through 2D real-time tracking of moving objects in video images and 3D real-time tracking of multiple human bodies in stereo range images.

Min-Cheng Pan et al. [16], In this paper, with this panoramic imaging system, propose a novel approach for motion tracking of multi-objects mainly based on the idea of tracking windows. With the low complexity of the algorithm and the use of image sampling, experimental results are presented with a sequence of frames to demonstrate the superior performance of the proposed algorithm.

Yankun Wei and Wael Badawy [17], This paper presents a mesh-based motion-tracking model with embedded residual image that improves the quality of predicted frames by the use of residual coding. Mesh-based motion tracking is a fast-emerging model for inter-frame video compression. The presented method reduces the rate of inter frame update. The use of residual image model in mesh based motion tracking is suitable for maintaining high quality P-frames at a low cost.

Hao Jiang and Mark S. Drew [18], This paper present a modified snake model for the problem of general video object tracking. It introduce a new external force into the snake equation based on the predictive contour such that the active contour is attracted to a shape similar to the one in the previous video frame. New methods of contour prediction and contour smoothing are presented. The proposed methods can deal with the problem of an object's Stopping movement temporarily and can also avoid the problem of the snake tracking into the object interior. Experimental results show that the proposed method exhibits increased robustness over a traditional snake algorithm and works well for general video object tracking.

Jiingecin Lim et al. [19], this paper improved video object-tracking algorithm using a motion re-estimation scheme. Experimental results improved a semi-automatic algorithm for semantic video object tracking by introducing object segmentation and motion re-estimation (segment selection) procedures. The

proposed algorithm is composed of two steps: object projection and boundary refinement. Furthermore a simple segmentation method is proposed, that is useful in a video sequence using temporal correlation between adjacent frames to reduce the computational complexity. The simulation results show that the proposed algorithm provides good results even though a video object has complicated movement.

#### IV. CONCLUSIONS AND FUTURE WORK

We have presented the survey of object tracking methods and all categories of moving objects that is object representation, object tracking and object detection from any type of video. This will help us to significantly improve and facilitate the performance of certain computer vision tasks, such as tracking, video surveillance, motion-based recognition, video indexing, human-computer interaction, traffic monitoring, and vehicle navigation. Our future work will focus on:

1) Movement detection and capable of finding the objects which are in motion in every frame with respect to the previous frame.

Till now in many proposed works, we have been plotting the movement of objects through videos but none of them is capable of plotting graph of the moving object if the background is changed at any instant.

2) The coordinates of tracked video in real time can import to any other software to work with the results calculated by our algorithm.

3) Our algorithm is capable of tracking the objects and making a 3d graph in mesh based scope in between any number of fps (frames per second) video.

#### REFERENCES

- [1] Pengwei LIU, Huiyuan WANG et al., Motion Compensation Based Detecting and Tracking Targets in Dynamic Scene, *IEEE*, 2010, 703-706.
- [2] Sajjad Torkan, Alireza Behrad, A New Contour Based Tracking Algorithm Using Improved Greedy Snake, *IEEE*, 2010.
- [3] Baiyang Liu, Lin Yang et al., An Adaptive Tracking Algorithm Of Lung Tumors In Fluoroscopy Using Online Learned Collaborative Trackers, *IEEE*, 2010, 209-212.
- [4] Alexander Toshev, Ameesh Makadia, Kostas Daniilidis, Shape-based Object Recognition in Videos Using 3D Synthetic Object Models, *IEEE*, 2009, 288-295.
- [5] Ming-Yu Shih, Yao-Jen Chang, Bwo-Chau Fu, and Ching-Chun Huang, Motion-based Background Modeling for Moving Object Detection on Moving Platforms, *IEEE*, 2007, 1178-1182.
- [6] Mark Ritch, Nishan Canagarajah, Motion-Based Video Object Tracking In The Compressed Domain, *IEEE*, 2007, 301-304.
- [7] Alper Yilmaz, Omar Javed, Mubarak Shah, Object Tracking: A Survey, *ACM Computing Surveys*, 38(4), 2006.
- [8] Rajan Sehgal, *Video Image Enhancement and Object Tracking*, A Thesis, Thapar Institute of Engineering and Technology, Patiala, ME, 2006.
- [9] Minglun Gong, A GPU-based Algorithm for Estimating 3D Geometry and Motion in Near Real-time, *IEEE*, 2006.
- [10] Huiqiong Chen, Derek Rivait and Qigang Gao, Real-Time License Plate Identification by Perceptual Shape Grouping and Tracking, *IEEE*, 2006, 1352-1357.
- [11] Mohammed Sayed and Wael Badawy, A novel motion estimation method for mesh- based video motion tracking, *IEEE*, 2004, 337-340.
- [12] Li-Qun Xu, Simultaneous Tracking And Segmentation Of Two Free Moving Hands In A Video Conferencing Scenario, *IEEE*, 2003, 49-52.
- [13] Beiji Zou, Xiaoning Peng and Liqin Han, Particle Filter With Multiple Motion Models For Object Tracking In Diving Video Sequences, *IEEE*, 2008, 224-228.
- [14] Harini Veeraraghavan, Paul Schrater and Nikolaos Papanikolopoulos, Switching Kalman Filter-Based Approach for Tracking and Event Detection at Traffic Intersections, *IEEE*, 2005, 1167-1172.
- [15] Yumi Iwashita, Ryo Kurazume et al., Fast Implementation of Level Set Method and Its Real-time Applications, *IEEE*, 2004, 6302-6307.
- [16] Min-Cheng Pan, Yau-Liang Tsai and Ying-Jieh Huang, Motion Tracking of Multi-objects for Access Surveillance Based on the Panoramic Imaging System, *IEEE*, 2003, 208-215.
- [17] Yankun Wei and Wael Badawy, On Reducing The Inter Frame rate: An Embedded Residual Image Model For Mesh-Based Motion Tracking, *IEEE*, 2002, 677-680.
- [18] Hao Jiang and Mark S. Drew, A Predictive Contour Inertia Snake Model For General Video Tracking, *IEEE*, 2002, 413-416.
- [19] Jiingecin Lim, Hong K. Cho, and Jong Beom Rat, An Improved Video Object Tracking Algorithm Based On Motion Re-estimation, *IEEE*, 2000, 339-342.