# A SURVEY ON MOVING OBJECT TRACKING IN VIDEO

Barga Deori[1] and Dalton Meitei Thounaojam[2]

[1] Computer Science and Engineering, National Institute of Technology, Silchar, India
[2]Computer Science and Engineering, National Institute of Technology, Silchar, India

## ABSTRACT

*The ongoing research on object tracking in video sequences has attracted many researchers. Detecting the objects in the video and tracking its motion to identify its characteristics has been emerging as a demanding research area in the domain of image processing and computer vision. This paper proposes a literature review on the state of the art tracking methods, categorize them into different categories, and then identify useful tracking methods. Most of the methods include object segmentation using background subtraction. The tracking strategies use different methodologies like Mean-shift, Kalman filter, Particle filter etc. The performance of the tracking methods vary with respect to background information. In this survey, we have discussed the feature descriptors that are used in tracking to describe the appearance of objects which are being tracked as well as object detection techniques. In this survey, we have classified the tracking methods into three groups, and a providing a detailed description of representative methods in each group, and find out their positive and negative aspects.*

## KEYWORDS

*Feature Descriptor, background modeling, Gaussian Mixture model, Kalman filter, Particle filer, Mean*

## 1. INTRODUCTION

Tracking objects in video sequences of surveillance camera is nowadays a demanding application. Tracking objects is much more challenging in video sequences to improve recognition and tracking performances. There are many existing methods of object tracking but all has some drawbacks. Some of the existing models for object tracking are contour-based models, region-based models and feature point-based models.

A. Contour-based object tracking model

Active contour model is used for finding object outline from an image [1]. In the contour-based tracking algorithm, the objects are tracked by considering their outlines as boundary contours. Thereafter these contours are updated dyna mically in successive frames. The discrete version of this approach is represented in active contour model. The discrete version of this approach takes the advantage of the point distribution model to limit the shape. However, this algorithm is highly sensitive to the initialization of tracking, making it difficult to start tracking automatically.

B. Region-based object tracking model

The region based object model bases it's tracking of objects on the color distribution of the tracked object [2, 3]. It represents the object based on the color. Hence, it is computationally efficient. However, its efficiency is degraded when several objects move together in the image sequences. It is not possible to achieve accurate tracking when multiple objects move due to occlusion. Also, in the absence of any object shape information, the object tracking is largely

DOI : 10.5121/ijit.2014.3304 

dependent on the background model used in the extraction of the object outlines.

C. Feature point based tracking algorithm

In Feature point based model feature points is used to describe the objects [4, 5]. There are three basic steps in feature point based tracking algorithm. The first step is to recognize and track the object by extracting elements. The second step is to cluster them into higher level features. The last step is to match these extracted features between images in successive frames. Feature extraction and feature correspondence are the important steps of feature based object tracking. The challenging problem in feature point based tracking is feature correspondence because a feature point in one image may have many similar points in another image, and hence results in feature correspondence ambiguity.

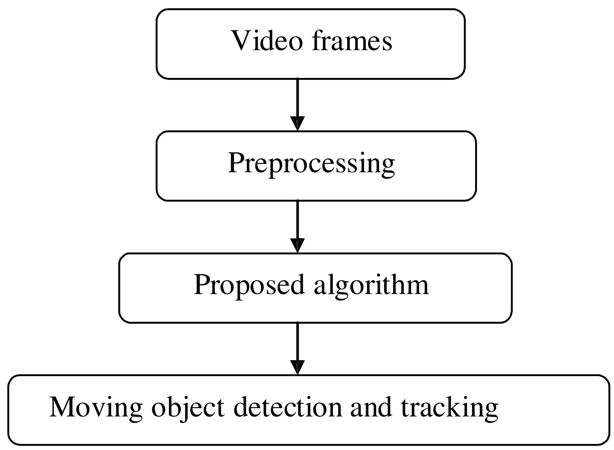


Figure 1: A typical object tracking system in video

## 2. LITERATURE SURVEY

### A. CONTOUR BASED OBJECT TRACKING:

Xu and Ahuja [6] proposed a contour based object tracking algorithm to track object contours in video sequences. In their algorithm, they segmented the active contour using the graph-cut image segmentation method. The resulting contour of the previous frame is taken as initialization in each frame. New object contour is found out with the help of intensity information of current frame and difference of current frame and the previous frame.

Dokladal et al.[7] the proposed approach is active contour based object tracking. For the driver's-face tracking problem they used the combination of feature-weighted gradient and contours of the object. In the segmentation step they computed the gradient of an image. They proposed a gradient-based attraction field for object tracking

Chen[8] models an active contour based object tracking by Neural Fuzzy network.Contour-based model is used to extract object's feature vector. For training and recognizing moving objects their approach uses the self-constructing neuralfuzzy inference network. In this paper, they have taken the histograms of the silhouette of human body in horizontal and vertical projection and then transform it by Discrete Fourier Transform (DFT).

Chen [9] proposed object tracking method consisting of two-stages. Firstly the kernel-based method is used to locate the object in complex environment like partial occlusions, clutter, etc. To improve the tracking result they again used contour based method and tracked the object contour precisely after the target localization. In the target localization step with the of Kalman filter and the Bhattacharyya coefficientthe initial targetposition is predicted and evaluated

Zhou et al.[10] uses the integration of color feature and contour information in the particle filter based multi-hypothesis tracking algorithm. For the contour detection they have used sobel operator and the shape similarity is evaluated between the observing position and the sample position by corresponding points matching in the two contour images.

Ling et al.[11] given an object tracking approach based on contours. The object rough location is found though multi-feature fusion strategy. For accurate and robust object contour tracking, they have extracted the contours with the help of region-based object contour extraction. In their model the object rough location is obtained by color histogram and Harris corner features fusion method. In the particle filer method they have used the Harris corner feature fusion method. Their model of region-based temporal differencing is applied in object contour detection step, and the resultant is the rough location tracking result.

Hu et al.[12] proposed an effective framework for tracking object contours. Their proposed framework integrated different model such as tracking initialization algorithm, color-based contour evolution algorithm, and adaptive shape-based contour evolution and Markov model-based dynamical shape model.

Optical flow detection is used in automatic and fast tracking initialization algorithm. In color-based contour evolution algorithm the correlations between values of neighboring pixels for posterior probability estimation is measured using Markov random field (MRF) theory and the correlations are incorporated into the estimation of the posterior probability of segmentation. Their adaptive shape-based contour evolution algorithm combines the color feature alone and the shape priors to obtain the final contour. A new incremental PCA technique is applied to update the shape model, making the shape model updating flexible. In the Markov model-based dynamical shape model, the dominant set clustering is used to obtain the typical shape modes of a periodic motion.

Rajabi and Nahvi [13] proposed a modified contour-based multiple object tracking algorithm using point processing. This approach has the advantage of multiple objects tracking. Their system can detect and track the peoples in indoor environments videos. In their method they have used Gaussian mixture model (GMM) based background modeling for background estimation.

**B. FEATURE BASED OBJECT TRACKING:**

Li et al. [14] proposed a corner feature based object tracking method using Adaptive Kalman Filter. To represent moving object corner feature are firstly used. Then, the number of corner point variation across consecutive frames to is used to automatically adjust the estimate parameters of Kalman Filter.

Xue et al. [15] uses the discriminative features which are chosen by object/background separation, using a voting strategy. With the help of discriminative features they presented an improved mean-shift algorithm for object tracking.

Yang et al. [16] proposes a object tracking framework for forward-looking infrared (FLIR) imagery based on mean shift algorithm and feature matching. In Feature matching step they used Harris detector to extract the feature points of template object and candidate area. Moreover they have developed an improved Hausdorff distance to measure the similarity of the feature points.

Aibin et al. [17] puts forward a new self-adaptive tracking algorithm views based on target center location and NMI feature. The normalized moment of inertia (NMI) features are combined to locate the center of tracking object in real-time. Mean shift algorithm is here for tracking the object.

Rahman et al.[19] presented an improved tracking method which can track both single object and multiple objects in video sequences where the object movement may be fast or slow. The proposed method is based on background subtraction and feature matching of SIFT features. With the help of background subtraction Object is detected. Matching of motion features and SIFT features helps in detection and tracking of an object.

Fazliet.al [20] proposed a new framework for object tracking combining sift feature and combination of color features and particle filter. SIFT features are used for target representation and localization. Local feature vector is obtained by the transformation of an image. Each of the feature vectors is invariant to image scaling, translation and rotation and illumination changes. The particle filter (PF) is used to find an approximation of the solution to the sequential estimation.

Bai [21] presented a novel object tracking algorithm based on Mean Shift and on-line feature selection. In a 4-D state space, the target object is defined. Feature space is created depending on the color pixel values in R, G and B channels. The best feature space is selected during the tracking which can distinguish objects and background scenes most. In their algorithm, state estimation of the tracking objects is done with the help of Kalman filter**.**

Miao et. al[22] proposed a new robust feature-based tracking method via online boosting by applying adaptive classifiers to match the detected key points in consecutive frames. The proposed approach shows that by integrating the robust local feature and the adaptive online boosting algorithm can help cater to changes between successive frames.

Fan et al.[23] presented a robust object tracking for processing images on mobile devices in real-time. They employ a holistic haar-like feature matching method to track objects of interests. With the help of online feature updating scheme, robustness was achieved in their method. A feature detection method is integrated with color filtering is used to recover tracking.

Kim et al.[25] proposed an algorithm combining background information based motion detection, feature extraction and block matching. In their method a set of features called shape control points (SCPs) are generated by detecting edges in the neighboring four directions. They have reduced the weakness of block matching algorithm with the help of an adaptive background generation method.

Fan [26] proposed a robust tracking method. During tracking they stored the representative object appearances as candidate templates and to match new frames the best template is selected. This procedure of template adding and switching via online strategy keeps update with new object appearances. They have shown that feature-based methods can be extended to non-planar objects or objects undergoing large pose changes.

Alvarez and Regazzoni [27] extended their feature based method of object tracking by using sparse shape points. The possible data association events are sampled with the particle filter to Also, the particle filter helps in estimating the global position and object velocity.

Biresaw et al.[28] developed a feature point tracker. To improve the performance of the tracker they have used time reversed back tracking evaluation criteria together with Partial Least Square regression.

Hossain et al.[29] proposed a multi-Part SIFT feature based rotating object tracking object tracking observation model. The reference and target object are represented to extract the potential key points for measurement the similarity. They have used the Particle filter for solving state space estimation when the state equation is non-linear and the posterior density is non-Gaussian.

Shen et al.[30] adopted the particle filter for tracking, which is useful for non-linear and non-Gaussian problems. To find the posterior probability of particle filter they have used Bhattacharyya distance of object and the predicted position of the object obtained by the particle filter. The posterior probability is used to update the state of the filter. Their experiment proved that HSV is the optimal color space for scale variation, occlusion, and illumination change.

Mahendran et al.[31] proposed a new tracking framework which uses Distance Metric Learning (DML) in combination with Nearest Neighbor (NN) classification for object tracking. In order to detect the object they used canny edge detector Using Nearest Neighbor classifier it is able to distinguish the object from other Objects and subtract the background from the frame using the Nearest Neighbor (NN) algorithm. The Nearest Neighbor algorithm uses the distance between the object and the background to subtract it. Then using a blob detector the object is identified on the basis of the skin color. For the identified object a bounding box is built. By applying Distance Metric Learning (DML) algorithm they tracked the object. For each frame the process was applied to track the object in real time.

Liu et al.[32] proposes an improved Markov chain Monte Carlo(MCMC) named optical flow MCMC(OF-MCMC) sampling algorithm for vehicle tracking. To get the moving direction of the vehicle in initial frames they used the optical flow method, which can solve the problem of scale change and the moving object speed is obtained by autoregressive motion model. To deal with vehicle tracking in low resolution of the video data and to get better tracking results they have generated more accurate feature template with different weighted features.

## C. REGION BASED OBJECT TRACKING:

Xu et al.[33] presented a new method for supervised object segmentation in video sequence. In the proposed method the user input object outline is considered as video object. In moving object tracking, the model incorporated the object's region segmentation and the motion estimation. Active contour model is also employed for contour fine-tuning.

Gu and Lee [34] introduced video object tracking system using backward region based classification. Their system consists of five steps, pre-processing of region, region extraction, motion estimation based on region, region classification and post-processing of the region. Semantic video object boundary is found using a combination of morphological segmentation tool and human assistance. Motion estimation, semantic video object compensation and I-frames boundary information is taken to find out other video objects in the remaining frames.

The object tracking algorithm proposed by Hariharakrishnan and Schonfeld [36] avoids segmentation only for the reason that object partition is initialized in the initial frame. Tracking is done by object boundary prediction using block motion vectors and then updating of object contour by occlusions/disocclusion detection method. For estimating motion between frames they used an adaptive block-based approach. The modification of disocclusion detection algorithm helps in developing occlusion detection algorithm by considering duality principle.

Andrade et al.[37] introduced a novel technique with the help of region derived descriptors for segmentation and tracking. The homogeneous regions of an image are obtained by partitioning the image into a series. Thus, the problem of object extraction changes from pixel based to database analysis.

Wei et al.[38] proposed an object extraction scheme mainly consists of two trackers. Using Adaboost-based global color feature selection the pixel-wise tracker extracts an object. To regionalize each frame *K*-means clustering is performed by the region-wise tracker at the beginning. Using a bidirectional labeling scheme region tracking is achieved.

Kim and Sim [39] proposed a region-based tracking method for the detection of multiple moving objects which uses a differential image. A method of background image update is applied to ensure accurate object detection in unconstrained environment. They have applied the particle filter which provides a robust object tracking framework under complex conditions and greatly improved estimation accuracy for complicated tracking problems.

Khraief et al.[40] presented algorithm for detecting and tracking moving objects using automatic initialization based on background modeling. Their proposed region competition level-set method was used for motion detection and tracking based on the statistical information of image intensity within each subset instead of searching geometrical boundaries. Before going to object segmentation and tracking background modeling is done.

Varas and Marques [41] presented a region-based particle filter for generic object tracking and segmentation. Their algorithm combines color based particle filter and region based particle filter. The algorithm tracked objects in a reliable manner and also provides an accurate segmentation of the target during the sequence. The particle filters uses multiple hypotheses for tracking objects.

Wu et al.[42] developed a robust 3D tracking model which is capable of extract object independent motion trajectory under uncontrolled environment. They have designed two novel algorithms, including a motion-based segmentation and a region-based Mean-shift tracking approach. A Kalman filter is applied to fuse their tracking results of the two algorithms.

## 3. FEATURE DESCRIPTORS

In video object tracking, selection of the right features plays important role. To clearly distinguish the objects in the feature space we need find the object visual feature uniqueness.

A. **Color features:** To increase the discriminative power of intensity based descriptors color feature descriptors are used [43]. Two physical factors primarily influenced the apparent color of an object- 1) the spectral power distribution of the illuminant and 2) object's surface reflectance property. To describe the color information of an object RGB color space is usually used. But RGB color space is not a perceptually uniform color space. Other color space like *L*a*b*and*

*L*u*v** are perceptually uniform. However the HSV (Hue, Saturation and Value) is an approximately uniform color space.

There is no efficient color space which can define the features of an object. So color descriptors in recent studies can be classified into novel histogram-based color descriptors and SIFT-based [44] color descriptors.

In HSV color space, hue becomes unstable near the grey axis. To prove that the certainty of hue is inversely related to saturation an error propagation analysis is applied to the hue transformation. Therefore, the hue histogram is made more robust by weighing each sample of the hue by its saturation. Therefore with respect to intensity of light, HSV color model is scale-invariant as well as shift-invariant.

For the detection and extraction of local feature descriptors a technique called Scale Invariant Feature Transform (SIFT) is used. In SIFT descriptor the intensity channel is a combination of R, G and B channels. Therefore SIFT descriptor is variant to light color changes,

B. **Gradient features:** Gradient features are important in human detection in video sequences. To represent objects like human body, shape/contour of the human body is used in gradient based methods.

C. **Edges features**: The change in intensities of an image is strongly related to object boundaries because after just after the object boundary the intensity instantly changes. To identify the instant change edge detection techniques are used. Compared to color features, edge features illumination changes are less sensitive. Canny Edge detector is mostly used in finding the edges of an object because of it is optimal. Roberts operator, Sobel operator and Prewitt operator are also used for finding the edges.

D. **Texture features**: In Comparison to color features and edge features, a processing step is required to generate the descriptors for the texture features. Local Binary Patterns (LBP) texture feature are known as one of the efficient features. The LBP are gone through an analysis operator is defined as a grayscale invariant texture measure, derived from a general definition of texture in a local neighbor-hood. The most important property of the LBP operator is its tolerance against illumination changes.

E. **Optical flow**: The translation of each pixel in a region can be found out by a dense field of displacement vectors defined as optical flow. Brightness constraint is taken as a measure while computing optical flow, assuming that brightness of corresponding pixels is constant in consecutive frames. Optical flow feature is mostly used in motion-based object segmentation and tracking applications. Furthermore it is also used in video segmentation algorithms.

F. **Spatio-temporal features:** In recent times local spatio-temporal features are mostly used. These features provide a visual representation for recognition of actions and visual object detection [45]. Salient and motion patterns characteristics in video are captured by local spatio-temporal features. These features provide relative representation of events independently. While presenting events the spatio-temporal shifts and scales of events, background clutter and multiple motions in the scene are considered. To show the low level presentation of an object such as pedestrian space-time contours are used. To covert a one-dimensional contour into three-dimensional space a 3D distance transform is used.

G. **Multiple features fusion**: The multi-feature fusion scheme has achieved high boosting performance or robustness, in the field of computer vision, multimedia and audio–visual speech processing, etc [45].

H. **Biological features:** Biological features are important in describing the biological characteristics of human. Attention Regions (ARs) and Biologically Inspired Model (EBIM) features the recent used biological features. Humans biological vision mechanism can be described by these biological and hence to achieve robust recognition.

# 4. OBJECT DETECTION

## 1. Segmentation Based

Segmentation based algorithms are used to segment the image frame into segments to find out the objects of interest. Criteria for good partition and efficient partitioning method plays important role in segmentation algorithms. Later on the segmented objects are considered for tracking.

**A.Graph cut:** In graph cut method the input image is considered as a graph. The segmentation of the objects in the image is considered as the graph partitioning problem. For a graph G (image), the vertices (i.e. pixels), $\mathbf{V} = \{u, v, \ldots..\}$, are partitioned into $N$ disjoint sub-graphs (regions),$A_i$, $\cup_{i=1}^{N} A_i = \mathbf{V}$, $A_i \cap A_j$ =null , $i \neq j$ , by pruning the weighted edges of the graph. Based on the similarity of color, brightness and texture, weight between the nodes is computed. The minimum cut criterion for partitioning an image proposed by Wu and Leahy uses color similarity for weight calculation but their method suffers from over segmentation.

Yi and Moon [46] considered graph cut image segmentation as pixel labeling problems. The label of the foreground object (s-node) is set to be 1 and the background (t-node) is set to be 0. By minimizing the energy-function with the help of minimum graph cut the process of pixel labeling can be done.

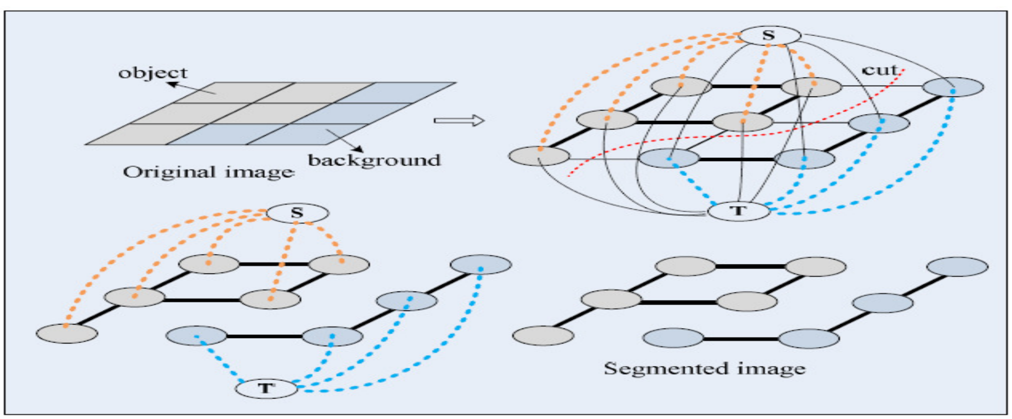


Figure 1: Illustration of graph cut for image segmentation [46]

Shi and Malik [47] propose the *normalized cut* to overcome the over segmentation problem. The 'cut 'of their method depends on the sum of weights of the edges in the cut and on the ratio of the total connection weights of nodes in each partition to all nodes of the graph. For image-based segmentation, the product of the spatial proximity and color similarity defines the weights between the nodes.

**B. Mean-shift clustering:** Mean shift clustering is used to find the clusters of image pixels in the image frame. Comaniciu and Meer [48] used the Mean-shift clustering for the image segmentation problem to find clusters in the joint spatial and color space, $[l, u, v, x, y]$, where $[l$ , *u*, *v*] denotes the color and [*x*, *y*] is the spatial location. For an input image, the algorithm is initialized by randomly choosing a large number of clusters centers from the data. In the next step each of the cluster centers is moved to the mean of the data. The mean of the data is lying inside the multi-dimensional ellipsoid. The multi-dimensional ellipsoid is centered on the cluster center. *Mean-shift vector* is a vector which is defined by the old and the new cluster centers.

**Active Contours:** The boundary of an object can be defined as contours. In active contour framework, a closed contour is evolved to the object's boundary so that the contour covers object region. Hence object segmentation is achieved. An energy function governs the evolution of the contour. The energy defines the fitness of the contour to the exact object region. The following energy function defines the contour evolution:

$$E(\Gamma) = \int_0^1 E_{int}(V) + E_{im}(V) + E_{ext}(V)ds$$

Where $s$ is the arc-length of the contour $\Gamma$, $E_{int}$ includes regularization constraints, $E_{im}$ includes appearance-based energy, and $E_{ext}$ specifies additional constraints. $E_{int}$ usually includes a curvature term, first-order ($\nabla V$) or second-order ($\nabla^2 V$) continuity terms to find the shortest contour.

## 2. Background modeling Based Object Detection

**Gaussian Mixture Model**: Knowing the moving object distribution in the first frame of the video sequences, we can localize the object in the next frames by tracking its distribution. Gaussian Mixture Model is a popular technique for modeling dynamic background as it can represent complex distribution of each pixel. But GMM suffers from slow convergence at the starting stage of detecting backgrounds. Also it sometimes leads to false motion detection in complex background.

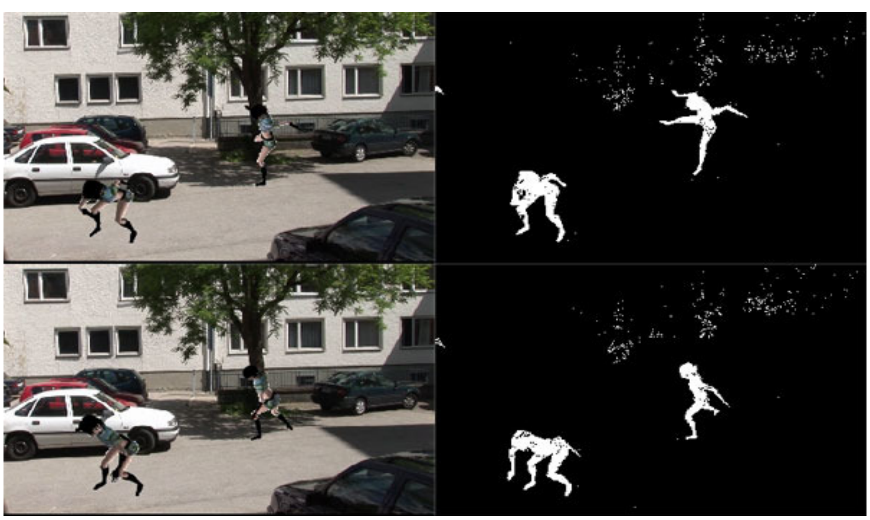

Figure2: Segmentation using GMM based background modeling. Foreground pixels are marked in white [49]

Stauffer and Grimson [50] use a mixture of Gaussians to model the pixel color. In this method, every pixel value of current frame is checked against the existing Gaussian distributions of the background model. Until a matching Gaussian is found the pixel value are checked continuously in the model. The mean and variance of the matched Gaussian is updated when a match is found. If this pixel value does not fit into any one of the Gaussian distributions, the distribution with the least weight is replaced by a new distribution mean as current pixel value, with high variance at initial stage, and a low weight. Classification of pixels is done based on whether matched distribution represents the background process.

B. **Eigen-space Decomposition of Background**: Another approach for background modeling based object detection is Eigen-space decomposition. It is less sensitive to illumination. Here, by projecting the current image to the eigen-space and calculating the difference between the reconstructed and actual images, the foreground objects are detected.

Suppose there are k input frames, $I^i$ : i= 1 · · · k , of size n×m. Now by cascading m rows in each frame one after another a background matrix **B** of size k ×l is formed, where l = (n×m). The eigenvalue decomposition is applied to the covariance of **B**, i.e. $\mathbf{C} = \mathbf{B}^T\mathbf{B}$. The most descriptive η eigenvectors $\mathbf{u}_i$, where i<η< k, represents the background, which encompass all possible illuminations in the field of view (FOV).

C. **Hidden Markov Model**: In recent days Hidden Markov Model is widely used for background subtraction. Corresponding to the events in the environment it represents the intensity variations of a pixel in an image sequence as discrete states. As for example, image pixels can be in the background state, the foreground (car) state, or the shadow state in highway car tracking scenario. Hidden Markov Models (HMM) used by Rittscher et al. [51] classified small blocks of an image into the above three states. Stenger et al. [52] use HMMs for the background subtraction in the context of detecting light on /off events in a room. Those events which are hard to model correctly using unsupervised background modeling approach advantage HMMs are used for training samples.

## 3. Supervised Learning based Background Subtraction

Supervised learning based background subtraction method can also be used for object detection. Supervised learning mechanism helps in learning of different objects view from a set of examples automatically. Supervised learning methods generate a function that maps inputs to desired outputs for a given set of learning examples. Classification problem is the standard formulation of supervised learning, where the learner approximates the behavior of a function. This approximation is done by generating an output in the form of either a continuous value. This process is called *regression*, or a class label, which is called *classification*. Some of the learning approaches are boosting .Viola et al. [53], support vector machines [Papageorgiou et al. [54] etc.

A. **Adaptive Boosting**: Boosting is done by combining many base classifiers to find accurate results. In the first step of training phase of the Adaboost algorithm is an initial distribution of weights over the training set is constructed. The first step of Adaptive boosting is that the boosting mechanism selects the base classifier with least error. The error of the classifier is proportional to the misclassified data weights. Next, the misclassified data weights are increased which are selected by the base classifier. In the next iteration the algorithm selects another classifier that performs better on the misclassified data.

B. **Support Vector Machines**: For a linear system, the available data can be clustered into two classes or groups by finding the maximum marginal hyperplane that separates one class from the other with the help of Support Vector Machines. The distance of

hyperplane and the closest data points helps in defining the margin of the maximized hyperplane
The data points that lie on the hyperplane margin boundary are called the support vectors. For object detection purpose the objects can be included in two classes, object class (positive samples) and the non-object class (negative samples). For applying SVM classifier to a nonlinear system, a kernel trick has to be applied to the input feature vector which is extracted from the input.

### 4. Point Detectors

There may be much interest point in the image frame that we are considering. To find interest points in image frame point detectors are used. The interest points should have expressive texture. In motion, stereo, and tracking problems interest points are mostly used. Detection of interest point is not feasible because of its invariance to illumination change and camera viewpoint.

## 5. OBJECT TRACKING

The importance of an object tracker is that it finds out the motion trajectory of an object as video frames progresses along with time by identifying the object position in every frame of the video. The complete region that is occupied by the object in the image at every time instant can also be found out by the object tracker. The detected objects in frames are being tracked in the subsequent frames. The object detection task and object correspondence establishment task between the instances of the object across frames can be done separately or jointly. In the first scenario, with the help of object detection algorithm possible object regions in every frame are obtained, and objects correspondence across frames is performed by object tracker. In the latter scenario, information obtained from previous frames helps in finding the the object region and correct estimation of correspondence is done jointly by iterative updating of object region and its location.

### STATISTICAL METHODS OF TRACKING:

A. **Kalman filters**: It is a single object state estimation procedure**.** Kalman filter is used as an estimator to predict and correct system state. It helps in studying system dynamics, estimation, analysis, control and processing. It is not only powerful practically but also very well precise theoretically. Kalman filter predicts the states of past, present, and future of an object or variable efficiently. For a linear system Kalman filter finds the correct estimation, with white Gaussian noise. For a linear system the discrete time process can be described by the following equation

1) Process equation

$$x_{k+1} = \mathbf{A}x_k + w_k$$

Where $x_k$ is the system state vector, $w_k$ is Gaussian process noise vector and *A* is the process transition matrix.

2) Measurement equation

$$z_k = Hx_k + \vartheta_k$$

Where $z_k$ is measurement vector, $\vartheta_k$ is the Gaussian measurement noise vector and *H* is the measurement matrix.

The two most important steps of Kalman filter are prediction (time update) step and correction (measurement update) step. A state model is used by the prediction step to predict the new state of the variables.

Prediction Step: The process equation and measurement equation describes a linear model. As $x_{k+1}$ is not measured directly, therefore the information provided by measured $z_k$ is used to update the unknown state $x_{k+1}$. A priori estimate of state $\hat{x}_{k+1}^{-}$ and covariance error $\hat{x}_k$ estimate is obtained for the next time step.

$$\hat{x}_{k+1}^{-} = A_k \hat{x}_k$$

$$P_{k+1}^{-} = A_k P_k A_k^T + Q_k$$

Correction Step: A new observation is incorporated into a priori estimate from the time update in the measurement update equations to obtain an improved posteriori estimate. In the time and measurement update equations, $\hat{x}_k$ is an estimate of the system state vector $x_k$ and $K_k$ is the kalman gain and $P_k$ is the covariance matrix of the state estimation error

$$K_k = P_k^- H_k^T (H_k P_k^- H_k^T + V_k R_k V_k^T)^{-1}$$
$$\hat{x}_k = \hat{x}_k^- + K_k (z_k - H_k \hat{x}_k^-)$$
$$P_k = (I - K_k H_k) P_k^-$$

The Extended Kalman filter (EKF) is a nonlinear version of Kalman Filter. Extended Kalman filter uses Kalman filters to linearize about the current mean and covariance. The result of Extended Kalman Filtering shows faster convergence in the terms of iterations in comparison to traditional methods, though each iteration cost is higher. There might also be some cases where EKF finds better or more robust solutions. In recent days Extended Kalman Filtering (EKF) along with ANN is being used in training.

B. **Particle Filters**: The problem with Kalman filter is that the state variables are normally distributed (Gaussian). So, those state variables that do not follow Gaussian distribution Kalman filter will give poor estimations for those state variables. This problem of the kalman filter can be solved with the help of particle filtering [55].

In particle filtering, the conditional state density $p(X_t \mid Z_t)$ at time $t$ is represented by a set of samples $\{s_t^{(n)} : n = 1, \ldots, N\}$ (particles) with weigh $\pi_t$ (sampling probability). The weights define the importance of a sample, that is, its observation frequency [56]. Particle filter uses a common sampling scheme i.e. *importance sampling* to find new samples. The *importance* sampling scheme can be done in three steps, i.e. selection (selection of random samples), prediction (generate new sample from selected sample) and correction (Weights corresponding to the new sample are computed using the measurements $Z_t$).

C**. Multiobject Data Association and State Estimation**: Kalman filter, extended kalman filter and particle give very good results when the objects are not close to each other. For tracking multiple objects in the video sequences by using Kalman or particle filters, the most likely measurement for a particular moving object needs to be associated with the object's state. This is called the correspondence problem. So for multiple object tracking the most important step we have solve is the correspondence problem before kalman or particle filters are applied. Nearest neighbor approach is the very simplest method to solve the correspondence problem. Data Association algorithms are used to associate the objects state like position, velocity, size with the available filters. Some of the methods to solve the data association are Linear Assignment problem (LAP), Stable Marriage problem (SMP) and Munkers algorithm etc. However the correspondence problem is hard to deal with when the moving objects are close to each other, and then the correspondence shows incorrect results. These filters fail to converge when incorrectly associated measurement occurs. There exist several statistical data association techniques to tackle this problem. Two mostly used techniques for data association in this complex scenario are Joint Probability Data Association Filtering (JPDAF) and Multiple Hypothesis Tracking (MHT).

# 6. CONCLUSIONS

In this article, we present a literature survey of object tracking approaches and also give a brief review of related topics. We divide the tracking approaches into three categories, contour based, region based and feature based approach. In our survey we have seen that moving tracking is a kind of motion tracking. Tracking object motion is done by object detection and then using tracking strategy. In this paper, we survey the various approaches of object tracking, including feature descriptors and object segmentation technique in video frames and various tracking methodologies. We expect that this survey on moving object tracking in video with rich theoretical details of the tracking methods along with bibliography contents will give valuable contribution to research works on object tracking and encourage new research.

# REFERENCES


[1] D. Serby, E. K. Meier, and L. V. Gool, "Probabilistic Object Tracking Using Multiple Features", IEEE Proc. of International Conf on Pattern Recognition Intelligent Transportation Systems, Vol. 6, pp. 43-53, 2004.

[2] L. Li, S. Ranganath, H. Weimin, and K. Sengupta, "Framework for Real-Time Behavior Interpretation From Traffic Video", IEEE Tran. On Intelligen Transportation Systems, , Vol. 6, No. 1, pp. 43-53, 2005.

[3] P. Kumar, H. Weimin, I. U. Gu, and Q. Tian, "Statistical Modeling of Complex Backgrounds for Foreground Object Detection", IEEE Trans. On Image Processing, Vol. 13, No. 11, pp. 43-53, November 2004.

[4] Z Zivkovi, "Improving the selection of feature points for tracking", In Pattern Analysis and Applications, vol.7, no. 2, Copyright Springer-Verlag London Limited, 2004.

[5] J. Lou, T. Tan, W. Hu, H. Yang, and S. J. Maybank, "3D Model-Based Vehicle Tracking", IEEE Trans. on Image Processing, Vol. 14, pp. 1561-1569, October 2005.

[6] N Xu, N Ahuja, 'Object contour tracking using graph cuts based active contours', International Conference on Image Processing, pp. 277-280 vol.3, 2002.

[7] P Dokladal, R Enficiaud, E. Dejnozkova , 'Contour-based object tracking with gradient-based contour attraction field', EEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP '04) ,pp. 17-20, 2004.

[8] T. Chen, 'Object Tracking Based on Active Contour Model by Neural Fuzzy Network', IITA International Conference on Control Automation and Systems Engineering, pp. 570-574, 2009.

[9] Q. Chen, Q. S.Sun , P. A. Heng ,De S. Xia, 'Two-Stage Object Tracking Method Based on Kernel and Active Contour', IEEE Transactions on Circuits and Systems for Video Technology, , pp. 605-609, 2010.

[10] Bing Pu ,Fugen Zhou ,Xiangzhi Bai, 'Particle Filter Based on Color Feature with Contour Information Adaptively Integrated for Object Tracking', Fourth International Symposium on Computational Intelligence and Design, pp.359-362, 2011.

[11] X. Lu, Li Song, Songyu Yu, Nam Ling,' Object Contour Tracking Using Multi-feature Fusion based Particle Filter', IEEE Conference on Industrial Electronics and Applications (ICIEA), pp. 237 – 242, 2012.

[12] W. Hu, Xue Zhou, Wei Li, W. Luo, X. Zhang, and S. Maybank, 'Active Contour -Based Visual Tracking by Integrating Colors, Shapes, and Motions', IEEE Transactions on Image Processing, Volume:22, Issue: 5 ,pp. 1778 – 1792, 2013.

[13] H. Rajabi, M. Nahvi. 'Modified contour-based algorithm for multiple objects tracking and detection', 3th International eConference on Computer and Knowledge Engineering (ICCKE),pp 235 - 239 2013.

[14] Ning Li ,Lu Liu , De Xu, Corner feature based object tracking using Adaptive Kalman Filter , Signal Processing, 2008. ICSP 2008. 9th International Conference ,pp. 1432 – 1435,2008

[15] Chen Xue ; Image Process. Lab., CAS, Changchun ; Ming Zhu ; Ai-hua Chen, A Discriminative Feature-Based Mean-shift Algorithm for Object Tracking , Knowledge Acquisition and Modeling Workshop, 2008. KAM Workshop 2008. IEEE International Symposium ,pp. 217 – 220,2008

[16] Wei Yang ,Junshan Li ; Jing Liu ; Deqin Shi, A Novel Layered Object Tracking Algorithm for Forward-looking Infrared Imagery Based on Mean Shift and Feature Matching, Computer Science and Information Technology, 2009. ICCSIT 2009. 2nd IEEE International Conference,pp. 188 – 191,2009.

[17] Chen Aibin ;Cai Zixing ; Dong Deyi, An Image Tracking Algorithm Based on Object Center location and Image NMI Feature, Fuzzy Systems and Knowledge Discovery, 2009. FSKD '09. Sixth International Conference (Volume:5 ), pp. 359 – 363,2009

[18] Rahman, M.S. ; Saha, A. ; Khanum, S., "Multi-Object Tracking in Video Sequences Based on Background Subtraction and SIFT Feature Matching", Computer Sciences and Convergence Information Technology, 2009. ICCIT'09 Fourth International Conference, pp. 457 – 462,2009

[19] Fazli, S.; Pour, H.M. ; Bouzari, H. , Particle Filter based Object Tracking with Sift and Color Feature, Machine Vision, 2009. ICMV '09. Second International Conference , pp. 89 – 93, 2009

[20] Ke-Jia Bai , A new object tracking algorithm based on Mean Shift in 4-D State Space and On-line Feature Selection, Information and Computing (ICIC), 2010 Third International Conference , pp. 39 – 42 , 2010

[21] Quan Miao ; Guijin Wang ; Xinggang Lin ; Yongming Wang , Scale and rotation invariant feature-based object tracking via modified on-line boosting , Image Processing (ICIP), 2010 17th IEEE International Conference, pp. 3929 – 3932, 2010

[22] Lixin Fan , Riihimaki, M. ; Kunttu, I., A feature-based object tracking approach for realtime image processing on mobile devices Image Processing (ICIP), 2010 17th IEEE International Conference ,pp. 3921 – 3924,2010

[23] T. Kim ; S. Lee ; J. Paik Combined shape and feature-based video analysis and its application to non-rigid object tracking, IET Image Processing, Volume 5, Issue 1,pp. 87 –100,2011

[24] Lixin Fan,'A Feature-based Object Tracking Method Using Online Template Switching and Feature Adaptation , Image and Graphics (ICIG), 2011 Sixth International Conference, pp. 707 – 713, 2011

[25] Alvarez, M.S., Regazzoni, C.S., 'Extended feature-based object tracking in presence of data association uncertainty', International Conference on Advanced Video and Signal-Based Surveillance (AVSS), 2011 ,8th IEEE ,pp. 136 – 141, 2011

[26] Biresaw, T.A, Online failure detection and correction for Bayesian sparse feature- based object tracking, Advanced Video and Signal-Based Surveillance (AVSS), 2011 8th IEEE International Conference , pp. 320 – 324,2011

[27] Chi-min Oh , Chi-Woo Lee , Guee-Sang Lee, 'Multi-Part SIFT feature based particle filter for rotating object tracking', Informatics, Electronics & Vision (ICIEV), 2012 International Conference, pp. 1016 – 1020,2012

[28] Hong-Ying Shen, Shui-Fa Sun ; Xian-Bing Ma ; Yi-Chun Xu ; Bang-Jun Lei, 'Comparative study of color feature for particle filter based object tracking', International Conference on Machine Learning and Cybernetics (ICMLC), pp. 1104 – 1110,2012

[29] Qinkun Xiao ; Dept. of Electron. Inf. Eng., Xi'an Technol. Univ., Xi'an, China ; Xiangjun Liu ; Mina Liu, 'Object Tracking Based on Local Feature Matching Computational Intelligence and Design (ISCID), 2012 Fifth International Symposium on (Volume:1 ), pp. 399 – 402, 2012

[30] Mahendran, S. ; Dept. of Electron. & Commun. Eng., Anna Univ., Chennai, India ; Vaithiyanathan, D. ; Seshasayanan, R., 'Object tracking system based on invariant features ' International Conference on Communications and Signal Processing (ICCSP), pp. 1138 – 1142, 2013

[31] J Ning, W Yu, S Yang, "An active contour tracking method by matching forground and background simultaneously,2013

[32] Ya Liu , Yao Lu ; Qingxuan Shi ; Jianhua Ding, 'Optical Flow Based Urban Roawd Vehicle Tracking, Computational Intelligence and Security (CIS), 2013 9th International Conference on, pp. 391-395, 2013

[33] D. Xu, Jenq-Neng Hwang, Jun Yu, 'An accurate region based object tracking for video sequences', IEEE 3rd Workshop Multimedia Signal Processing, pp. 271-276,1999.

[34] Gu Chuang, Lee Ming-Chieh, 'Semantic video object tracking using region-based classification, International Conference on Image Processing (ICIP'98)- Volume 3, pp.643, 1998.

[35] Hariharakrishnan, K., Schonfeld, D.,' Fast object tracking using adaptive block matching ' IEEE Transactions on Multimedia, Volume:7 , Issue: 5 , pp 853 – 859, 2005.

[36] Andrade, E.L. , Woods, J.C. ; Khan, E. ; Ghanbari, M.,' Region-based analysis and retrieval for tracking of semantic objects and provision of augmented information in interactive sport scenes', IEEE Transactions on Multimedia , Volume:7 , Issue: 6 , Dec. pp. 1084-1096, 2005.

[37] Fan-Tung Wei, Sheng-Ting Chou ; Chia-Wen Lin, 'A region-based object tracking scheme using Adaboost-based feature selection', IEEE International Symposium on Circuits and Systems, Ipp. 2753 – 2756, 2008.

[38] Hyung-Bok Kim, Kwee-Bo Sim, 'A particular object tracking in an environment of multiple moving objects ' International Conference on Control Automation and Systems (ICCAS), pp. 1053 – 1056, 2010.

[39] Khraief, C. , Bourouis, S. ; Hamrouni, K., Unsupervised video objects detection and tracking using region based level-set, International Conference on Multimedia Computing and Systems (ICMCS), pp. 201 – 206, 2012.

[40] Varas, D., Marques, F., 'A region-based particle filter for generic object tracking and segmentation', 19th IEEE International Conference on Image Processing (ICIP), pp. 1333 – 1336, 2012.

[41] Kumar,S. , Narayanan, M.S. , Singhal, P. , Corso, J.J. ; Krovi, V.' Product of tracking experts for visual tracking of surgical tools, IEEE International Conference on Automation Science and Engineering (CASE),pp. 480 – 485, 2013 .

[42] Xingyu Wu ,Xia Mao, Lijiang Chen , Compare, A.,'Combined Motion and Region-Based 3D Tracking in Active Depth Image Sequence ' Green Computing and Communications (GreenCom), 2013 IEEE and Internet of Things (iThings/CPSCom), IEEE International Conference on and IEEE Cyber, Physical and Social Computing, pp .1734 – 1739, 2013.

[43] Alper Yilmaz, Omar Javed, Mubarak Shah, "Object Tracking: A Survey", ACM Computing Surveys, Vol. 38, No. 4,2006

[44] Wei-Bin Yang, Bin Fang ; Yuan-Yan Tang ; Zhao-Wei Shang ; Dong-Hui Li, "Sift features based object tracking with discrete wavelet transform ", International Conference on Wavelet Analysis and Pattern Recognition,ICWAPR,pp. 380-385,2009.

[45] Hanxuan Yang , Ling Shao, Feng Zheng , Liang Wangd, Zhan Song,"Recent advances and trends in visual tracking: A review", Elsevier Neurocomputing 74 (2011) pp. 3823–3831, 2011

[46] Faliu Yi, Inkyu Moon," Image Segmentation: A Survey of Graph-cut Methods",International Conference on Systems and Informatics (ICSAI 2012), 2012.

[47] Shi, J. And Malik, J. 2000. Normalized cuts and image segmentation. IEEE Trans. Patt. Analy. Mach. Intell. 22, 8, pp.888–905.

[48] Comaniciu, D. And Meer, P. 2002. Mean shift: A robust approach toward feature space analysis. IEEE Trans. Patt. Analy. Mach. Intell. 24, 5, pp.603–619.

[49] Vu Pham, Phong Vo, Hung Vu Thanh, Bac Le Hoai, "GPU Implementation of Extended Gaussian Mixture Model for Background Subtraction", IEEE-RIVF 2010 International Conference on Computing and Telecommunication Technologies, Nov. 01-04, 2010

[50] Stauffer, C. And Grimson, W. 2000. Learning patterns of activity using real time tracking. IEEE Trans. Patt. Analy. Mach. Intell. 22, 8, pp.747–767.

[51] Rittscher, J., Kato, J., Joga, S., AND Blake, A. 2000. A probabilistic background model for tracking. In European Conference on Computer Vision (ECCV). Vol. 2. pp.336–350.

[52] Stenger, B., Ramesh, V., Paragios, N., Coetzee, F., AND Burmann, J. 2001. Topology free hidden markov models: Application to background modeling. In IEEE International Conference on Computer Vision (ICCV). pp.294–301.

[53] Viola, P., Jones, M., And Snow, D. 2003. Detecting pedestrians using patterns of motion and appearance. In IEEE International Conference on Computer Vision (ICCV). pp.734–741.

[54] Papageorgiou, C., Oren, M., And Poggio, T. "A general framework for object detection". In IEEE International Conference on Computer Vision (ICCV). pp.555–562, 1998.

[55] Tanizaki, H. , "Non-gaussian state-space modeling of nonstationary time series." J. Amer. Statist. Assoc.82, pp.1032–1063, 1987.

[56] Isard, M. And Blake, A. , “Condensation - conditional density propagation for visual tracking. Int. J. Comput. Vision 29, 1, pp.5–28, 1998.
[57] Djamal Ziani”Feature Selection on Boolean Symbolic Objects”, International Journal of Computational Science and Information Technology (IJCSITY) Vol.1, No.4. , 2013.
[58] Yoonsuk Choi, Ershad Sharifahmadian, Shahram Latifi, “Performance analysis of contourlet-based hyperspectral image fusion methods”, International Journal on Information Theory (IJIT), Vol.2,No.1/2/3/4, October 2013.