Moving Object Tracking Techniques: A Critical Review

Sandeep Kumar Patel
Department of Electronics and Communication Engineering,
Jabalpur Engineering College, Jabalpur (Madhya Pradesh), Pin: 482011, India
Email: sandeeppatel87@sify.com

Agya Mishra
Department of Electronics and Communication Engineering,
Jabalpur Engineering College, Jabalpur (Madhya Pradesh), Pin: 482011, India
Email: agyamishra@gmail.com

Abstract

Moving Object Tracking is one of the challenging problems in the field of computer vision, surveillance, traffic monitoring, video compression etc. The goal of object tracking is to locate a moving object in consecutive video frames. Normally a video tracking system combines three stages of data treating; object extraction, object recognition & tracking, and decisions about activities. This paper presents a critical review of various video object tracking techniques like point tracking, kernel tracking and Silhouette tracking algorithms. Comparison of all the techniques concludes the better approach for its future research.

Keywords: Moving Object Tracking, Object extraction, Object recognition, Occlusion, Daubechies Complex Wavelet Transform (CxWT), Dual-Tree Complex Wavelet Transform (CxWT).

1. Introduction

Object tracking in video sequences is an important topic in the field of computer vision and various research fields. Object tracking aims at deriving the trajectory over time of moving object in video sequences [1]. Object tracking has various applications in the areas like security, surveillance, clinical applications, education, entertainment, biomechanical applications, human robot interaction etc.

There are two key steps in object tracking process:

- Object detection: detection of an object in a given scenario.
- Object tracking: frame by frame tracking of object.

Tracking of objects is very complex in nature due to several problems. Following difficulties come during object tracking:

- The object's shape and size may vary from frame to frame.
- Partial and full object occlusion.
- Presence of noise and blur in video.
- Luminance and intensity changes.
- Object's abrupt motion.

To perform tracking in video sequences, an algorithm analyses sequential video frames and outputs the movement of target between the frames. Many tracking algorithms have been proposed so far. These object tracking methods are classified according to their tracking behaviour.

1.1 *Classification is based on*

- What kind of feature is to be extracted from an image?
- How should we represent the motion, appearance, and shape of the object?

In the following literature review various moving object tracking techniques are broadly classified which provide inclusive descriptions of the illustrative methods in each classification.

1.2 *Features for Object Tracking*

Selecting the accurate features plays a critical part in the tracking. Feature selection is strictly associated to the object representation. For instance, color is used as a feature for Histogram based appearance representations, whereas for contour-based representation object edges are usually used as features. In general, many tracking algorithms use combinations of these features. Common visual features are described below.

1.2.1 *Color:*

All video frame formats are based on different color spaces model. The data of different frame can be stored in dissimilar color spaces ranging from gray scale, RGB, YCbCr and HSB (hue, saturation, value) color spaces. The data is stored in each frame is the brightness in each spectral band. Color images are denoted as red(R), green (G) and blue (B) layers or RGB. Other distinctive color spaces used in the area of object tracking are YcbCr and HSV [7].

1.2.2 *Edges*:

Object boundaries generate strong changes in image intensities. These changes are identified by Edge Detection. A key Property of edges is that they are less sensitive to illumination changes compared to color features. Most of the algorithms use edge as the main agent in object tracking.

1.2.3 *Texture:*

Texture gives the intensity difference of a surface which helps in measuring properties such as smoothness and regularity. On comparing with color space model, texture requires a processing step. On basis of color, the texture features are less sensitive to illumination changes as same as to edge features.

1.2.4 *Optical flow*:

Optical flow is a heavy field of displacement vectors which defines the translation of each pixel in a region. Optical flow is commonly used as a characteristic in motion-based segmentation and tracking applications.

1.3 *Categories of Object tracking*

The goal of an moving object tracking is to create the path for an object above time by finding its position in every single frame of the video. The job of detecting the object and creating correspondence between the object occurrences through frames can either be accomplished separately or jointly. In the first stage, Region of interest (ROI) in each frame is achieved by means of an object detection algorithm, and then tracking corresponds to objects across frames. In final stage, the object region is projected by iteratively updating object location obtained from previous frames [2]. Few Object Tracking methods have been categorized below:

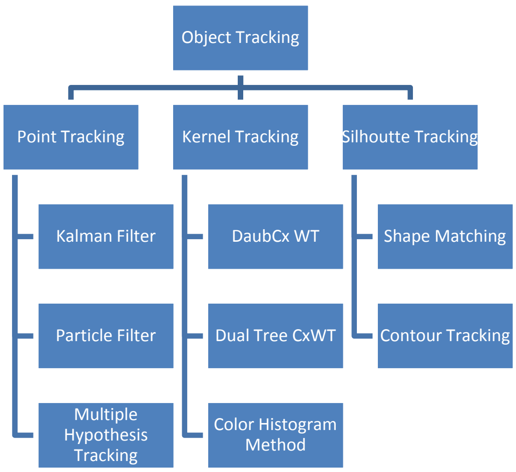

Fig. 1. Object Tracking Methodologies

This report is organized as follows. Section II will give the brief introduction about the various Object tracking methodologies, Section III will provide the performance analysis of the Tracking algorithms used in various tracking methods, section IV will give the Comparison between various Tracking algorithms in tabular form and finally, the conclusion is presented in section V.

2. Introduction of various Tracking Methods

2.1 *Point Tracking Approach*

Point Tracking can be defined as the correspondence of detected objects represented by points across the frames [2]. There are two methods of correspondence methods namely – deterministic and statistical methods [2]. Point Tracking is a difficult problem particularly in the existence of occlusions, false detections of object. Recognition of points can be done simply by thresholding, at of identification of these points [8].

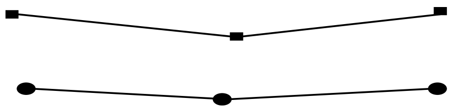

Fig. 2. The lines represent the point correspondence in time [8]

Point Tracking is capable of dealing with tracking very small objects only. Some approaches based on Point Tracking are described below:

2.1.1 *Kalman Filter:*

Kalman filter are based on Optimal Recursive Data Processing Algorithm [2]. Here Gaussian state distribution is assumed. Kalman filtering is composed of two stages, prediction and correction [9]. 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. It is shown below in fig 3.

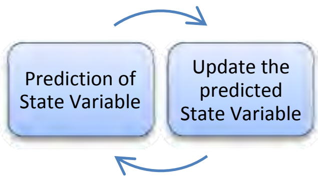

Fig. 3. Basic steps of Kalman Filter [9]

Kalman filter tries to find a balance between predicted values and noisy measurements. The value of the weights is decided by modelling the state equations. Kalman filter track the system in discrete interval of time [9]. The flowchart for the algorithm is drawn below.

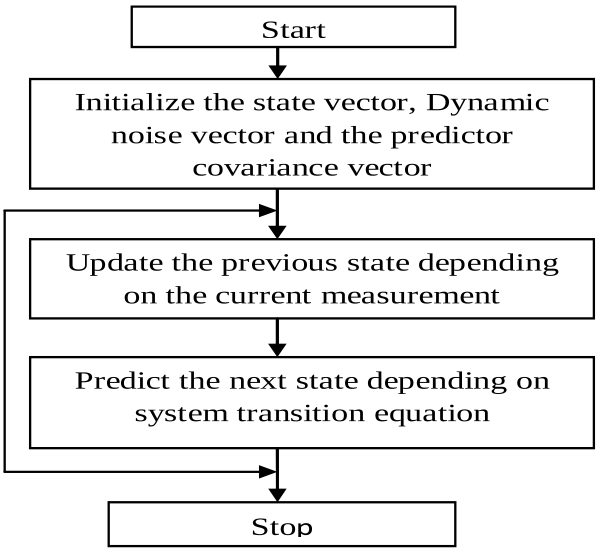

Fig. 4. Algorithm for Kalman Filter [9]

2.1.2 *Particle Filter based on codebook:*

Particle filter is used to track non-linear, non- Gaussian moving objects. As in [5] algorithm based on codebook and particle filter is used to detect moving objects in difficult scenes. 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.

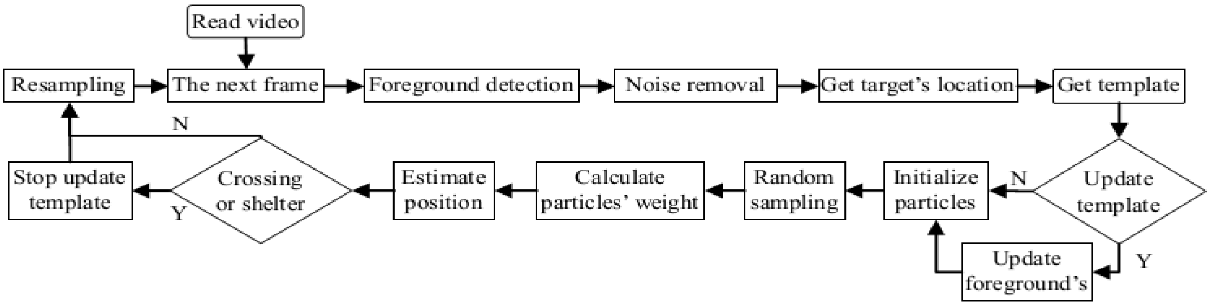

Fig. 5. The improved particle filter algorithm processing [5]

Simulation is performed on the video sequences of size 320×240 pixels, frame rate 25fps. This algorithm helps in solving the problem of particle degradation which was arising in the case of traditional particle filter.

2.1.3 *Multiple Hypothesis Tracking (MHT):*

The MHT algorithm is based on motion correspondence of several frames together. Better results are obtained if correspondence is established observing several frames rather than using only two frames. The MHT algorithm upholds several suggestions for each object at each time. The final track of object is the most likely set of correspondences over time period of its observation [2].

MHT is an iterative algorithm. Iteration begins with a set of existing track hypotheses. Each hypotheses is a crew of disconnect tracks. For each hypothesis, a prediction of object's motion in the succeeding frame is made. The predictions are then compared by calculating a distance measure.

2.2 *Kernel Based Tracking Approach*

Kernel tracking is usually performed by computing the moving object, which is represented by a potential object region, from one frame to another. The object motion is usually in the form of parametric motion such as translation, conformal, affine, etc.

These algorithms appear different in terms of the representation used, the number of objects tracked, and the method used for approximating the object motion [2]. There are several techniques based on representation of object, object features, appearance and shape of the object [10]. Few of the tracking technique based on Kernel tracking approach:

2.2.1 *Dual-Tree Complex Wavelet Transform Technique:*

Real Wavelet Transform suffers from shift variance and poor directionality. As in [3] Object tracking method based on complex wavelet transform is used. Real Filter is used to obtain shift invariance. Two steps are followed; Segmentation and Tracking.

2.2.1.1 *Segmentation Process***:**

Optical flow computation for finding moving object is used. Segmentation algorithm proceeds as follows:

- Take first and tenth frame of video sequence.
- Convert them to grey level image; further determine optical flow using Horn-Schunck method between these two images.
- Thus we find the magnitude square value of the optical flow $|V|^2$.
- Find the mean value of the $|V|^2$ for the first image and compare its value with magnitude square value of the optical flow at each pixel location in the image.
- If $|V|^2$ at any pixel is greater than or equal to the mean value. Then keep its pixel value 1, otherwise assign it 0.

2.2.1.2 *Tracking Using Dual–Tree Complex Wavelet Transform:*

Tracking algorithm can be obtained from [3]. In different video frames centroid of the moving object is calculated.

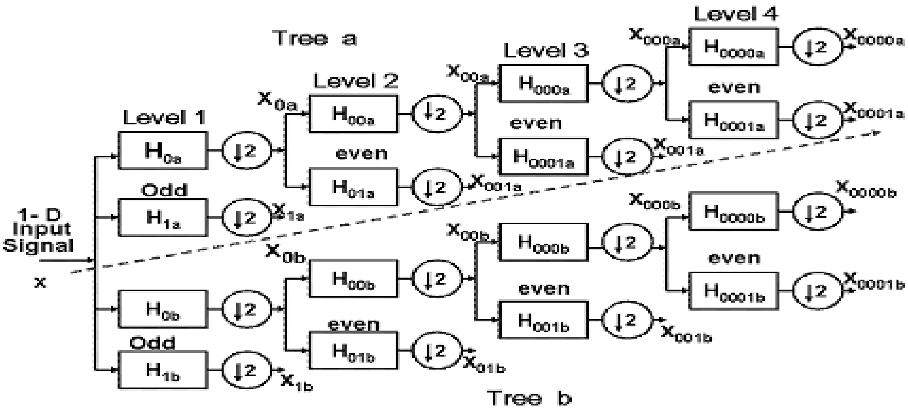

Fig. 6. Dual- Complex Wavelet Transform over 2 levels

2.2.2 *Daubechies Complex Wavelet Transform Technique:*

This technique uses only complex wavelet coefficients. Complex valued filters are used here instead of real valued filter [3] to obtain shift invariance making it a true complex wavelet transform. The use of only one parameter makes tracking accurate and efficient. The scaling equation of Multiresolution theory is given by:

$$\emptyset(x) = 2\sum_k a_k \emptyset(2x - k) \quad (1)$$

Where a_k are the coefficients. The a_k can be real as well as complex valued and $\sum a_k=1$. Daubechies's wavelet bases $\{\Psi_j,k(t)\}$ in one dimension are defined through the above scaling function and multiresolution analysis of L2(R). To provide general solution, Daubechies considered a_k to be real valued only. If we consider a_k and $\emptyset$ both to be complex valued as well, we will get Daubechies complex wavelet transform.

The generating wavelet $\psi(t)$ is given by:

$$\Psi(t) = 2\sum_n (-1)^n \overline{a_{1-n}} \emptyset(2t - n) \quad (2)$$

Here $\psi(t)$ and $\emptyset(t)$ are the same compact support [-N, N+1]. Any function and mother wavelet as:

$$f(t) = \sum_k c_k^{j_0} \emptyset_{j_{0,k}}(t) + \sum_{j=j_0}^{j_{max}-1} d_k^j \Psi_{j,k}(t) \quad (3)$$

Where j_0 is a given resolution level, $\{c_k^{j_0}\}$ and $\{d_k^j\}$ are known as approximation and detailed coefficients.

2.2.3 *Object tracking using Color Histogram:*

Here feature space is chosen to charecterize the target [6]. The reference target model is represented by its probability density function in the feature space. In the subsequent frame, a candidate model is defined at location y and is charecterised by the probability density function *p(y)*. The *pdfs* are estimated from the m – bin histograms. A similarity function $\hat{\rho}(y)$ called as the Bhattacharyya coefficient between $\hat{\mathbf{p}}$and $\hat{\mathbf{q}}$ plays the role of likelihood and its local maxima in image indicate the presence of objects.

The target mode $\hat{\mathbf{q}} = \{\hat{\mathbf{q}}_u\}_{u=1...m}$,

$$\sum_{u=1}^{m} \hat{\mathbf{q}}_u = 1 \quad (4)$$

The candidate model $\hat{\mathbf{p}}(\mathbf{y}) = \{\hat{\mathbf{p}}_u(y)\}_{u=1............m}$,

$$\sum_{u-1}^{m} \hat{\mathbf{p}}_u = 1 \quad (5)$$

The distance between two discrete distributions is defined as

$$d(y) = \sqrt{1 - \rho[\hat{\mathbf{p}}(\mathbf{y}), \hat{\mathbf{q}}]} \quad (6)$$

Where $\rho[\hat{\mathbf{p}}(\mathbf{y}), \hat{\mathbf{q}}]$ is the Bhattacharyya coefficient

$$\hat{\rho}(y) = \rho[\hat{\mathbf{p}}(\mathbf{y}), \hat{\mathbf{q}}] = \sum_{u-1}^{m} \sqrt{\hat{\mathbf{p}}_u(\mathrm{y})\hat{\mathbf{q}}_u} \quad (7)$$

2.3 *Silhouette Based Tracking Approach*

Objects having composite shapes for example, hands, head, and shoulders, are cannot be well defined by geometric shapes. Silhoue0tte based approach will give perfect description of shape of those objects. The aim of the silhouette based tracking is to find the object region by means of an object model. This model verifies the object region in each frame. Model can be represented in the form of color histogram, object edges or contour. We classify silhouette tracking into two categories, namely, shape matching and contour tracking.

2.3.1 *Contour Tracking:*

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 [2].

- *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.

2.3.2 *Shape Matching:*

This approach checks for object model in the existing frame [2]. Shape matching performance is similar to template based tracking in kernel approach. Another approach to Shape 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 [2].

3. Performance Analysis

3.1 *Point Tracking*

3.1.1 *Kalman Filter:*

Here author has analysed that Kalman filtering approach is capable in dealing with noise. It is applicable only for single object, multiple objects [2], [9]. Kalman Filter always gives optimal solution. It is used in vision community tracking.

3.1.2 *Particle Filter based on codebook:*

Here author has mainly emphasized on information loss in the detection and tracking. This method makes the particle filter reflect the objects more accurately and timely. Here average processing time per frame is calculated on the video sequences of 320×240 pixels, frame rate is 25fps and the result was 94ms.

3.1.3 *Multiple Hypotheses Tracking:*

MHT focuses on FOV (field of view) of object entering and object leaving. It can handle occlusion. It can track multiple objects.

3.2 *Kernel Tracking*

3.2.1 *Dual- Tree CxWT based object Tracking:*

Dual –Tree CxWT is an efficient way of implementing an analytic wavelet transform. It has properties like directionality selectivity, Shift invariance and perfect reconstruction. Experiment is performed on video sequence (240×320) and minimum energy difference & corresponding boundary values are calculated. Then centroid corresponding to this boundary value is calculated.

3.2.2 *Daub CxWT based Object Tracking:*

Here Author has used Daubechies complex wavelet coefficients as a feature of object. DaubCxWT helps in preserving the edges making it useful in tracking edge sensitive tracks. This technique has reduced shift sensitivity. DaubCxWT keeps the shape of object intact. Experimental results are performed on video clips of frame size 288×352 and centroid of the object window at each frame is computed.

3.2.3 Color Histogram Technique:

This method is capable of dealing with complex camera motion, partial occlusion of the target, presence of significant clutter, and large variations in target scale and appearance. Minimum Bhattacharyya distance is calculated here. Detection rate is calculated on the basis of (Frame Missed/Total number of frames) and it is about 96.55%, image size is 320×240.

3.3 *Silhouette Tracking*

3.3.1 *Contour Tracking:*

In [2] 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.

3.3.2 *Shape Matching:*

Edge based templates has been used by the author. Here temporal spatial velocity in 4D image per frame is calculated. It can track only single object. Occlusion handling is performed using Hough Technique [2].

4. Comparison between various Tracking Techniques

Comparison of Observations given in all references is discussed here. Qualitative comparison for tracking methodologies is presented in the table drawn below. The Following table concludes that different tracking techniques have been applied for object tracking for different challenging situations.

Table 1.Comparison between different moving object tracking techniques

S.No.	Methodology	Type of Tracking	Algorithms used	¥	Occlusion Handling	Efficiency/ Measurement	Advantages	Limitation
1	Kalman Filter[2],[9]	Point Tracking	Kalman Filtering algorithm	S	No	Efficiency in terms of total time elapsed(in seconds) for processing certain frames	Used to track points in noisy images	State variables are normally distributed (gaussian)
2	MHT(Multiple Hypotheses Tracking) [2]	Point Tracking	MHT algorithm	M	Yes	Distance measure is calculated	Able to deal with entries of new object and exit existing object	Computationally exponential both in time and memory
3	Particle Filter based on Codebook background[2]	Point Tracking	Improved particle filtering algorithm	M	Yes	Get the min. Variance estimate, avg. Processing time/frame is 94ms	Solves the problem of particle degradation of traditional particle filter, the background color interference	---
4	Dual – Tree Complex Wavelet transform [3]	Kernel Tracking	Dual-Tree CxWT algorithm	S	P	Centroid of the moving object bounding box in each frame is calculated	Good directional selectivity and shape matching	It uses real Filter
5	Daub Complex Wavelet transform [4]	Kernel Tracking	Daub CxWT algorithm	S	P	Min. Difference of energy of wavelet coefficients between frames	Reduced phase sensitivity and false tracking of objects, helps in preserving the edges	Object shape and size should not change b/w successive frames
6	Color Histogram[6]	Kernel Tracking	Histogram based algorithm	S	P	Search takes about 500ms and detection rate is 96.5%	Runs very fast, suitable for models having dominant colors	Spatial information of the target is lost, cannot give good performance when an object & its background have similar color
7	Contour Tracking[2]	Silhouette Tracking	Gradient Descent Algorithm	M	F	Region Statics is calculated using grid points	Object Shape is Implicitly modeled	---
8	Shape Matching [2]	Silhouette Tracking	Hough Transform	S	P	TSV(Temporal Spatial Velocity) in 4D(x,y,u,v) image per frame is calculated	Less sensitive to appearance variations	It requires Training

¥: number of objects being tracked; S: single object tracking; M: multiple object tracking;
F: full occlusion handling; P: partial occlusion handling; CxWT: Complex Wavelet transform

In case of point tracking Particle filter using codebook background gives best results in case of object crossing and blocking, judge abnormal behaviour of the object and detected object is tracked effectively. In case of kernel tracking Daub-CxWT uses complex filter which makes it a true Complex Transform while Dual-Tree CxWT uses real filter. In case of Silhouette tracking complete region of an object is required. Silhouette Tracking has most important advantage in handling variety of object shapes and they have the capability to deal with object split and merge. The most important conclusion can be drawn that chance of occlusion must be reduced to greater extent.

5. Conclusion

In this paper, a widespread literature survey on various moving object tracking methods has been presented. Object is tracked mainly on the bases of object extraction, object recognition and tracking and decisions about activities. Comparative analysis is being done for point tracking, kernel and Silhouette tracking algorithms. For instance point tracking involves detection in every frame, while kernel based or contour based tracking requires detection when object first appears in the scene. In case of point tracking particle filter using codebook give best results with respect to average processing time per frame. In kernel tracking approach, various estimating methods are used to find corresponding region to target object. Color histogram technique has good efficiency in terms of frame detection but spatial information is lost, if features like spatial information, texture information etc is included then kernel tracking will give better results. DaubCxWT gives accurate tracking results but object must move with a constant velocity between successive frames, Dual-Tree CxWT provides better directionality and shift invariance and further it can be used to track multiple objects. Silhouette tracking forms the bases on the type of representation which can be motion models or appearance models. This survey highlights the features of algorithm for researchers in the area of moving object tracking.

References

[1] Bovik Al –"*Essential guide to video processing*" Elsevier 2009 .inc.

[2] Yilmaz Alper, Javed Omar and Shah Mubarak, 2006, "*Object Tracking: A survey*" ACM Compt. Surv.38,4, Article 13, December, pp. 1-45.

[3] Singh Roshan, Purwar Ravindra Kumar, Rajpal Navin, "*A better approach for object tracking using Dual-Tree Complex Wavelet Transform*" , International Conference on Image Information Processing (ICIIP 2011)pp. 1-5.

[4] Prakash Om, Khare Manish, Sharma Chandra Mani, Kushwaha Alok K. Singh, "*Moving object tracking in video sequences based on energy of Daubechies Complex Wavelet Transform*", National Conference on Communication Technologies & its impact on Next Generation Computing (CTNGC 2012), pp. 6-10.

[5] Zhong QU, Quingquing ZHANG, Tengfei GAO, "*Moving Object Tracking based on Codebook and Particle Filter*", Internatinal Workshop on Information and Electronics Engineering 2012(IWIEE) Elsevier, pp. 174-178.

[6] Dash Prajna Parimita, Patra Dipti, Behera SatyaRanjan, Behera Subhendu Kumar, Mishra Sudhanshu Kumar, "*Comparative Performance Evaluation of Three Object Tracking Methods*", International Journal of merging Technology and Advanced Engineering(May 2012), pp. 297-302.

[7] Sebastian Patrick, Voon Yap Vooi, Comley Richard, "*Color Space effect on Tracking in video Surveillance*", International Journal On Electrical Engineering and Informatics 2010, vol-2, number-4, pp. 298-312.

[8] Veenman C.J., Hendrics E.A., Reinders M.J.T, "*A fast Robust Point Tracking Algorithm*" IEEE International Conference on Image Processing, volume III (1998).

[9] Banerjee Salil P., Pallipuram Krish "*Multi Person Tracking using Kalman Filter*" December 10, 2008.

[10] Baviskar Swati P., Ujjare Nitin "*Kernel Based Object Tracking using Mean Shift Method Method*", IJCA proceedings on International conference in Computational Intelligence (CICCIA 2012).