International Journal of New Technology and Research (IJNTR)
ISSN:2454-4116, Volume-2, Issue-5, May 2016 Pages 96-100

Moving Vehicle Detection and Tracking Using Modified Mean Shift Method and Kalman Filter and Research

Nikita Rawat, Rohit Raja

Abstract— **The goal of object tracking is segmenting a region of interest from a video scene and keeping track of its motion, positioning and occlusion. The object detection and object classification are preceding steps for tracking an object in sequence of images. Mean shift algorithm is recently widely used in tracking clustering, etc. First phase of the system is to detect the moving objects in the video. Second phase of the system will track the detected object. In this paper, detection of the moving object has been done using simple background subtraction and tracking of single moving object has been done using modified mean shift method and Kalman filter. Further result of both algorithm is compared on basis on time and accuracy.**

Index Terms— **object tracking; kalman filter; mean shift method..**

I. Introduction

Object tracking is an important task within the field of computer vision. 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. There are three key steps in video analysis: detection of interesting moving objects, tracking of such objects from frame to frame, and analysis of object tracks to recognize their behavior. Therefore, the use of object tracking is pertinent in the tasks of motion-based recognition, that is, human identification based on gait, automatic object detection, etc; and in traffic monitoring, that is, real-time gathering of traffic statistics to direct traffic flow. In its simplest form, 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. Tracking objects can be complex due to loss of information caused by projection of the 3D world on a 2D image and noise in images, Mean shift, which was proposed in 1975 by Fukunaga and Hostetler[1], is a nonparametric, iterative procedure that shifts each data to local maximum of density function. In spite of its good properties, it has been ignored until Cheng's paper[2] renews our interest in it. Cheng in [2] revisited mean shift, developing a more general formulation and demonstrating its potential uses in clustering and global optimization. Since then, mean shift has been widely used in object tracking[3-7], image segmentation[8,9], pattern recognition and clustering[10,11], filtering[12], information fusion[13] and etc.

Nikita Rawat, ME (CTA), Research Scholar, SSTC Bhilai, india
Rohit Raja, Assitant Professor, SSTC SSGI Bhalai, India,

Kalman filter is an optimal Recursive Data Processing Algorithm. It consists of the following two phases- (i) prediction and (ii) correction. The first refers to the prediction of the next state using the current set of observations and update the current set of predicted measurements. The second updates the predicted values and gives a much better approximation of the next state. It attempts to achieve a balance between predicted values and noisy measurements. The values of the weights are determined by modeling the state equations.

II. LITERATURE SURVEY

Cheng [19] introduced Mean shift algorithm to the field of computer vision. In his paper he has briefly about Mean shift, Mean Shift is a simple interactive procedure that shifts each data point to the average of data points in its neighborhood is generalized and analyzed in the paper. This generalization makes some k-means like clustering algorithms its special cases. It is shown that mean shift is a mode-seeking process on the surface constructed with a "shadow" kernal. For Gaussian kernels, mean shift is a gradient mapping. Convergence is studied for mean shift iterations. Cluster analysis if treated as a deterministic problem of finding a fixed point of mean shift that characterizes the data. Applications in clustering and Hough transform were demonstrated. Mean shift is also considered as an evolutionary strategy that performs multistate global optimization.

Bradski [20] modified Mean Shift Algorithm developed by Cheng [19], and developed the Continuously Adaptive Mean Shift (CAMSHIFT) algorithm for face tracking. As a first step towards a perceptual user interface, a computer vision color tracking algorithm was developed and applied towards tracking human faces. Computer vision algorithms that are intended to form part of a perceptual user interface must be fast and efficient. They must be able to track in real time yet not absorb a major share of computational resources: other tasks must be able to run while the visual interface is being used. The new algorithm developed here was based on a robust non-parametric technique for climbing density gradients to find the mode (peak) of probability distributions called the mean shift algorithm. In his case, they want to find the mode of a color distribution within a video scene.

Therefore, the mean shift algorithm was modified to deal with dynamically changing color probability distributions derived from video frame sequences. The modified algorithm was called the Continuously Adaptive Mean Shift (CAMSHIFT) algorithm. CAMSHIFT's tracking accuracy was compared against a Polhemus tracker. Tolerance to noise, distracters and performance was studied. CAMSHIFT was then used as a computer interface for controlling commercial computer games and for exploring immersive 3D graphic worlds. Comaniciu and Meer successfully applied mean shift algorithm to image segmentation [21] and object tracking. They have developed a new method for real time tracking of non-rigid objects seen from a moving camera was proposed. The central computational module is based on the mean shift iterations and finds the most probable target position in the current frame. The dissimilarity between the target model (its color distribution) and the target candidates were expressed by a metric derived from the Bhattacharyya coefficient. The theoretical analysis of the approach shown, that it relates to the Bayesian framework while providing a practical, fast and efficient solution. The capability of the tracker to handle in real time partial occlusions, significant clutter, and target scale variations was demonstrated for several image sequences. Comaniciu and Meer [22] then modified their approach and developed a general non-parametric technique for the analysis of a complex multimodal feature space and to delineate arbitrarily shaped clusters. The basic computational module of the technique was an old pattern recognition procedure: the mean shift. For discrete data, they proved the convergence of a recursive mean shift procedure to the nearest stationary point of the underlying density function and, thus, it's utility in detecting the modes of the density. The relation of the mean shift procedure to the Nadaraya-Watson estimator from kernel regression and the robust M-estimators; of location was also established. Algorithms for two low-level vision tasks discontinuity-preserving smoothing and image segmentation were described as applications. In those algorithms, the only user-set parameter was the resolution of the analysis, and either gray-level or color images are accepted as input. Extensive experimental results illustrated their excellent performance.

Comaniciu et. al.[23], developed Vision based tracking, it was a challenging engineering problem is one of the hot research areas in machine vision. At that time Kernel based tracking using Bhattacharya similarity measure was shown to be an efficient technique for non-rigid object tracking through the sequence of images. In their paper they presented a robust and efficient tracking approach for targets having larger motions as compared to their sizes. Their tracking approach was based on calculating the Gaussian pyramids of the images and then applying mean shift algorithm at each pyramid level for tracking the target. Model based tracking often suffers abrupt changes in target model, which is compensated by the model updates of target. This leads to a very efficient arid robust nonparametric tracking algorithm the new method was easily able to track the fast moving targets and is more robust and environment independent as compared to original kernel based object tracking.

Collins R [24], the mean-shift algorithm is an efficient technique for tracking 2D blobs through an image. Although the scale of the mean-shift kernel was a crucial parameter, there was presently no clean mechanism for choosing or updating scale while tracking blobs that are changing in size. He adapted Lindeberg's (1998) theory of feature scale selection based on local maxima of differential scale-space filters to the problem of selecting kernel scale for mean-shift blob tracking. He had shown that a difference of Gaussian (DOG) mean-shift kernel enables efficient tracking of blobs through scale space. Using this kernel requires generalizing the mean-shift algorithm to handle images that contain negative sample weights.

Zivkovic Z. and Krose B [25], the iterative procedure called 'mean-shift' is a simple robust method for finding the position of a local mode (local maximum) of a kernel-based estimate of a density function. A new robust algorithm was developed that presented a natural extension of the 'mean-shift' procedure. The new algorithm simultaneously estimates the position of the local mode and the covariance matrix that describes the approximate shape of the local mode. They applied the new method to develop new 5-degrees of freedom (DOF) color histogram based non-rigid object tracking algorithm.

Kalman filter technique is used to estimate the state of a linear system where state is assumed to be distributed by a Gaussian [18]. In 1960, R.E. Kalman [14] published his famous paper describing a recursive solution to the discrete-data linear filtering problem[1].Object tracking is performed by predicting the object's position from the previous information and verifying the existence of the object at the predicted position.

Secondly, the observed likelihood function and motion model must be learnt by some sample of image sequences before tracking is performed [15]. The Kalman filter is a set of mathematical equations that provides an efficient computational (recursive) means to estimate the state of a process in several aspects: it supports estimations of past, present, and even future states, and it can do the same even when the precise nature of the modelled system is unknown [16-17]. The Kalman filter estimates a process by using a form of feedback control. The filter estimates the process state at some time and then obtains feedback in the form of noisy measurements. The equations for Kalman filters fall in two groups: time update equations and measurement update equations. The time update equations are responsible for projecting forward (in time) the current state and error covariance estimates to obtain the a priori estimate for the next time step. The measurement update equations are responsible for the feedback. That is used for incorporating a new measurement into the a priori estimate to obtain an improved a posteriori estimate. The time update equations can also be thought of as predictor equations, while the measurement update equations can be thought of as corrector equations.

III. METHODOLOGY:

A. Modified Mean Shift Tracking (MMST) Algorithm

1. Initialization: calculate the target model $\hat{q}$ and initialize the position y_0 of the target candidate model in the previous frame.

The probability of the feature u (u=1, 2… m) in the target model is computed as [9].

$$\begin{cases}\hat{q}=\{\hat{q}_u\}_{u=1\cdots m}\\ \hat{q}_u=C\sum_{i=1}^{n}k\left(\|x_i^*\|^2\right)\delta\left[b(x_i^*)-u\right]\end{cases}$$

(1)

Where $\hat{q}$ is the target model, $\hat{q}_u$ is the probability of the u^{th} element of $\hat{q}$, δ is the Kronecker delta function, $b\{X_i^*\}$ associates the pixel X_i^* to the histogram b_{in}, and $k(x)$ is an isotropic kernel profile.

2. Initialize the iteration number $k \leftarrow 0$.

3. Calculate the target candidate model $\hat{p}(y_0)$ in the current frame.

The probability of the feature u in the target candidate model from the candidate region centered at position y is given by

$$\begin{cases}\hat{p}(y)=\{\hat{p}_u(y)\}_{u=1\cdots m}\\ \hat{p}_u(y)=C_h\sum_{i=1}^{n_h}k\left(\left\|\frac{y-x_i}{h}\right\|^2\right)\delta\left[b(x_i)-u\right]\end{cases} \quad (2)$$

Where C is constant

4. Calculate the weight vector $\{w_i\}_{i=1\cdots n}$

$$w_i=\sum_{u=1}^{m}\sqrt{\frac{\hat{q}_u}{\hat{p}_u(y_0)}}\delta[b(x_i)-u]$$

(3)

5. Calculate the new position y_1 of the target candidate model

In the mean shift iteration, the estimated target moves from y to a new position y_1, which is defined as

$$y_1=\frac{\sum_{i=1}^{n_h}x_i w_i g\left(\left\|\frac{y-x_i}{h}\right\|^2\right)}{\sum_{i=1}^{n_h}w_i g\left(\left\|\frac{y-x_i}{h}\right\|^2\right)}$$

(4)

6. Let $d \leftarrow \| y_1\text{-}y_0 \|$, $y_0 \leftarrow y_1$.Set the error threshold ε (default 0.1) and the maximum Iteration number N (default 15).

If $(d < \varepsilon \text{ or } k \geq N)$ (5) Stop and go to step 7;

Otherwise $k \leftarrow k+1$, and go to step 3.

7. Estimate the width, height and orientation from the target candidate model (Cov)

Bhattacharyya coefficient can be used to adjust M_{00} in estimating the target area, denoted by A

$$A=c(\rho)M_{00} \quad (6)$$

Where $c(\rho)$ is a monotonically increasing function with respect to the Bhattacharyya coefficient $\rho(0 \leq \rho \leq 1)$, M_{00} is estimated frame, If λ_1 and λ_2 height of the target that

$$k=\sqrt{A/(\pi\lambda_1\lambda_2)} \quad (7)$$

$$a=\sqrt{\lambda_1 A/(\pi\lambda_2)} \qquad b=\sqrt{\lambda_2 A/(\pi\lambda_1)}$$

(8)

Now the covariance matrix becomes

$$Cov=\begin{bmatrix}u_{11} & u_{12}\\ u_{21} & u_{22}\end{bmatrix}\times\begin{bmatrix}a^2 & 0\\ 0 & b^2\end{bmatrix}\times\begin{bmatrix}u_{11} & u_{12}\\ u_{21} & u_{22}\end{bmatrix}^T$$

(9)

8. Determining the Candidate Region in Next Frame

Once the location, scale and orientation of the target are estimated in the current frame, we need to determine the location of the target candidate region in the next frame. With Eq. (9), we define the following covariance matrix to represent the size of the target candidate region in the next frame

$$Cov_2=U\times\begin{bmatrix}(a+\Delta d)^2 & 0\\ 0 & (b+\Delta d)^2\end{bmatrix}\times U^T \quad (10)$$

where Δd is the increment of the target candidate region in the next frame. The position of the initial target candidate region is defined by the following ellipse region

$$(x-y_1)\times Cov_2^{-1}\times(x-y_1)^T\leq 1$$

(11)

B. Algorithm of Kalman Filter For Object Tracking

The Kalman filter is a set of mathematical equations that provides an efficient computational (recursive) means to estimate the state of a process, in a way that minimizes the mean of the squared error. The filter is very powerful in several aspects: it supports estimations of past, present, and even future states, and it can do so even when the precise nature of the modeled system is unknown[14].

The Kalman filter is the best filter among the subset of all linear filters and the best filter among the set of all filters when the noise processes are Gaussian type[15].

The Kalman filter is essentially a set of mathematical equations that implement a predictor-corrector type estimator that is optimal in the sense that it minimizes the estimated error covariance-when some presumed conditions are met [17].

Tracking of moving object has been done using Kalman filter. Here tracking of any object can be done by providing the frame number from which tracking has to be started. From the selected frame any object can be picked for tracking by setting the position of the mask and then the object can be tracked in subsequent frames.

Following steps have been implemented for tracking a single object.

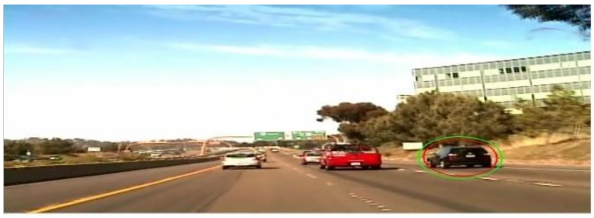
1. Background frame has been calculated by taking average of all the pixels.
2. Frame number has been selected from which tracking of any object has to be started.
3. From selected frame object to be tracked has been selected by repositioning the mask.

4. For selected object its centroid position has been found out and from centroid information all the equation of time and measurement update have been calculated. For selected frame the actual position X and error P has been calculated.

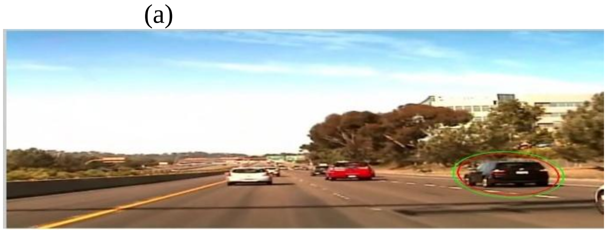
For all remaining frames following steps have been repeated.

1. Background subtraction has been done to find out all the moving regions in the frame.
2. From the found regions, region with the lowest distance from the region selected in previous frame has been selected.
3. Selected region's centroid and other parameter have been used to calculate time and measurement update equations.
4. Obtained state position values X has been stored in Array for every frame.
5. Line joining each stored point has been drawn in every frame which shows the trajectory of the selected moving object.

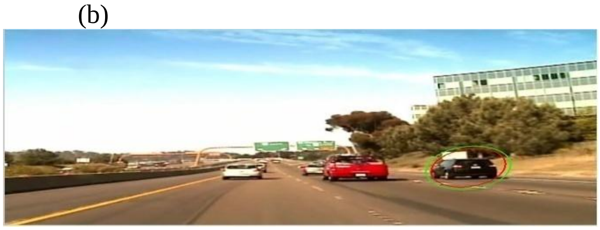
IV. RESULT ANALYSIS



(a)



(b)



(c)

Figure1: (a, b ,c) Tracking results of vehicle using MMST algorithms. The frames 20, 40 and 80 are displayed.

(a)

(b)

(c)

Figure 2: (a,b,c) Tracking results of vehicle for Kalman algorithms. The frames 20, 40 and 80 are displayed.

Above figure shows the output on applying modified mean shift algorithm and kalman algorithm on different objects for tracking and time required for tracking is given in table 1

Table 1: Comparison of MMST and Kalman Filer Object Tracking Methods

S. No.	**Video**	**No.of Frames**	**MMST**	**Kalman Filter**
			Time (in sec)	
1	Car1	300	324	244
2	One stop move	720	100	73

V. CONCLUSION

Form figure 1 and 2 it is clear that both the algorithm can track the object clearly, table 1 shows the time required by both the algorithms for tracking of objects and it is very clear from the table that kalman filter generates a good output with respect to time.

REFRENCES

[1] Fukunaga K, and Hostetler LD, "The estimation of the gradient of a density function, with applications in pattern recognition," IEEE Trans. Information Theory, vol. 21, pp.32-40, 1975.

[2] Cheng Y, "Mean shift, mode seeking, and clustering," IEEE Trans. on Pattern Analysis and Machine Intelligence, vol.17, no.8, pp.790-799, 1995.

[3] Comaniciu D, and Ramesh V, "Mean shift and optimal prediction for efficient object tracking," In: Mojsilovic A, Hu J, eds. Proc. of the IEEE Int'l Conf. on Image Processing (ICIP), pp.70-73, 2000.

[4] Comaniciu D, Ramesh V, and Meer P, "Real-Time tracking of non-rigid objects using mean shift," In: Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition (CVPR), pp.142-149, 2000.

[5] Collins RT, “Mean shift blob tracking through scale space,” In: Proc. of the Conf. on Computer Vision and Pattern Recognition (CVPR), pp.18-20, 2003.

[6] Shan C, Wei Y, and Tan T et al, “Real Time Hand Tracking by Combining Particle Filtering and Mean Shift,” In: Proc. of the 6th IEEE International Conf. on Automatic Face and Gesture Recognition, 17-19, May, pp.669-674, 2004.

[7] Maggio E, and Cavallaro A, “Hybrid Particle Filter and Mean Shift tracker with adaptive transition model,” In: Proc. of IEEE Signal Proc. Society Int. Conf. on Acoustics, Speech, and Signal Processing (ICASSP), Philadelphia, PA, USA, March pp.19-23, 2005.

[8] Comaniciu D, “Image segmentation using clustering with saddle point detection,” In: Proc. of the IEEE Int’l Conf. on Image Processing (ICIP), pp.297-300, 2002.

[9] Wang J, Thiesson B, and Y. Xu et al, “Image and Video Segmentation by Anisotropic Kernel Mean Shift,” In: Proc. European Conf. on Computer Vision (ECCV), 2004.

[10] Comaniciu D, Ramesh V, and A. D. Bue, “Multivariate saddle point detection for statistical clustering,” In: Proc. Of the European Conf. Computer Vision (ECCV). Pp.561-576, 2002.

[11] Georgescu B, Shimshoni I, and Meer P, “Mean Shift Based Clustering in High Dimensions: A Texture Classification Example,” In: Proc. ICCV, Oct. pp. 456-463, 2003.

[12] Comaniciu D, and Meer P, “Mean shift analysis and applications,” In: Proc. of the IEEE Int’l Conf. on Computer Vision (ICCV), pp.1197-1203, 1999.

[13] Comaniciu D, “Nonparametric information fusion for motion estimation,” In: Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition (CVPR), pp.59-66, 2003.

[14] Kalman, R. E.: A new approach to linear filtering and prediction problems, Transaction of the ASMEJournal of Basic Engineering, 35–45, March 1960.

[15] Watada JunzoMusa Zalili , Jain Lakhmi C., and Fulcher John., .," Human Tracking: A State-of-Art Survey.", KES 2010, Part II, LNAI 6277, pp. 454–463, 2010.

[16] Welch, G. and Bishop, G.: An introduction to the Kalmann Filter, In University of North Carolina at Chapel Hill, Department of Computer Science. Tech. Rep.95-041, available at http://russelldavidson.arts.mcgill.ca/e761/kalman-intro.pdf,2004.

[17] Xu, Sheldon, and Anthony Chang. "Robust Object Tracking Using Kalman Filters with Dynamic Covariance." available at http://www.cs.cornell.edu/Courses/cs4758/2011sp/final_projects/spring_2011/Xu_Chang.pdf

[18] Yilmaz ,Alper, Omer Javed and Mubarak Shah. " Object tracking : A survey." Acm Computing Surveys (CSUR) 38.4 (2006):13.

[19] Cheng Y.: ‘Mean Shift, Mode Seeking, and Clustering’, IEEE Trans on Pattern Anal. Machine Intell., 1995, 17, (8), pp. 790-799.

[20] Bradski G.: ‘Computer Vision Face Tracking for Use in a Perceptual User Interface’, Intel Technology Journal, 1998, 2(Q2), pp. 1-15.

[21] Comaniciu D., Ramesh V., Meer P.: ‘Real-Time Tracking of Non-Rigid Objects Using Mean Shift’. Proc. IEEE Conf. on Computer Vision and Pattern Recognition, Hilton Head, SC, June, 2000, vol. 2, pp. 142-149.

[22] Comaniciu D., Meer P.: ‘Mean Shift: a Robust Approach toward Feature Space Analysis’, IEEE Trans Pattern Anal. Machine Intell., 2002, 24, (5), pp. 603-619.

[23] Comaniciu D., Ramesh V., Meer P.: ‘Kernel-Based Object Tracking’, IEEE Trans. Pattern Anal. Machine Intell. 2003, 25, (2), pp. 564-577.

[24] Collins R.: ‘Mean-Shift Blob Tracking through Scale Space’, Proc. IEEE Conf. Computer Vision and Pattern Recognition, Wisconsin, USA, 2003, pp. 234-240.

[25] Zivkovic Z., Krose B.: ‘An EM-like Algorithm for Color-Histogram-Based Object Tracking’, Proc. IEEE Conf. Computer Vision and Pattern Recognition, Washington, DC, USA, 2004, vol.1, pp. 798-803.