

REAL TIME MOVING OBJECT TRACKING USING CORRECTED BACKGROUND WEIGHTED HISTOGRAM AND KALMAN FILTER FOR AN AIRCRAFT

Gargi G¹, Sudesh K Kasyap², VPS Naidu², Ajesh M A³

¹College of Engineering Kolloppara (Under Cochin University of Science and Technology CUSAT), ² National Aerospace Laboratory, ³College of Engineering, Kolloppara (Under CUSAT)

gargi.nrd@gmail.com

Abstract- The paper presents an approach for the real time tracking of the moving object in a video stream obtained from a moving airborne platform. The real time tracking means that the user can select the target from any frame at any instant. The Corrected Background Weighted Histogram (CBWH) based Mean Shift algorithm is used for designing the tracking algorithm. Which can effectively reduce the background interferences from the target representation and has high tracking accuracy than general tracking algorithms.

Keywords: Mean shift; Kalman filter; Corrected Background Weighted histogram; Object tracking; Occlusion.

I. INTRODUCTION

In recent years a lot of work has been done in Image Processing field. This paper is going to discuss the use of image processing on object tracking. Object tracking is one of the important fields in Image Processing. It is the process of locating a moving object over time using a camera. The innovation of high-powered computer, high quality camera and the need of automated video analysis have brought much attention to the object tracking algorithms. Here the objective is to track a moving object in a video sequence captured by a camera mounted on a moving aircraft. When an operator observes an object moving in a way that will impact on aircraft operations, they take the necessary steps to avoid undesired interaction. Their response depends on recognizing the type of object and its track. There are three key steps in video analysis: detection of the interesting objects, tracking of such objects from frame to frame, later depending on the application the tracks can be analyzed to recognize their behavior. This work is concentrated on single object tracking. Most of the papers are dealing with the selection of target from the starting still frame. If the first frame or the still frame does not contain the target then we cannot select the particular target for the tracking. So this is a disadvantage of all those papers. Here in this paper we can select the target from the video sequence at any instant or from any frame. Mean shift tracking algorithm [1] is used to design the tracking. Different features like color, texture, edges etc. can be used to represent the target here color feature is used. In order to reduce the

background information from the target model Background weighted Histogram (BWH) [1], [2] is used. But here we proposed a method called Corrected Background weighted Histogram (CBWH) [3], [4], which is same as Background Weighted Histogram (BWH) but CBWH can work robustly even if the target model contains much background interference. So CBWH has better tracking accuracy than simple Mean shift and BWH algorithms. For an efficient tracking we need to update the background. But this method can track the target under normal conditions but it will fail in partial or full occlusion, illumination changing conditions, scale and orientation changes of the target, clutter etc. Kalman filter [5] can be used to eliminate the partial or full occlusion. So here in this paper we will discuss both CBWH based mean shift tracking and Kalman filter.

II. MEAN SHIFT TRACKING

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 [6], [7]. The MS algorithm was originally proposed by Fukunaga in 1975 [1], and with the concept of kernel function, which became the practical foundation of MS [2]. The popularity of the MS method is due to its ease of implementation, real time response efficiency and robust tracking performance. It is a non-parametric feature space analysis technique for locating the maxima of a density function but it is prone to local minima when some of the target features are present in the background. Therefore Comaniciu et al. [2] further proposed a method Background Weighted Histogram (BWH). The strategy of BWH is to derive a simple background features and to use it to select the salient components from the target model and target candidate model. BWH decreases the probability of background features from the target model and target candidate model thus reduces the background interferences from the target. The proposed corrected background-weighted histogram (CBWH) can truly achieve what the original BWH method wants: reduce the interference

of background in target localization. An important advantage of the proposed CBWH method is that it can work robustly even if the target model contains much background information. Thus, it reduces greatly the sensitivity of mean-shift tracking to target initialization.

A. Target Representation

In the object tracking usually a rectangle or an ellipse is used to select the target. In order to minimize the error due to different target dimensions we need to normalize all targets. Let $\{X_i\}$ be the normalized pixel locations in the region defined as the target model. The target model can be considered as centered at the spatial location 0. The reference target model is represented by its pdf q in the feature space. In the subsequent frame, a target candidate is defined at location y , and is characterized by the pdf $\hat{P}(y)$. Both pdfs are to be estimated from the data.

Target Model [2]:

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

Target Candidate Model [2]:

$$\hat{P}(y) = \{\hat{P}_u(y)\}_{u=1..m} \quad \sum_{u=1}^m \hat{P}_u = 1 \quad (2)$$

A similarity function [1] between \hat{P} and \hat{q} is $\hat{\rho}(y)$.

$$\hat{\rho}(y) = \rho[\hat{P}(y), \hat{q}] \quad (3)$$

This function plays the role of a likelihood and its local maxima in the image indicate the presence of objects in the second frame having representations similar to \hat{q} defined in the first frame.

B. Target Model

An isotropic kernel, with a convex and monotonic decreasing kernel profile $k(x)$, assigns smaller weights to pixels farther from the center. Using these weights increases the robustness of the density estimation since the peripheral pixels are the least reliable, being often affected by occlusions (clutter) or interference from the background.

The probability of the feature $u = 1..m$ in the target model is then computed as

$$\hat{q}_u = C \sum_{i=1}^n k(\|x_i^*\|^2) \delta[b(x_i^*) - u] \quad (4)$$

Here δ is the Kronecker delta function. The normalization constant C is derived by imposing the condition $\sum_{u=1}^m \hat{q}_u = 1$, from where

$$C = \frac{1}{\sum_{i=1}^n k(\|x_i^*\|^2)} \quad (5)$$

Since the summation of delta function for $u = 1..m$ is equal to one.

C. Target Candidate Model

Let $\{X_i\}$ be the normalized pixel locations of the target candidate, centered at y in the current frame. The normalization is inherited from the frame containing the target model. Using the same kernel profile $k(x)$, but with bandwidth h , the probability of the feature $u = 1..m$ in the target candidate is given by [2],

$$\hat{P}_u(y) = C_k \sum_{i=1}^n k\left(\left\|\frac{y-x_i}{h}\right\|^2\right) \delta[b(x_i^*) - u] \quad (6)$$

Where

$$C_k = \frac{1}{\sum_{i=1}^n k\left(\left\|\frac{y-x_i}{h}\right\|^2\right)} \quad (7)$$

This normalization constant does not depend on y . Therefore, C_k can be precalculated for a given kernel and different values of h .

D. Mean Shift Tracking Algorithm

The aim of the mean shift algorithm is to compute the offset from the current location to the new location in accordance to the mean shift iteration equation.

The mean shift iteration equation is

$$y_1 = \frac{\sum_{i=1}^n X_i w_i g\left(\left\|\frac{y-X_i}{h}\right\|^2\right)}{\sum_{i=1}^n w_i g\left(\left\|\frac{y-X_i}{h}\right\|^2\right)} \quad (8)$$

Where $g(x)$ is the shadow of the kernel profile $k(x)$

$g(x) = -k'(x)$

$$w_i = \sum_{u=1}^m \sqrt{\frac{\hat{q}_u}{\hat{P}_u}} \delta[b(X_i) - u] \quad (9)$$

$$y_1 = \frac{\sum_{i=1}^n X_i w_i g_i}{\sum_{i=1}^n w_i g_i} \quad (10)$$

With this equation, the mean shift tracking algorithm can find the most similar region to the target object in the new frame.

E. Corrected Background Weighted Histogram

The name itself has the definition that it is a correction of the first model BWH. During the tracking process often background information may present in the target model. If the target features and background features are same then the tracking accuracy will be decreased. Therefore Jifeng Ning, Lei Zhang, David Zhang and Chengke Wu came with CBWH [3].

The background is represented as

$$\{\hat{\sigma}_u\}_{u=1..m} \text{ with } \sum_{i=1}^m \hat{\sigma}_u = 1 \quad (11)$$

And it is calculated by the surrounding area of the target. The background region is two times the size of target. $\hat{\sigma}^*$ is the minimum non-zero value in $\{\hat{\sigma}_u\}_{u=1..m}$.

$$\left\{v_u = \min\left(\frac{\hat{\sigma}^*}{\hat{\sigma}_u}, 1\right)\right\}_{u=1..m} \quad (12)$$

The coefficients $\{v_u\}_{u=1..m}$ is employed to transform only the target model but not the target candidate model. That is to say, we reduce the prominent background features only in the target model but not in the target candidate model.

From Eq. (9)

$$w_i' = \sum_{u=1}^m \frac{\hat{q}_u'}{\hat{p}_u'(y)} \delta[b(x_i) - u] \quad (13)$$

Let u be the bin index in the feature space which corresponds to point x_i in the candidate region. We have $\delta[b(x_i) - u] = 1$.

So Eq. 13 can be simplified to

$$w_i' = \sqrt{\hat{q}_u' / \hat{p}_u'(y)} \quad (14)$$

But the new target model,

$$\hat{q}_u' = C' v_u \sum_{i=1}^n k(\|x_i^*\|^2) \delta[b(x_i^*) - u] \quad (15)$$

With the normalization constant

$$C' = \frac{1}{\sum_{i=1}^n k(\|x_i^*\|^2) \sum_{u=1}^m v_u \delta[b(x_i^*) - u]} \quad (16)$$

Then the new candidate model is,

$$\hat{p}_u'(y) = C_k' v_u \sum_{i=1}^{n_h} k\left(\left\|\frac{y - x_i}{h}\right\|^2\right) \delta[b(x_i) - u] \quad (17)$$

Where

$$C_k' = \frac{1}{\sum_{i=1}^n k\left(\left\|\frac{y - x_i}{h}\right\|^2\right) \sum_{u=1}^m v_u \delta[b(x_i) - u]} \quad (18)$$

Then

$$w_i' = \frac{C' v_u \sum_{j=1}^n k(\|x_j^*\|^2) \delta[b(x_j^*) - u']}{C_k' v_u \sum_{j=1}^{n_h} k\left(\left\|\frac{y - x_j}{h}\right\|^2\right) \delta[b(x_j) - u']} \quad (19)$$

By removing the common factor v_u' from the numerator and denominator and substituting the normalization factors C and C_k into the above equation, we have

$$w_i' = \frac{C C_k \sum_{i=1}^n k(\|x_i^*\|^2) \delta[b(x_i^*) - u']}{C C_k \sum_{i=1}^{n_h} k\left(\left\|\frac{y - x_i}{h}\right\|^2\right) \delta[b(x_i) - u']} \quad (20)$$

$$= \sqrt{\frac{C' C_k}{C C_k'}} \cdot \sqrt{\frac{\hat{q}_u'}{\hat{p}_u'}} \quad (21)$$

$$= \sqrt{\frac{C' C_k}{C C_k'}} w_i \quad (22)$$

We define a new weight formula

$$w_i'' = \sqrt{\hat{q}_u' / \hat{p}_u'(y)} \quad (23)$$

We can also obtain

$$w_i'' = \sqrt{C' / C} \cdot \sqrt{v_u'} \cdot w_i \quad (24)$$

Since $\sqrt{C' / C}$ is a constant scaling factor, it has no influence on the mean shift tracking process. We can omit it and simplify the above equation as

$$w_i'' = \sqrt{v_u'} w_i \quad (25)$$

Here we can see that the weight calculated by using the usual target representation and the weight calculated by exploiting the background information are directly proportional. If the color at a point i in the background region is prominent, then the corresponding value of v_u' is small.

Hence in the above equation this point's weight is decreased and its relevance for target localization is reduced. This will then speed up mean shift's convergence towards the salient features of the target.

The background is initialized at the beginning of the tracking process. During the tracking the background may changing due to the illumination variations, scene content, view point etc. If the original background color model is still used without updating, the tracking accuracy may be reduced because the current background may be very different from the previous background model. Therefore it is necessary to dynamically update the background model for a robust CBWH tracking performance.

Here we propose a simple background model updating method. First, the background features $\{\hat{\sigma}'_u\}$ and $\{\hat{v}'_u\}$ in the current frame are calculated. Then the Bhattacharya similarity function between $\{\hat{\sigma}'_u\}$ and the old background model $\{\hat{\sigma}_u\}$ is computed by

$$\rho = \sum_{u=1}^m \sqrt{\hat{\sigma}_u \hat{\sigma}'_u}$$

If ρ is smaller than a threshold, this implies that there are considerable changes in the background, and then we update $\{\hat{\sigma}_u\}$ by $\{\hat{\sigma}'_u\}$ and update $\{\hat{v}_u\}$ by $\{\hat{v}'_u\}$. The transformed target model is the calculated using $\{\hat{v}'_u\}$. Otherwise we do not update the background model.

F. Kalman Filter (KF) for Occlusion Handling

The Kalman Filter was first introduced by Rudolf Kalman at NASA to track the trajectory of spacecraft. The KF [9] is a state estimation algorithm which compares state prediction against state measurements to get an accurate estimation of the true state. It has two stages, predict stage and update stage. In predict step, location of state of an object is predicted based on previous observations. In the update stage the measured location of the object from the tracking algorithm is combined with the predicted location in order to get an estimated location of the object. By combining CBWH and Kalman filter [10] we can improve the tracking efficiency.

i. Predict stage:

- a. Predicted (a priori) state:

$$\hat{x}_{t|t-1} = \mathbf{F}_t \hat{x}_{t-1|t-1} + \mathbf{B}_t \mathbf{U}_t$$

- b. Predicted (a priori) estimate covariance:

$$\mathbf{P}_{t|t-1} = \mathbf{F}_t \mathbf{P}_{t-1|t-1} \mathbf{F}_t^T + \mathbf{Q}_t$$

ii. Update stage:

- a. Optimal Kalman gain:

$$\mathbf{K}_t = \mathbf{P}_{t|t-1} \mathbf{H}_t^T (\mathbf{H}_t \mathbf{P}_{t|t-1} \mathbf{H}_t^T + \mathbf{R}_t)^{-1}$$

- b. Updated (a posteriori) state estimate:

$$\hat{x}_{t|t} = \hat{x}_{t|t-1} + \mathbf{K}_t (\mathbf{Z}_t - \mathbf{H}_t \hat{x}_{t|t-1})$$

- c. Updated (a posteriori) estimate covariance:

$$\mathbf{P}_{t|t} = \mathbf{P}_{t|t-1} - \mathbf{K}_t \mathbf{H}_t \mathbf{P}_{t|t-1}$$

The variables are used:

- X_t is the current state vector, as estimated by the Kalman filter, at time t .
- Z_t is the measurement vector from sensor.
- P_t measures the estimated accuracy of X_t at time t .
- F describes the motion of object from one state to another.
- H defines the mapping from the state vector, X_t , to the measurement vector, Z_t .
- Q and R define the Gaussian process and measurement noise, respectively, and characterize the variance of the system.

B and U are control Input parameters are only used in systems that have an input; these can be ignored in the case of an object tracker.

G. Algorithm Based on CBWH with Kalman filter

1. Calculate the target \hat{q} by Eq. (1) and the background weighted histogram $\{\hat{\sigma}'_u\}_{u=1..m}$ and then compute (26) $\{v_u\}_{u=1..m}$ by Eq. (12) and the transformed target model by Eq. (15). Initialize the position of the target candidate region in the previous frame.
2. Let $k \leftarrow 0$.
3. Calculate the target candidate model $\hat{p}(y)$ using Eq. (2) in the current frame.
4. Calculate the weights $\{w_i\}_{i=1..n_R}$ using Eq. (23).
5. Calculate the new position y_1 of the target candidate region using Eq. (10).
6. Let $d \leftarrow \|y_1 - y\|$, $y \leftarrow y_1$, $k \leftarrow k + 1$. Set error threshold ε_1 (default value is 0.1), the maximum iteration number N , and the background model update threshold ε_2 (default value: 0.5).

If $d < \varepsilon_1$ or $k \geq N$

Calculate $\{\hat{\sigma}'_u\}_{u=1..m}$ and $\{v'_u\}_{u=1..m}$ based on the tracking result of the current frame. If ρ by Eq. (26) is smaller than ε_2 , then $\hat{\sigma}_u \leftarrow \hat{\sigma}'_u$ and $v_u \leftarrow v'_u$ and \hat{q}'_u is updated by Eq. (15).

Stop iteration and estimate the object position then go to step 7 for next frame.

Otherwise

Go to step 3.

III. SETUP FOR GETTING THE VIDEO INPUT

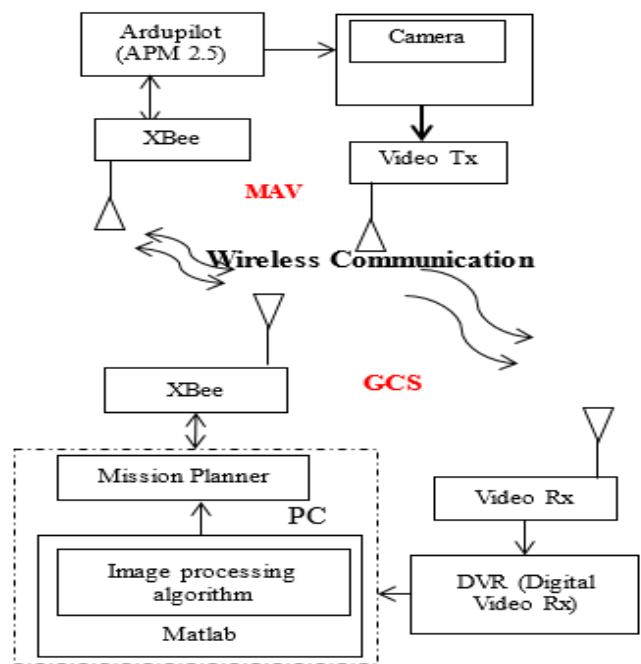


Fig.1. Setup for receiving the video sequence

A. Miniature Aerial Vehicle (MAV)

MAV is controlled with autopilot Arduplane mega (APM 2.5) [11] board. Gimbaled camera in MAV will provide video data which will be continuously transmitted to Ground Control Station (GCS) [12] through the video transmitter card. XBee PRO RF is used as transceiver for wireless communication. Transceiver is connected to APM 2.5 board to transmit navigational data and to receive control command from GCS.

B. Ground Control Station (GCS)

It sends ground commands to MAV, and also receive and displays the MAV flight parameters. GCS includes hardware and software. Hardware includes XBEE PRO RF transceiver, host computer system, auxiliary control system, power system and a user designed remote control system and input devices. The software includes Mission Planner, which helps to establish a wireless data link to obtain and display the MAV flight parameters and perform the ground control operation and sent the wireless data to the Ardupilot Mega (APM 2.5) through XBee PRO RF module [12]. XBee transceiver is used to transmit control commands and receive the navigational data from MAV. Video receiver is used to receive the video signal from camera in MAV. Video is received in PC by using TV toner card (DVR) which act as a frame grabber, that convert analog video to digital frames. DVR is connected to mission planner, by which it receives video. Operator/pilot can interact with MAV by giving commands as command line or by giving inputs through Joystick. To give input through Joystick, it should be connected to Mission Planner through USB port.

XBee PRO RF module meet the IEEE 802.15.4 standards and is operating with 2.4 GHz for worldwide development and with 900 MHz for long range development. It can transmit data at high speed with less delay.

For transmitting video sequence, the video transmitter operates at a frequency of 5.8 GHz. Here we are using the camera FPV 10X Zoom camera, providing images in NTPC format and in 700 TVL. The image produced by this camera is of 720X480.

IV. EXPERIMENTAL RESULT

The algorithms were implemented under the programming environment of MATLAB and several video sequences are used to evaluate those algorithms. Here we selected the Epanechnikov kernel as recommended in [2]. In the experiments, the RGB color space was used. Each color band was equally divided into 16 bins (16x16x16). The video is taken with a camera which is mounted on a moving aircraft. The aircraft is at a height of 300 feet above. This recorded video is used as the input for the MATLAB program. We can select any particular target from the live video, which is playing in the MATLAB figure window.

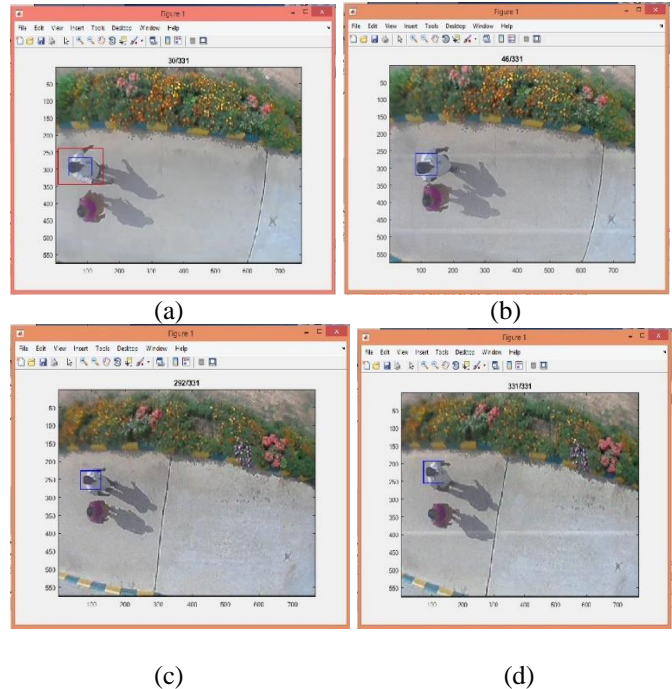
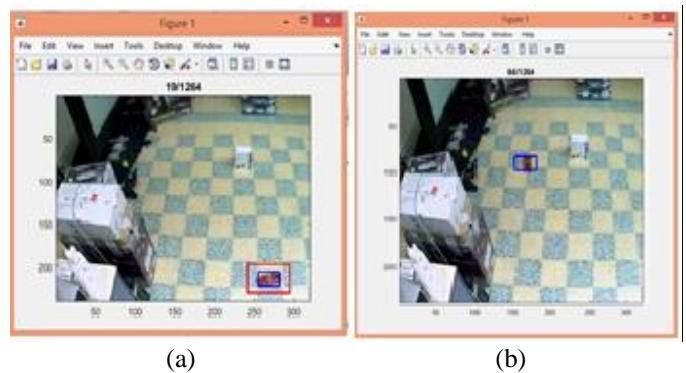


Fig. 2. Tracking with CBWH Based Mean shift Algorithm (a) selection frame, b-f tracking frames.

The object selection is done with mouse by drawing a rectangular box around the target. Here the advantage of this project is that we can select any object at any instant by pausing the video. The system is user friendly even with the people with less knowledge of Image processing. If we are using the automatic detection of the target more features may be used to detect the desired object.

Here total number of frames is 331. With the use of mouse, selected the target from the 30th frame. For the selection click and drag the mouse, which will create a rectangular box. Here the target is inside the blue rectangle which is the target window and the red rectangular box is the background window. The background window is two times larger than the target window. The selected frame is 30 so the tracking is from the frame 30 to the frame 331. The tracking window is changing its position from one frame to next frame in accordance with the movement of the target.



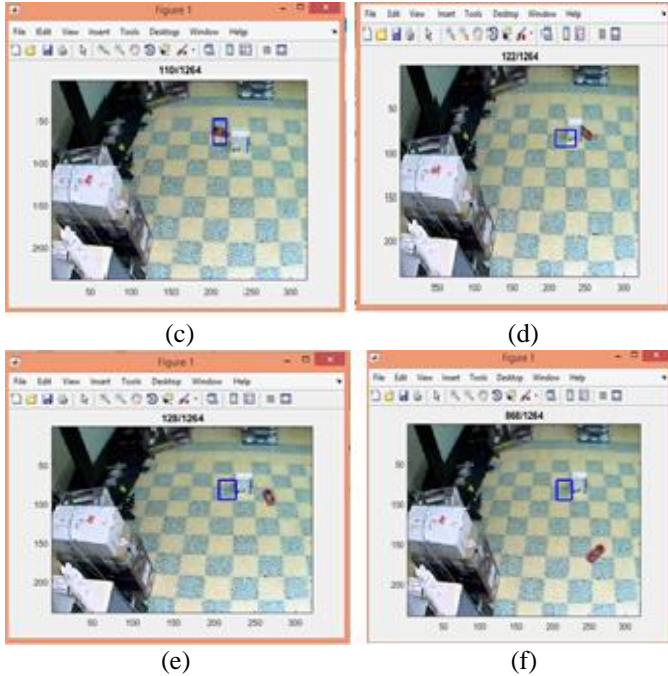


Fig. 3. Tracking without occlusion handling. (a) Selection frame, (b), (c) tracking frames before occlusion, (d), (e), (f) missed target while tracking after occlusion.

Here in the second experiment we can see that the tracking is carried out from the select frame and the tracking will remain until the car reaching the box or before the car is being occluded by the box. The tracking is missed at 221th frame. Then the CBWH algorithm cannot track the car in the next frames ie, from 221 to end frame. The traditional Mean shift algorithm also having the same effect with occlusion. So we are going for the incorporation of CBWH based Mean shift tracking and Kalman filter for the occlusion handling. The result of combined CBWH and Kalman filter is shown in the below figure. Here the tracking is carried out for the entire frames.

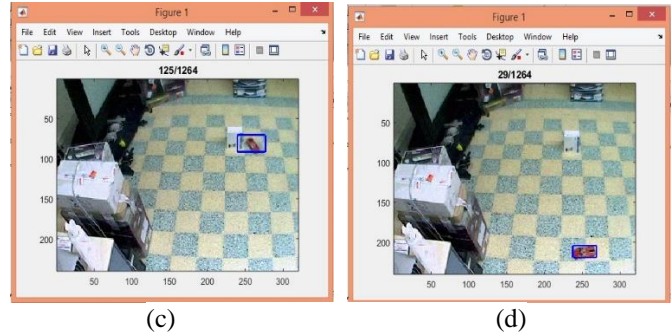
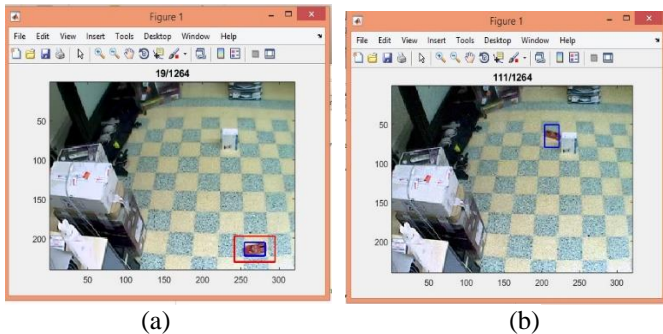


Fig. 4. a-f Tracking with CBWH and Kalman Filter

In the frame 111 we can see that the car is being occluded by the box. But the use of Kalman filter with CBWH handled the occlusion problem and we can see that the tracking is there after the frame 111 as in the previous frames. Hence we can obtain a successful tracking with CBWH and Kalman filter. The experimental results show that the proposed method is superior to the traditional MS tracking in the following aspects: 1) it provides consistent object tracking throughout the video; 2) it is not influenced by the objects with partial and full occlusions; 3) it is less prone to the background clutter.

A. Variations in Bhattacharya Coefficient (BC).

Here the Bhattacharya coefficient is plotted with frame number. Fig. 5 showing the variations of BC for the algorithm using CBWH only, for the tracking purpose. In this figure we can see that the BC is low for the frames in which the target is occlude by some other objects that means the BC is low when the tracking window loose the target.

Fig. 6 showing the variations of BC in the tracking of the target with combination of CBWH and Kalman filter for the occlusion removal. Here we can see that the BC is changing as we can see in the previous frames. Loose of tracking is not in this proposed method. The BC is getting decreased when the target is occluded but it is increasing again. That means the algorithm can track the target without any interference even in the occlusion also.

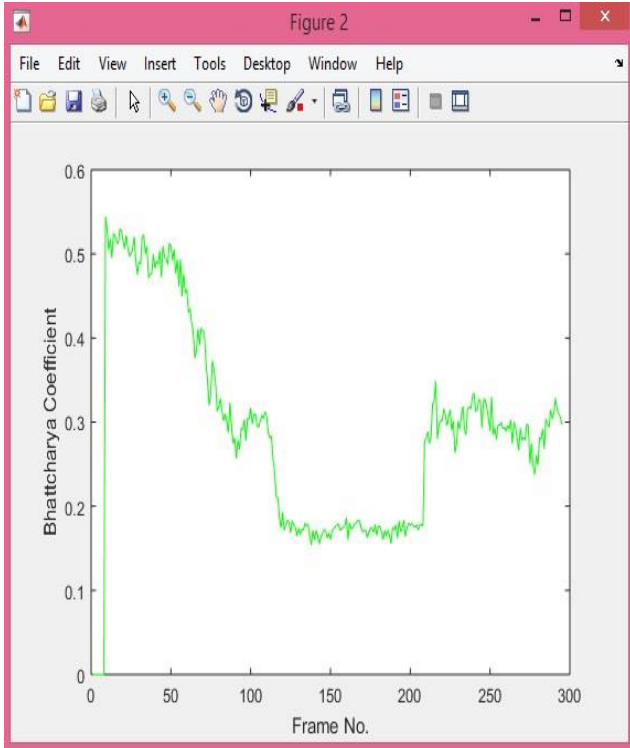


Fig. 5. Variations in BC with CBWH.

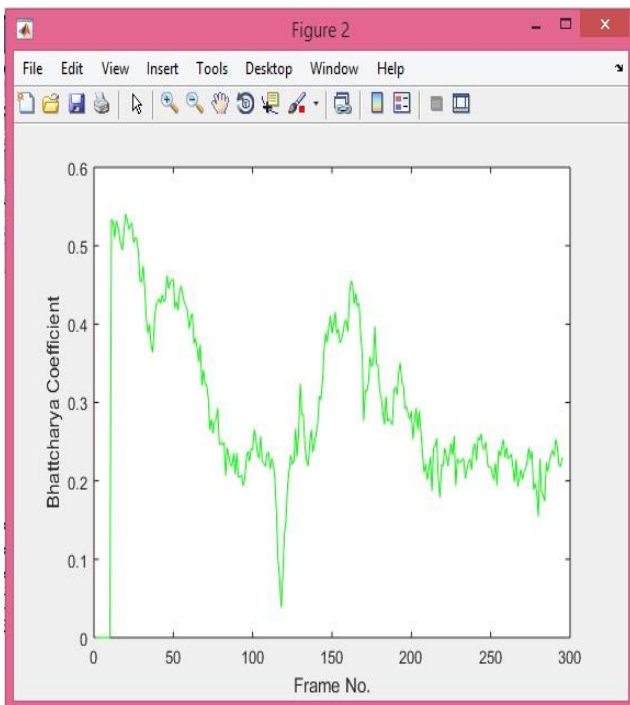


Fig. 6. Variations of BC in CBWH and Kalaman filter.

IV. CONCLUSIONS

Here proposed a CBWH method for the tracking to reduce the relevance of background information and improve

the target localization. The proposed algorithm only transforms the histogram of the target model and decreases the probability of target model features that are prominent in the background. The system successfully designed, tested and work as desired as the target object can be selected and being tracked in real time. Using this system now we can select and track any moving object from the video captured by the aircraft.

ACKNOWLEDGEMENT

We gratefully acknowledge the support from National Aerospace Laboratories, without which the work could not have been completed. For the completion of the paper we received a great support from Faculties in College of Engineering, Kalloppa under Cochin University of Science and Technology.

REFERENCES

- [1] Comaniciu, D., Ramesh, V., Meer, P.: 'Real-time tracking of non-rigid objects using mean shift'. Proc. IEEE Conf. Computer Vision and Pattern Recognition, Hilton Head, SC, USA, June 2000, pp. 142-149
- [2] Comaniciu, D., Ramesh, V., Meer, P.: 'Kernel-based object tracking', *IEEE Trans. Pattern Anal. Mach. Intell.*, 2003, 25 ,(2), pp.564-577
- [3] Jifeng Ning, Lei Zhang and Chengke Wu, Robust Mean Shift Tracking with Corrected Background-Weighted Histogram, IET Computer Vision. January 2012, vol. 6.
- [4] Michael Daniel Nachipyangu, "Real Time Corrected Background Weighted Histogram Based Mean Shift tracking:*International Journal of Science and Research (IJSR)*, Volume 4, issue 4, April 2015.
- [5] Greg Welch and Gary Bishop "An Introduction to Klamam Filter" *SIGGRAPH* 2001, Los Angeles, CA, August 12-17, 2001
- [6] Barga Deori and Dalton Meitei Thounaojam "A Survey On Moving Object Tracking In Video" *International Journal on Information Theory (IJIT)*, Vol.3, No.3, July 2014.
- [7] Alper Yilmaz, Omar Javed, Mubarak Shah, "Object Tracking: A Survey", *ACM Computing Surveys*, Vol. 38, No. 4, 2006.
- [8] Zivkovic Z., Kröse 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.
- [9] K. Nickels and S. Hutchinson, Estimating Uncertainty in SSD Based Feature Tracking, *Image and Vision Computing*, 2002, vol. 20, pp. 47-58.
- [10] Zhu, Z., Ji, Q., Fujimura, K., Lee, K.: Combining Kalman filtering and mean shift for real time eye tracking under active ir illumination. In: Proceedings

of 16th *International Conference on Pattern Recognition (ICPR 2004)*, vol. 4, p. 40318 (2002).

- [11] Ardupilot mega board- The APM 2.5 Board Overview <http://copter.ardupilot.com/wiki/apm2.5> board overview accessed on 4th September -2015.
- [12] Indhu, B and VPS Naidu (2015). Gimbaled Camera Control for On-Point Target Tracking. *American Research Journal of Electronics and Communication*, Vol. 1, Iss 1.