Implementation of Modified Mean-shift Tracking Algorithm for Occlusion Handling

Baber Khan[1], Ahmad Khalil Khan[1], Gulistan Raja[1], Muhammad Haroon Yousaf[2]

[1]Department of Electrical Engineering, UET, Taxila, Punjab, Pakistan
[2]Department of Computer Engineering, UET, Taxila, Punjab, Pakistan
ahmad.khalil@uettaxila.edu.pk

Abstract: Object tracking is critical and difficult task when it comes to unmanned air vehicles and traffic surveillance. The major challenge in object tracking is occlusion handling either partial occlusion or full occlusion. In this paper modified mean-shift tracking algorithm is proposed to tackle the problem of full occlusion. Mean-shift object tracking algorithm uses the color information to represent the target and to localize it in next frame. So when the object gets occluded with other object having similar colors, mean-shift tracking algorithm easily lost the target. In this modified mean-shift algorithm implementation, traditional mean-shift algorithm is lumped with the motion information associated with the spatial information of the moving object. Spatial information was exploited to handle the full occlusions present in the video. The object moves from one pixel to other in two consecutive frames its information about pixel index was stored in new variables. This information was used to capture the correct object when it reappears in the video, after the occlusion. Many videos were used to test the proposed tracking algorithm. Two examples were presented in this paper, which successfully cope with the partial occlusions, full occlusions and full occlusions when both the objects have exactly same colors.

[Baber Khan, Ahmad Khalil Khan, Gulistan Raja, Muhammad Haroon Yousaf. **Implementation of Modified Mean-shift Tracking Algorithm for Occlusion Handling**. *Life Sci J* 2013; 10(11s): 337-342]. (ISSN: 1097-8135). http://www.lifesciencesite.com 62

Keywords: Occlusion handling; spatial information; object tracking; mean-shift tracking

1. Introduction

The Object tracking means to follow the trajectory of the object when it moves around the scene in consecutive frames. Fast and efficient tracking of moving objects is important for applications like video monitoring, traffic navigation, traffic management (Prajna et al., 2012), video broadcasting, teleconferencing(Kai Du and Yongfeng Ju,Yinli Jin, 2012), human-computer interface, unmanned air vehicle attacks etc (Gargi phadke, 2011). During video monitoring, it helps understanding of motion carving to uncover doubtful events. In traffic navigation system it helps to prevent the accident by keeping the vehicles in their respective lanes. In traffic management systems it controls the flow and keeps the traffic smooth. Video broadcasting makes use this for better compression of data and improving download speed. In unmanned air vehicle attacks, tracking algorithms are implemented to track and hit the target accurately with minimum chance of error.

Many problems make the target tracking a complex problem. For example, loss of information: by projecting 3-dimenstional date on 2-dimensional image, noise in images, complex object motion, nonrigid or articulated nature of objects, partial and full object occlusions, occlusions with objects having same frequency of colors, complex object shapes, scene illumination changes, and real-time processing requirements (EmanueleTrucco and Konstantinos Plakas, 2006) and (AlperYilmaz, Omar Javed and Mubarak Shah, 2006). Tracking algorithms can be divided into two major categories (Gargi phadke, 2011): Mean-shift tracking(Comaniciu D, Ramesh V, Meer P, 2000) and particle filtering based tracking. Particle filtering based tracking is more efficient in terms of occlusion handling but have high computational time (Satoshi Y, 2008). Mean-shift tracking algorithm is preferred due to low processing time and invariance to object shape as given by (Comaniciu D, Ramesh V, Meer P, 2000). It is most suitable algorithm for real time object tracking according to (Cheng Y, 1995). According to (Fukunaga K, Hostetler L., 1975), mean-shift tracking algorithm is statistical method which is non parametric and has many advantages (D. Comaniciu, P. Meer, 1999) and (Comaniciu D, Ramesh V, Meer P, 2000). At first Mean-shift tracking algorithm was given in 1998 by (Bradski G R, 1998). It used the frequency of colors associated with the object to be tracked. It did not contain any spatial information associated with the object to be tracked, so when two objects in a same frame have same color information, Mean-shift algorithm suffer with a problem in such a scenario. Another problem in tracking algorithm arises, when object gets occluded. Two major types of occlusion are: partial occlusion and full occlusion. Problem of partial occlusion can be solved by using isotropic kernel (D. Comaniciu,V, Ramesh, and P.Meer, 2003). In this method spatial masking of object is done with isotropic monotonically decreasing kernel function, which gives more weight to the central components and less weight to the components which are farther from the center. As in case of partial occlusion, objects get occluded at corners mostly. Kernel based algorithms successfully cope with the problem of partial occlusion (D. Comaniciu, V.Ramesh, and P.Meer, 2003). A lot of effort has been made in the field of artificial intelligence to cater the problem of full occlusion. (De Villiers et al, 2006), used the mix features to get the descriptive representation of the object. Mean-shift was used to find the next location while kalman filtering was used to detect the object in case of occlusion. Object tracking is very useful to overcome the traffic congestion and hence accident, as most of the accidents occurs due to the intersections of roads. Background subtraction technique was applied on mean-shift tracking to track the trajectory of objects like car (Rawat et al., 2011). (Kai Du, Yongfeng Ju, Yinli Jin, 2012) developed an anti occlusion algorithm combination of mean-shift and kalman filter algorithm. Only frequency of colors was considered by mean-shift algorithm so when the color of object and background matches, the traditional mean-shift algorithm suffers with problems to correct object. (Kai Du,Yongfeng Ju,Yinli Jin, 2012) proposed new algorithm blend of mean-shift and SIFT.

As color feature was heavily used in the tracking algorithms (Prajna et al., 2012), (Cheng Y, 1995). Algorithms presented by (Rawat et al., 2011) and (de Villiers et al., 2012) are anti occlusion but have high computational cost. It is a challenging task to track the object when object is fully occluded with the other object having different frequency of color but it becomes even more difficult and challenging when objects are fully occluded with the other objects having same frequency of colors. To cater this problem of full occlusion, algorithm which exploit spatial information, was developed and proposed in the section II. Proposed algorithm is more robust in a sense of occlusion.

2. Modified Mean-shift Tracking Algorithm

Mean-shift is a very useful and versatile non parametric iterative algorithm that is applicable to numerous applications like finding modes, clustering etc (de Villiers et al., 2012),(Rawat, 2011). In last half decade, mean-shift tracking algorithm has been extended to be applicable in other fields such as Computer Vision and artificial intelligence(de Villiers, B. Z., W. A. Clarke Robinson, 2012) , (Xinguo Yu et al., 2006). The block diagram of proposed algorithm from input data (video) to output data (video) is shown in (Figure 1). This is described as follows:

In the very first block of (Figure 1), input video was fed to the tracker followed by feature-space-selection. Feature selection is an important block of the system diagram. Frequently used features are color, edges, texture, optical flow, gradient etc. In this paper, Color feature was used, as mean-shift algorithm was implemented in consequent block.

In a mean-shift tracker block, mean-shift algorithm iteratively find the new location of the target candidate (target candidate is a candidate to be target model) according to (2.1) and compare it with the target model's (target model is an object to track) probability density function. If both the probability density functions matches to the threshold value. That target candidate will become a target model for the next frame. y0 is the initial location of the target model in the previous frame and y1 is the location of the target candidate found by mean-shift tracking algorithm in the current frame. This process will be carried out on each two consecutive frames as depicted by figure (2).

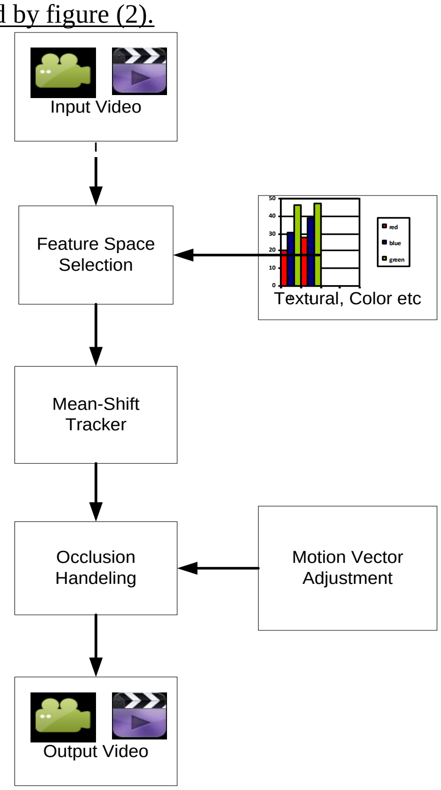

Figure1, System block diagram. Input and output to this diagram is video and proposed work was done in the block of occlusion handling

Three frames of the video were shown in (figure 2). In the very first frame target model will be identified by the user and in the next frame object will be searched, Once Mean-shift algorithm found the target that target will be verified with the proposed algorithm and if it found to be correct target, it will be locked and will be called target model in the first frame of next pair of frames. If the target found by mean-shift algorithm is not the actual target to be tracked then we have to apply mean-shift algorithm again to find the correct object. Algorithm is explained below with the help of mathematics.

Equation to localize the target in the current frame is given by (2.1) (AlperYilmaz, Omar Javed and Mubarak Shah, 2006) and (Kai Du, Yongfeng Ju, Yinli Jin, 2012) [5],[2]

$$y_1 = \frac{\sum_{i=1}^{n} x_i w_i}{\sum_{i=1}^{n} w_i} \qquad (2.1)$$

Where x_i is a value of the i^{th} pixel, n is the total number of bins in histogram, w_i is weighting function for i^{th} pixel and mathematically it is represented by (2.2).

$$w_i = \sum_{u=1}^{m} \delta[b(X_i) - u] \sqrt{\frac{\widehat{q_u}}{p_u(\widehat{y_o})}} \qquad (2.2)$$

Where i is index of pixel, (x_i) is the color bin index (for u=1 to m)of pixel x, q_u is the probability density function of target model and $p_u(y_o)$ is the probability density function of target candidate, δ is a delta function and its value becomes one when argument of this function becomes zero and m is total number of bins in the probability density function.

Last but not least block of the (figure 2) is occlusion handling. This is explained with the help of flow chart given by (figure 3). This flow chart is actually modified mean-shift tracking algorithm incorporated with mean-shift and motion vector adjustment. Flow chart starts with first frame in which object is chosen to be tracked in consecutive frames. In next block new location y1 is found by (2.1).Hereafter coordinates of the object were saved into variable C0 and C1, given by (2.3) and (2.4) respectively.

$$C0(x, y) = y0(x, y) \qquad (2.3)$$
$$C1(x, y) = y1(x, y) \qquad (2.4)$$

Tracker continuously finds the location of the center of the object in each frame and saves it. The center of the object C0(x, y) was compared to center C1(x, y).The mathematical expression for comparing the centers is given by (2.5).

$$C1(x, y) - C0(x, y) > 0 \qquad (2.5)$$

Expression (2.5) is valid for left to right motion. Where C0(x, y) is the center index location of the object found by the mean-shift Tracker in a current frame before occlusion, as mean-shift is an iterative algorithm it will find the new location y1 according to (2.1) in each frame therefore C1(x, y) is the center index location of the object found after occlusion in the current frame. Pictorially this phenomenon is given by (Figure 4).

Apply MMSTA Iteratively on frame # 1 and 2,if target found correctly in frame # 2 then apply MMSTA on frame # 2 and 3 ,this process will continue till the end of sequence/video.

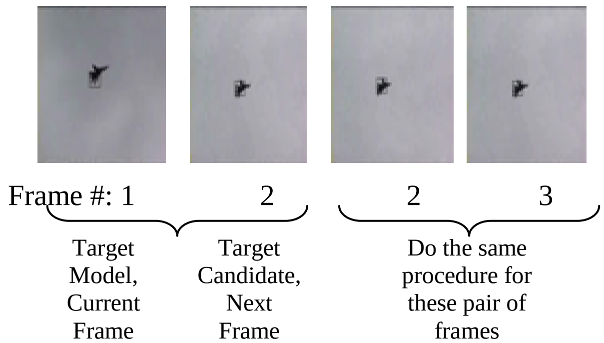

Figure 2, Modified Mean-shift procedure. Depicting that algorithm was applied on each pair of frames

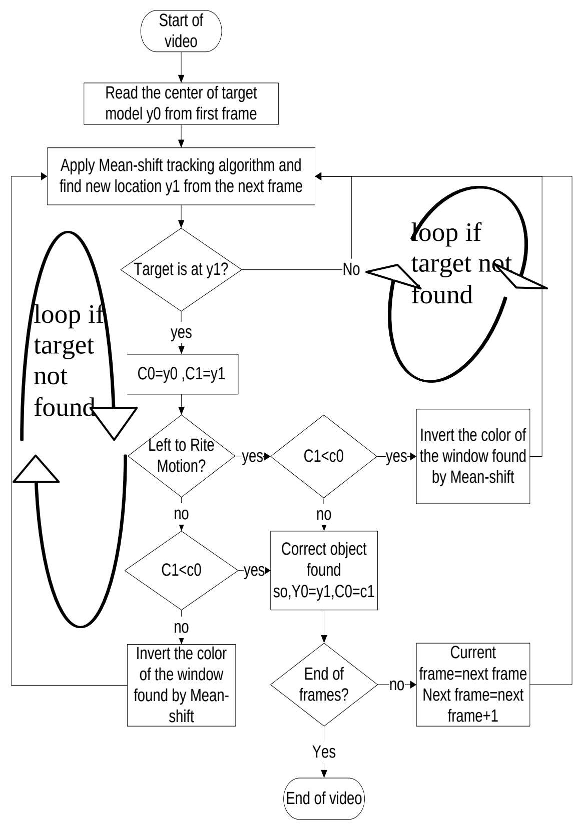

Figure 3, Flow chart of Modified Mean-shift object tracking algorithm

Figure4, Location of the tracked object before and after the occlusion.

According to equation (2.4) index of center found after the occlusion must be greater than the index of center found before the occlusion. The mathematical expression for right to left motion is given by (2.5).

$$C1(x, y)-C0(x, y) <0 \quad (2.6)$$

Similarly when object moves from right to left, the pixel index will be decreasing. Therefore, the index of central pixel found after occlusion must be smaller than the index found before occlusion, as given by (2.5).Now if occlusion occurs, this means that the mean-shift tracker did not following the expression (2.4) and (2.5), then complement around C1(x, y) was taken that is C1'(x, y).

The main idea behind this complement is that when next time mean-shift will be applied on same frame, this time it will not track the wrong target again and desired object will be tracked correctly. After localizing the correct target, actual values of the pixels around C1(x, y) was placed back. To make the solution robust algorithm was iterated infinitely though it can be iterated up to 4 or 5 times but in this case tracker may converge to wrong object. This phenomenon of infinite loop was depicted in (Figure 3).

3. Results and Discussions

Modified mean-shift tracking algorithm was tested on six different videos, all videos were in avi (audio video interleaved) format having different lengths. Proposed algorithm successfully tracks the correct target in all six videos. RGB color space was chosen as feature space and was quantized to 16*16*16 bins. The kernel was applied on histogram of target (Prajna et al., 2012). Kernel is nothing but smoothing function and its shape help in partial occlusion handling (D. Comaniciu,V.Ramesh, and P.Meer, 2003), mean-shift procedure was used to find the next location and the motion direction information was used to cope with the problem of full occlusion. Two out of six videos are presented in this paper.

(Figure 5) Video 1;An aero plane video with no occlusion having 50 frames, 352X240 pixels in each frame was initially tested. Aero plane was tracked with kernel based mean-shift object tracker and with modified kernel based mean-shift visual object tracker. Both the trackers were successful in tracking the aero plane with real time processing. Now the problem comes when target gets occluded as explained below with the help of results:

In figure 6 and 7, the white paper video with occlusions have 90 frames with 352X240 pixels in each frame. Both the algorithms were tested on this video. First experiment was done with simple kernel based mean-shift object tracker; as in this video there is no occlusion up to frame number 44 hence kernel based tracker successfully track the target up to the frame number 44. (Figure 6) clearly depicts that after frame number 44, target gets occluded with other body having same frequency of color, so after the occlusion kernel based mean-shift tracking algorithm loss its target and starts tracking the wrong target. Here it is concluded that kernel based mean-shift tracking algorithm fail to cope with the problem of full occlusion.

Figure 5, Video1; Aero plane video with no occlusion, having 50 frames

Figure6, Video 2;90 frames, target model fully occluded with other object of same appearance and shape, applied to traditional meanshift algorithm

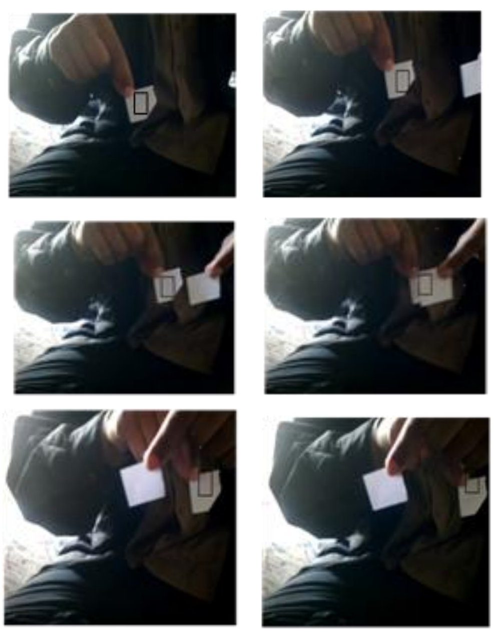

Figure7, Video 2; 90 frames, target model fully occluded with other object of same appearance and shape, applied to modified mean-shift algorithm

Again same experiment was done with modified mean-shift object tracking algorithm, results are given in (Figure 7). It is clear from the (Figure 7) that when the target gets occluded in the frame number 55 and reappears in the frame number 75, modified mean-shift tracking algorithm successfully track the target before and after the occlusion. From the above discussions and experimental results it is proved that modified mean-shift object tracker is more robust in terms of occlusion handling then simple kernel based mean-shift object tracker. Comprehensive comparison of both the techniques is given in (Table I). First column of (Table I) show the different algorithms used to track the object. Videoes on which the algorithms were tested is given in second column. Whereas, the third column represents the total number of frames in video sequence. Fourth column represent the total number of bins used in a histogram. As Bin form of histogram was used to minimize the processing time. Fifth field of Table 1 is representing the iterations used in an algorithms, successful were also taken for infinite iterations. Last but not least the fifth column is showing that either algorithm was capable of tracking the target or not.

Table 1. Comparison of modified mean-shift tracking algorithm with original mean-shift tracking algorithm

Algorithm	Sample Video	Total Frames	Bins	Iterations	Target Loss
Meanshift	Video 1 aeroplane	50	5	18	No
Meanshift	Video 2 paper video	90	2	18	Yes
Modified Meanshift	Video 1 aeroplane	50	5	18	No
Modified Meanshift	Video 2 paper video	90	2	18	No

4. Conclusion and Future work

From the section 3, it is concluded that modified Mean-Shift algorithm can perform better in situations where target to be tracked is occluded. Mean-Shift is color based algorithm, so the algorithm proposed in section 3 keeps the spatial information associated with the target model and can track the object efficiently when occlusion and the target have the same color. Scope of the proposed algorithm is limited to the situations where objects are moving 180 degree out phase in a straight line. As this algorithm can only be applied to the videos in which objects are moving in the opposite direction in straight line. When the objects are moving not at 180 degree then this algorithm suffers with problems. Algorithm can be enhanced further by dividing the frame/image into different quadrants and considering the angles other then 180 degree. So in future one can find the solution which cope with the problem discussed above to make the algorithm fully robust.

Corresponding Author:
Prof. Dr. Ahmad Khalil Khan
Department of Electrical Engineering
University of Engineering and Technology, Taxila, Punjab, Pakistan.
E-mail: ahmad.khalil@uettaxila.edu.pk

References

[1] Prajna Parimita Dash, Dipti patra, Sudhansu Kumar Mishra,Jagannath Sethi, “Kernel based Object Tracking using Color Histogram Technique “ International Journal of Electronics and Electrical Engineering ISSN : 2277-7040 Volume 2 Issue 4 (April 2012).
[2] Kai Du,Yongfeng Ju,Yinli Jin; Gang Li,Yanyan L, Shenglong Qian, "Object tracking based on improved Meanshift and SIFT," 2nd International Conference on, Consumer Electronics, Communications and Networks (CECNet), 2012 pp.2716,2719, 21-23 (April 2012).
[3] Gargi phadke “Robust Multiple Target Tracking Under Occlusion Using Fragmented Mean Shift and Kalman Filter” 2011 International Conference on Communications and Signal Processing (ICCSP) , PP 517 – 521, (2011).
[4] EmanueleTrucco and KonstantinosPlakas “Video Tracking: A Concise survey” IEEE Journal. Ocean Eng., vol. 31, no. 2, pp. 520-528. (2006).
[5] AlperYilmaz, Omar Javed and Mubarak Shah, “Object Tracking: A Survey” ACM Comput. Surv.38,4,Article 13, (2006).
[6] Comaniciu D, Ramesh V, Meer P “Real-time Tracking of Non-rigid Objects Using Mean Shif”t[C]. Proceedings of the IEEE Conference on CVPR, pp 142-149(2000).
[7] Satoshi Y, "Multiple tracking using mean shift with particle filter based Intialization". International conferenceon information visualization (2008).
[8] Cheng Y Mean Shift, “Mode Seeking, and Clustering” [J]. IEEE Trans on Pattern Analysis. Machine Intell., 17(8): 790(1995).
[9] Fukunaga K, Hostetler L. “The Estimation of the Gradient of a Density Function, with Applications in Pattern Recognition”[J]. IEEE Trans. On Inform. Theory, (1975).
[10] D. Comaniciu, P. Meer. “Mean shift analysis and application” [J]. Proceedings of the Seventh IEEE International Conference Computer Vison. (1999).
[11] Bradski G R. “Computer Vision Face Tracking for Use in a Perceptual User Interface” [EB/OL](1998).
[12] D. Comaniciu,V.Ramesh, and P.Meer. “Kernel-based object tracking”. IEEE Trans. Patt. Analy. Mach. Intell. 25, pp-564–575. (2003).
[13] Yuan, Baohong, Dexiang Zhang, Kui Fu, and Lingjun Zhang. "Video tracking of human with occlusion based on Mean-shift and Kalman filter." In IEEE 4th Electronic System-Integration Technology Conference (ESTC), pp. 148-151, (2012).
[14] Rawat, Manmohan Singh, N. S. Varun, and S. R. Vinotha. "Occlusion Detection and Handling In Video Surveillance." International Proceedings of Computer Science and Information Technology, (2011).
[15] de Villiers, B. Z., W. A. Clarke, and P. E. Robinson. "Mean Shift Object Tracking with Occlusion Handling." Proceedings of the Twenty-Third Annual Symposium of the Pattern Recognition Association of South Africa, (2012).
[16] Xinguo Yu, Leong, H.-W.; Changsheng Xu; Qi Tian, "Trajectory-Based Ball Detection and Tracking in Broadcast Soccer Video," Multimedia, IEEE Transactions on, vol.8, no.6, pp.1164,1178, (2006).
[17] Kai Du; Yongfeng Ju; Yinli Jin; Gang Li; Shenglong Qian; Yanyan Li, "Mean-shift tracking algorithm with adaptive block color histogram,"2nd International Conference onConsumer Electronics, Communications and Networks (CECNet), pp.2692,2695, (2012).
[18] Long, Yonghong, Xiyu Xiao, Xiaohua Shu, and Shenglan Chen. "Vehicle Tracking Method Using Background Subtraction and Mean-shift Algorithm." In IEEE International Conference on E-Product E-Service and E-Entertainment (ICEEE), pp. 1-4,(2010).
[19] H.Schweitzer, J.W.Bell, andF.Wu., “Very fast template matching”.European Conference on Computer Vision (ECCV).pp.358–372. (2002).
[20] P.Fieguth, and D.Terzopoulos.“Color-based tracking of heads and other mobile objects at videoframe rates”. In IEEE Conference on Computer Vision and Pattern Recognition (CVPR).pp.21–27. (1997).
[21] D.Comaniciu, P. Meer, “Mean shift: A robust approach toward feature space analysis”. IEEE Trans. Patt. Analy.Mach. Intell.24, 5, pp-603–619. (2002).
[22] J.Kang, I. Cohen, G. Medioni. “Object reacquisition using geometric invariant appearancemodel”. International Conference on Pattern Recongnition (ICPR). pp-759–762. (2004).

11/14/2013