IJCSE International Journal of Computer Sciences and Engineering Open Access
Research Paper Volume-4, Issue-11 E-ISSN: 2347-2693

Experimental analysis of Mean shift method of tracking objects

Renuka Devi S M

ECE Dept, GNITS, JNTUH
renuka.devi.sm@gmail.com

Available online at: www.ijcseonline.org

Received: Oct/09/2016 Revised: Oct/21/2016 Accepted: Nov/15/2016 Published: Nov/30/2016

Abstract— Real time object tracking is a perplexing task in computer vision. Many algorithms exist in literature like Mean shift, background-weighted histogram (BWH) and Corrected background-weighted histogram(CBWH) for tracking the moving objects in a video sequence.This paper attempts to do the comparative analysis of the three methods in terms of performance parameters like Normalised Centroid Distance , Overlap and number of iterations using two types of features i.e., color histogram and color texture histogram. Experimental results show that the performance of CBWH gives better performance when compared with basic Mean shift and BWH.

Index Terms— Object Tracking, Mean Shift Algorithm, Target Feature Modelling, Candidate Feature Modelling, Bhattacharya Coefficients

1. INTRODUCTION

All Real time object tracking is a critical task in computer vision. Some of the important applications of object tracking are: Automated video surveillance, Robot vision, Traffic Monitoring, Animation: 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. Many algorithms [1],[2] have been proposed to solve the various problems arisen from noises, clutters and occlusions in the appearance model of the target to be tracked. Among various object tracking methods, the mean shift tracking algorithm [3], is a popular due to its simplicity and efficiency.

Object tracking can be defined as the process of segmenting an object of interest from a video scene and keeping track of its motion, orientation, occlusion etc[4,5]. in order to extract useful information[6]. The tracking is performed by monitoring object's spatial and temporal changes during a video sequence, including its presence, position, size, shape, etc[6][7].

Mean shift algorithm[8] was originally proposed by Fukunaga and Hostetler[3] for data clustering.Also many algorithms were proposed on features taken from video sequence like color histogram, and color-texture histogram, extending to online features selection[11] . Comaniciu and Meer [12] successfully applied mean shift algorithm to image segmentation and object tracking. Mean Shift is an iterative kernel-based deterministic procedure which converges to a local maximum of the measurement function with certain assumptions on the kernel behaviours. Furthermore, mean shift is a low complexity algorithm, which provides a general and reliable solution to object tracking and is independent of the target representation.

Comaniciu et al. [12] further proposed the background-weighted histogram (BWH) to decrease background interference in target representation. The strategy of BWH is to derive a simple representation of the background features and use it to select the salient components from the target model and target candidate model. More specifically, BWH attempts to decrease the probability of prominent background features in the target model and candidate model and thus reduce the background's interference in target localization. Such an idea is reasonable and intuitive, and some works have been proposed to follow this idea. In [11], the object is partitioned into a number of fragments and then the target model of each fragment is enhanced by using BWH. Different from the original BWH transformation, the weights of background features are derived from the differences between the fragment and background colors. In [12,13], the target is represented by combining BWH and adaptive kernel density estimation, which extends the searching range of the mean shift algorithm. In addition, Ning et al. [13] proposed a corrected background-weighted histogram (CBWH) method of mean shift algorithm and demonstrated the efficiency of this technique. All the above BWH and CBWH based methods aim to decrease the distraction of background in target location to enhance mean-shift tracking. The research contribution of this paper is, to perform comparative analysis of these three methods of object tracking. Finally it is observed CBWH perform better compared to basic Mean Shift and BWH.

The paper is organised as follows: Section 2 deals with basics of mean shift tracking, Section 3 presents with BWH and CBWH tracking algorithms, Section 4 gives comparison of experimentation results of the three methods. Finally Section 5 presents the conclusion

2. TRACKING WITH MEAN SHIFT ALGORITHM

2.1 Mean Shift Tracking[3]: In [18][2] [20]exhaustive object tracking methods are compared.

A key issue in the mean shift tracking algorithm is the computation of an offset from the current location y to a new location y_1 according to the mean shift iteration equation.

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

Where

$$w_i = \sum_{u=1}^{m} \sqrt{\frac{\hat{q}_u}{\hat{p}_u(y_0)}}\ \delta[b(x_i) - u] \quad (2)$$

When we choose kernel g with the Epanechnikov profile, (3.7) is reduced to

$$y_1 = \frac{\sum_{i=1}^{n_h} x_i w_i}{\sum_{i=1}^{n_h} w_i} \quad (3)$$

By using (3), the mean shift tracking algorithm finds in the new frame the most similar region to the object [8].

2.2 Target Tracking Using Color-Texture Histograms[17]

Mean shift Tracking using Color-texture Histograms

We use the RGB channels and the LBP patterns extracted jointly to represent the target and embed it into the mean shift tracking framework. Obtain the color and texture distribution of the target region, to calculate the color and texture distribution of the target model $\hat{q}$, in which u = 16×16×16×5(for only color distribution of target region, u = 16×16×16). The first three dimensions (i.e. 16 ×16×16) represent the quantized bins of color channels and the fourth dimension (i.e. 5) is the bin of the modified LBP texture patterns. Similarly, the target candidate model $\hat{p}(y)$ is calculated. The whole tracking algorithm is summarized as follows.

Input: the target model $\hat{q}$ is calculated and its location y_0 in the previous frame.
(1) Initialize the iteration number k ← 0.
(2) In the current frame, calculate the distribution of the target candidate model $\hat{p}(y)$.
(3) Calculate the weights $\{w_i\}_{i=1\ldots\ldots n_h}$ (4) Calculate the new location y_1 of the target candidate
(5) Let k ← k + 1, d ← $\| y_1 - y_0$, $y_0 \leftarrow y_1$. Set the threshold ε and the maximum iteration number N. (Here threshold is set to 0.1 and maximum iterations are 15)
If d < ε or k ≥ N Stop and go to Step 6. Otherwise Go to step 2.
(6) Load the next frame as the current frame with initial location y_0 and go to Step 1.

3. TRACKING WITH BWH AND CBWH MEAN SHIFT ALGORITHM

Based on the mean shift iteration formula, the key to effectively exploit the background information is to decrease the weights of prominent background features. Therefore, a corrected background-weighted histogram (CBWH) [11]is proposed to transform only the target model but not the target candidate model. A new formula for computing the pixel weights in the target candidate region is then derived. The CBWH algorithm can truly reduce the interference of background in target localization. An important advantage of the 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. In the experiments, we can see that even when the initial target is not well selected, the CBWH algorithm can still correctly track the object, which is hard to achieve by the usual target representation.

3.1 BWH Mean Shift Tracking

In order to reduce the interference of salient background features in target localization, a representation model of background features was proposed by Comaniciu et al. [12,14] to select discriminative features from the target region and the target candidate region.

In [9], the background histogram is represented as $\{\hat{o_u}\}_{u=1\ldots..m}$and it is calculated by the surrounding area of the target. The background region is considered to be three times the size of the target as suggested in [9,12]. The minimal non-zero value in $\{\hat{o_u}\}_{u=1\ldots..m}$ is denoted by $\hat{o}^*$. The coefficients used to define a transformation between the representations of target model and target candidate model are given below.

$$\{v_u = \min(\hat{o}^*/\hat{o}_u\, , 1)\}_{u=1\ldots.m}$$

This transformation reduces the weights of those features with low V_u , i.e. the salient features in the background. Then the new target model is defined as:

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

Where

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

The new target candidate model is given as:

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

Where

$$C'_h = \frac{1}{\sum_{i=1}^{n_h} k\left(\| \frac{y-x_i}{h} \|^2\right) \sum_{u=1}^{m} v_u\, \delta(b(x_i) - u)}$$

The above BWH transformation aims to reduce the effects of prominent background features in the target candidate region on the target localization. In next section, however, it will prove that BWH cannot achieve this goal because it is equivalent to the usual target representation under the mean shift tracking framework.

3.2 CBWH Mean Shift Tracking[11]

Equivalence of BWH representation to usual representation: By the mean shift iteration formula, in the target candidate region the weights of points (determine the convergence of the tracking algorithm, only when the weights of prominent features in the background are decreased which will reduce relevance of background information for target localization.

Let's analyze the weight changes by using the BWH transform. w_i' denote the weight of point x_i computed by the BWH in the target candidate region. Further

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

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. can be simplified as

$$w_i' = \sqrt{\frac{\hat{q}_{u'}'}{\hat{p}_{u'}'(y)}}$$

Substituting the above equations we have

$$w_i' = \sqrt{\frac{C' v_u{}' \sum_{j=1}^{n} k(\| x_i^* \|^2)\ \delta[(b(x_i^*) - u')]}{C_h' v_{u'} \sum_{j=1}^{n_k} k\left(\| \frac{y - x_i}{h} \|^2\right)\ \delta[(b(x_j) - u')]}}$$

By removing the common factor $v_{u'}$ from the numerator and denominator and substituting the normalization factors C and C_h into the above equation, we have[11]

$$w_i' = \sqrt{\frac{'C' C_h}{C C_h'}}\, w_i$$

Where w_i calculated by Eq. (3.8) is the weight of point i in the target candidate region using the usual representation of target model and target candidate model. This suggests that w_i' is proportional to w_i. Moreover, by combining mean shift iteration,we have

$$y_i = \frac{\sum_{i=1}^{n_k} x_i g_i w_i}{\sum_{i=1}^{n_k} w_i g_i}$$

Above equaction shows that the mean shift iteration formula is invariant to the scale transformation of weights.

Therefore, BWH actually does not enhance mean shift tracking by transforming the representation of target model and target candidate model. Its result is almost same as that without using BWH.

CBWH Algorithm[11]

Although the idea of BWH is good, but the BWH algorithm does not improve the target localization. To truly achieve what the BWH wants to achieve, J.Ning proposed a new transformation method, namely the corrected BWH (CBWH) algorithm. In CBWH, Eq. 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

We then define a new weight formula,

$$w_i'' = \sqrt{\frac{\hat{q}_{u'}'}{\hat{p}_{u'}'(y)}}$$

Therefore, we can easily obtain that,

$$.\ w_i'' = \sqrt{v_{u'} w_i}$$

This clearly reflects the relationship between the weight calculated by using the usual target representation (i.e.w_i) and the weight calculated by exploiting the background information (i.e. w_i''). If the color of point i in the background region is prominent, the corresponding value of $v_{u'}$ is small. 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. Note that if we do not use the background information $v_{u'}$ will be 1 and w_i'' will degrade to with w_i the usual target representation.

The whole CBWH mean shift tracking algorithm is summarized as follows.

Input: the target model $\hat{q}$ is calculated and its location y_0 in the previous frame.

(1) Initialize the iteration number $k \leftarrow 0$.

(2) In the current frame, calculate the distribution of the target candidate model $\hat{p}(y)$.

(3) Calculate the weights $\{w_i''\}_{i=1......n_h}$.

(4) Calculate the new location y_1 of the target candidate.

(5) Let $k \leftarrow k + 1$, $d \leftarrow \| y_1 - y_0$, $y_0 \leftarrow y_1$. Set the threshold ε and the maximum iteration number N. (Here threshold is set to 0.1 and maximum iterations are 15)

If $d < \varepsilon$ or $k \geq N$. Calculate $\{o_{u'}\}_{u=1......m}$ and $\{v_{u'}\}_{u=1.........m}$ based on tracking result of current frame.

Stop and go to Step 6. Otherwise Go to step 2.

(6) Load the next frame as the current frame with initial location y_0 and go to Step 1.

4. RESULTS AND DISCUSSION

Several representative video sequences[19] are used to evaluate the CBWH mean shift method in comparison with the BWH based mean shift tracking, and mean shift tracking with usual target representation. The comparison of mean shift tracking, BWH based mean shift tracking and CBWH based mean shift tracking is done by considering two cases

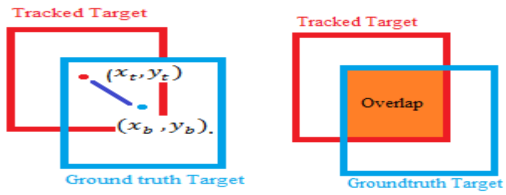


Fig 1 : The position of tracker (red) and its associated ground truth bounding box (blue). Centroid distance is represented by dark blue line. Overlap is represented by Orange.

Table 1: Comparison of target localization accuracies of three methods using color histogram

Equences	Mean shift		BWH Mean shift		CBWH Mean shift	
	Normalised Centroid Distance	Overlap	Normalised Centroid Distance	Overlap	Normalised Centroid Distance	Overlap
Table tennis	0.8900	0.8192	0.8900	0.8192	0.5543	0.8777
Pinpang ball	2.0080	0.5632	2.0080	0.5632	1.2952	0.7024
Basketball	0.5427	0.8741	0.5427	0.8741	0.3841	0.9201
Face	0.5730	0.8780	0.5730	0.8780	0.3123	0.9512

Table 2: Comparison of target localization accuracies of three methods using color-texture histogram

Sequences	Mean shift		BWH Mean shift		CBWH Mean shift	
	Normalised Centroid Distance	Overlap	Normalised Centroid-Distance	Overlap	Normalised Centroid Distance	Overlap
Table tennis	1.3457	0.7126	1.3457	0.7126	0.5837	1
Pinpang ball	0.3198	0.9003	0.3198	0.9003	0.2085	0.9565
Basketball	0.4240	0.9508	0.4240	0.9508	0.2206	1
Face	0.7880	0.8330	0.7880	0.8330	0.4951	0.8793

The Target localization accuracies for this algorithm are:

Normalised Centroid Distance: For a tracker cantered at (x_t, y_t) and a ground truth bounding box with centre (x_b, y_b). The normalized centroid distance in terms of the width w_b and the height h_b of the bounding box is given as

$$\text{Normalised Centroid distance} = \left(\frac{x_t - x_b}{w_b}\right)^2 + \left(\frac{y_t - y_b}{h_b}\right)$$

Overlap: The proportion of the ground truth bounding box that is occupied by the tracker in a given frame is another useful measure of the tracker's accuracy. This metric is referred to as the overlap which is a given by

$$\text{Overlap} = \frac{\text{area}_{\text{Common}}}{\text{area}_{\text{bounding\_box}}}$$

Case 1: In the analysis of mean shift tracking using color histogram, RGB color model is used as feature space in all the experiments and it is quantized into 16×16×16 bins. The tracking results of all the four video sequences for color histogram for BWH and CBWH are shown in Fig 2.

Table 1 show that the Comparison of target localization accuracies of three methods using color histogram. CBWH model has more accurate localization accuracy than the BWH model and usual mean shift model because the CBWH model truly exploits the background information in target localization with low Centroid distance and high overlap. Table 3 show the average number of iterations by three methods using color histogram. It can be seen that average number of iterations is less for CBWH compared to BWH and usual mean shift model. The salient features of target model are enhanced while the background features are suppressed in CBWH so that the mean shift algorithm can more accurately locate the target.

(a) Mean shift tracking

(b) BWH mean shift tracking

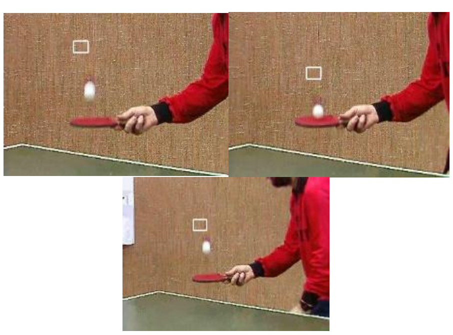
(c) CBWH mean shift tracking

Fig.2 Comparison of three methods using color histogram for pinpang ball sequence. Frames 10,24,40 are displayed.

Sequences	Mean shift	BWH mean shift	CBWH mean shift
Table tennis	3.51	3.51	2.74
Pinpang ball	7.92	7.92	4.32
Basketball	4.30	4.32	3.86
Face	4.12	4.12	3.30

Table 3 Average number of iterations for different video sequences using color histogram

Case 2: In the analysis of mean shift using color-texture histogram, RGB color model and LBP texture pattern are used as feature space in all the experiments and both are quantized into 16×16×16×5 bins. The tracking results of all the video sequences in Fig 3. Table 2 show the target localization accuracies of mean shift, BWH and CBWH based mean shift algorithms. Table 4 show the average number of iterations by three methods. Figure 4 gives the plot of number of iterations for pinpang ball video sequence using color-texture histogram.



(a) Mean shift tracking

(b) BWH mean shift tracking

(c) CBWH mean shift tracking

Fig 3: Comparison of three methods using color-texture histogram for pinpang ball sequence. Frames 10,27,39 are displayed.

Sequences	Mean shift	BWH mean shift	CBWH mean shift
Table tennis	3.24	3.24	3.00
Pinpang ball	6.05	6.09	5.48
Basketball	4.86	4.86	3.93
Face	4.20	4.25	3.51

Table 4: Average number of iterations for different video sequences using color-texture histogram

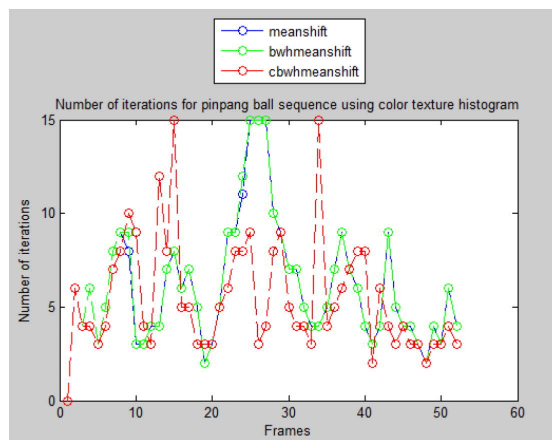


Fig 4: Plot of number of iterations for pinpang ball sequence using color-texture histogram

5. CONCLUSION:

In this paper, mean shift tracking based on color histogram and color-texture histogram is analysed using three methods of object tracking algorithm. In order to improve the target localization and to decrease the background interference in target representation, a background weighted histogram (BWH) and a corrected BWH (CBWH) method has been analysed. The experimental results validate that CBWH, not only reduces the mean shift iteration number but also provides improved tracking accuracy. Further advantage of CBWH is that it has reduced sensitivity than mean shift tracking to the target initialization i.e., CBWH can robustly track the target even if it is not well initialized.

REFERENCES

[1] Babenko B, Yang M H, Belongie S. "Robust object tracking with online multiple instance learning", IEEE Transactions on

Pattern Analysis and Machine Intelligence" 2011,33 (8): 1619-1632.

[2] Gandham. Sindhuja, Renuka Devi S.M.: "A Survey on detection and tracking of objects in a Video sequence", International Journal of Engineering Research and General Science Volume 3, Issue 2, Part 2, March-April, 2015, pp.418-426.

[3] K. Fukunaga, L.D. Hostetler, "The Estimation of the Gradient of a Density Function with Application in Pattern Recognition", *IEEE Trans. Information Theory*, vol. 21, no. 1, pp. 32-40, Jan. 1975.

[4] Saravanakumar, S. Vadivel and A. Saneem Ahmed, "Multiple human object tracking using background subtraction and shadow removal techniques", The International Conference on Signal and Image Processing ,2010, pp. 15-17.

[5] A. Yimaz, O. Javed and M. Shah, "Object tracking: A survey," ACM Computing Surveys, Vol. 38, No. 4, Article 13, December 2006, pp. 13-20.

[6] Gandham Sindhuja, Renuka Devi S.M. Comparative analysis of mean shift in object tracking. IEEE Conference on Power, Control, Communication and Computational Technologies for Sustainable Growth (PCCCTSG),2015, 283-287.

[7] Meng G, Jiang G, "Real-time illumination robust maneuver-ing target tracking based on color invariance", Proceedings of the 2nd International Congress on Image and Signal, 2009, pp. 15.

[8] Gammeter S, Bossard L, Gassmann A, et al. "Server-side object recognition and client-side object tracking for mobile augmented reality", CVPR, IEEE Computer Society Conference, 20 I 0: 1-8.

[9] N. A Gmez. "A Probabilistic Integrated Object Recognition and Tracking Framework for Video Sequences", University at Rovira I Virgili, PHD thesis, Espagne, 2009.

[10] Collins,R T, Yanxi Liu and Leordeanu, M, "Online selection of discriminative tracking features", IEEE Transactions on Pattern Analysis and Machine Intelligence , 2010, Vol. 10, pp. 1631-1643.

[11] Ying-JiaYeh, Chiou-Ting Hsu, "Online Selection of Tracking Features Using AdaBoost", IEEE Transactions on Circuits and Systems for Video Technology, 2009, VoU, pp. 442-446.

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

[13] Ning, Lei Zhang, David Zhang and C. Wu, "Robust Mean Shift Tracking with Corrected Background-Weighted Histogram," to appear in lET Computer Vision.(2011).

[14] D. Comaniciu and P. Meer. Mean shift: "A robust approach toward feature space analysis". PAMI,24(5):603-619, 2002.

[15] K. Nummiaro, E. Koller-Meier, and L. 1. Van Gool. "Object tracking with an adaptive color-based particle filter. In Proc. Of the 24th DAGM Symposium on Pattern Recognition, pages 353-360, London, UK, 2002. Springer-Verlag.

[16] Ess A, Schindler K, Leibe B, "Object detection and tracking for autonomous navigation in dynamic environments". IJRR, 2010, 29(14): 1707-1725.

[17] Shah M, Saleemi I, Hartung L, "Scene understanding by statistical modeling of motion patterns", IEEE Conference CVPR, 2010: 2069-2076.

[18] Patel, Sandeep Kumar, and Agya Mishra. "Moving object racking techniques: A critical review." Indian Journal of Computer Science and Engineering 4.2 (2013): 95-102.

[19]]cmp.felk.cvut.czl-vojirtom/datasetl, www.iai.unibonn.de/-kleindltracking. clickdamage.coml..Jcv _ datasets.php

[20] Ning J, Zhang L, Zhang D, et al. "Robust object tracking using joint color-texture histogram." International Journal of Pattern Recognition and Artificial Intelligence, 2009,23(07): 1245-1263.

[21] Pietikäinen, Matti, et al. "Local binary patterns for still images." Computer vision using local binary patterns. Springer London, 2011. 13-47.

