ACSIJ Advances in Computer Science: an International Journal, Vol. 3, Issue 3, No.9 , May 2014
ISSN : 2322-5157
www.ACSIJ.org

ACSIJ
WWW.ACSIJ.ORG

A Color-Adaptive and Robust Visual Object Tracking Method Based on MeanShift Algorithm

Arman Heydarian[1] , Milad Heydarian [2]

[1] **Electrical and Computer Engineering Department, Shahid Rajaee Teacher training University**
Tehran, Iran
Cdebugger@gmail.com

[2] **Electrical and Computer Engineering Department, *Isfahan University* of Technology**
Esfahan, Iran
Mld.heydarian@yahoo.com

Abstract

Visual object tracking is a key component in video analysis and surveillance system. In this paper we propose a novel and robust video object tracking method based on kernel tracking approach .MeanShift algorithm is a Kernel Tracking approach based on color histogram modeling. Because of changing of the color and shape of target model, it cannot track the object as much as possible. in some video streams with changeable color and brightness this method encounters with failure. So we manipulated some essential changes in original MeanShift and made it adaptive and more powerful in the realm of color, brightness and shape changes .The result of applying this method illustrates the high precision in our method for non-rigid objects in long videos.

Keywords: *Kernel Tracking, Color Weighted Histogram, Target Model, Target candidate.*

1. Introduction

Automatic object tracking is a digital image processing method, which consists of finding and tracking various moving objects through successive frames of the video sequence [1]. This function of computer vision is widely used in computer-human interfaces, robotics, medical applications, surveillance, etc. [1, 2].
The efficient tracking of visual features in complex environments is a challenging task for the vision community [3].
The most abstract formulation of the filtering and data association process is through the state space approach for modeling discrete-time dynamic systems [4]. in Object Tracking Field some algorithms has been developed and one of them is MeanShift in Kernel Tracking Method which it is based on Color Histogram Modeling.
Creation of object tracking systems with high precision has got 2 problems:
1.Tracking non-rigid (changeable) objects in the whole of the video stream .
2. robustness against noise when observation changes in sequence frames.
In MeanShift each image frame of the sequence is converted into a probability distribution image frame relatively to the histogram of the object to be tracked [5].
MeanShift initializes the model of target just one time in frame #0 then target candidate would be compared with target model in the whole of video stream. Based on the correspondence between both of them a weighted kernel would be created and the position of candidate moves to a heavyweight area.in the process of tracking if target observation changes due to changeling rate, it could be lost very fast.
In this paper we present a new Adaptive Menshift tracking algorithm that is robust to said problems and tracks non-rigid and changeable object in any environment (ex team sport players).
The principle of the Original MeanShift algorithm is given in section 2.

2. Original MeanShift Algorithm

The MeanShift is an algorithm trying to locate the object by finding the local maximum of a function [6]. The object target pdf is approximated by a histogram of m bins $\hat{q} = \{\hat{q}_u\}_{u=1...m}$, $\sum_{u=1}^{m} \hat{q}_u = 1$, with $\hat{q}_u$ being the u-th bin.
To form the histogram, only the pixels inside an box surrounding the object are taken into account. The center of the box is assumed to be at the origin of the axes. Due to the fact that the box contains both object pixels and background pixels a kernel with profile k(x), $k : [0,\infty) \to R$ is applied to every pixel to make pixels near the center of the box to be considered more important. To reduce the influence of different length of the box axes on the weights, the pixel locations are normalized by dividing the pixel's coordinates with the box's semi-axes dimension hx and hy.
Let $\{x^*_i\}_{i=1...n}$ be the normalized pixel's spatial location.
The u-th histogram bin is given by

ACSIJ Advances in Computer Science: an International Journal, Vol. 3, Issue 3, No.9 , May 2014
ISSN : 2322-5157
www.ACSIJ.org

$\hat{q}u = k(\_x* i \_2)\ \delta\ [b(x* i) - u]$ (1)

where b : $R^2 \rightarrow \{1 . . .m\}$ associates each pixel with each bin in the quantized feature space, δ is the delta function and C is a normalization factor such as $\sum_{u=1}^{m} \hat{q}u = 1$.In the next image, the object candidate is inside the same box with its center at the normalized spatial location y. Let {xi} 1...n be the normalized pixel coordinates inside the target candidate box. The pdf of the target candidate is also approximated by an m-bin histogram $\hat{p}(y) = \{\hat{p}u(y)\}u=1...m$, $\sum_{u=1}^{m} \hat{p}u(y) = 1$,
with each histogram bin given by

$\hat{p}u(y) = Cc\ n \sum_{i=1}^{n} k\ (|y - xi|^2)\ \delta[b(xi) - u]$ (2)

where Cc is a normalization factor such as $\sum_{u=1}^{m} \hat{p}u(y) = 1$.
The distance between $\hat{q}$ and $\hat{p}(y)$ is defined as:

$d(y) = \sqrt{1 - \rho[\hat{p}(y), \hat{q}]}$ (3)

where

$\rho[\hat{p}(y), \hat{q}] = \sum_{u=1}^{m} \sqrt{pu(y)\hat{q}u}$ (4)

is the similarity function between $\hat{q}$ and $\hat{p}(y)$, called Bhattacharyya coefficient.
To locate the object correctly in the image, the distance in (3) must be minimized, which is equivalent to maximize (4). The box center is initialized at a location $\hat{y}0$ which is the box center in the previous image frame. The probabilities $\{\hat{\rho}u(\hat{y}0)\}$ u=1...m are computed and using linear Taylor approximation of (4) around these values:

$\rho[\hat{p}(y), \hat{q}] \approx 1\ 2 \sum_{u=1}^{m} \sqrt{pu(y)\hat{q}u} + Cc2$

$\sum_{i=1}^{n} k\,(|y - xi|^2)\ wi\ k\,(|y - xi|^2)$ (5)

Where

$wi = \sum_{u=1}^{m} \sqrt{\hat{q}u/pu(y)}\ \delta[b(xi) - u]$ (6)

As the first term of (5) is independent of y, the second term of (5) must be maximized. The maximization of this term may be accomplished by employing Algorithm 1.
Maximizing Bhattacharyya coefficient $\rho[\hat{p}(y), \hat{q}]$
Input: The target model $\{\hat{q}u\}$u=1... m and its location $\hat{y}0$ in the previous frame.

2.1 MeanShift Algorithm

1. Initialize the center of the box at $\hat{y}0$ position,

2. Compute $\{\hat{p}u(\hat{y}0)\}$ u=1...m using (2) and evaluate $\rho[\hat{p}(\hat{y}0), \hat{q}]$ using (4).

3. Compute the weights {wi}i=1...n according to (6).

4. Compute the next location of the target candidate according to (7).

5. Compute $\{\hat{p}u(\hat{y}1)\}$u=1...m using (2) and evaluate $\rho[\hat{p}(\hat{y}1), \hat{q}]$ using (4).

6. If $\|\hat{y}1 - \hat{y}0\| < \epsilon$ Stop .Otherwise set $\hat{y}0 \leftarrow \hat{y}1$ and go to Step 2.

the MeanShift algorithm [1], which yields the following update:

$$\hat{y}1 = \frac{\sum_{i=1}^{n} xi\ wi\ g\ (||\hat{y}0 - xi||^2)}{\sum_{i=1}^{n} wi\ g\ (||\hat{y}0 - xi||^2)}$$

where g(x) =-k' (x) and k(x) is kernel with Epanechnikov profile. The complete algorithm [1] is summarized in algorithm 1.

3. The proposed approach

Original MeanShift Algorithm just can track the objects for short times (2-3 minutes).
But in long time videos like sport videos and other surveillance videos because of the changing of shape and color of the object original MeanShift loses the object.
so we developed original MeanShift Algorithm and created an Adaptive MeanShift approach which is more robust against changing of shape, brightness and color histogram. Adaptive MeanShift updates object model at periods of frames, like any 10 or 20 frame. so the correspondence between model and candidate Color histogram will be increased and probability of missing object will be decreased. to prevent noise influence in model histogram we eliminated the background pixels in both model and candidate kernel.
Also this approach is robust to tracking fast moving object (as soccer players, moving ball, fast moving car). If the proposed method missed the target objects, it can redetect the missed object through the target model updating process.

Process Model of Adaptive MeanShift Algorithm is given in fig. 1.

ACSIJ Advances in Computer Science: an International Journal, Vol. 3, Issue 3, No.9 , May 2014
ISSN : 2322-5157
www.ACSIJ.org

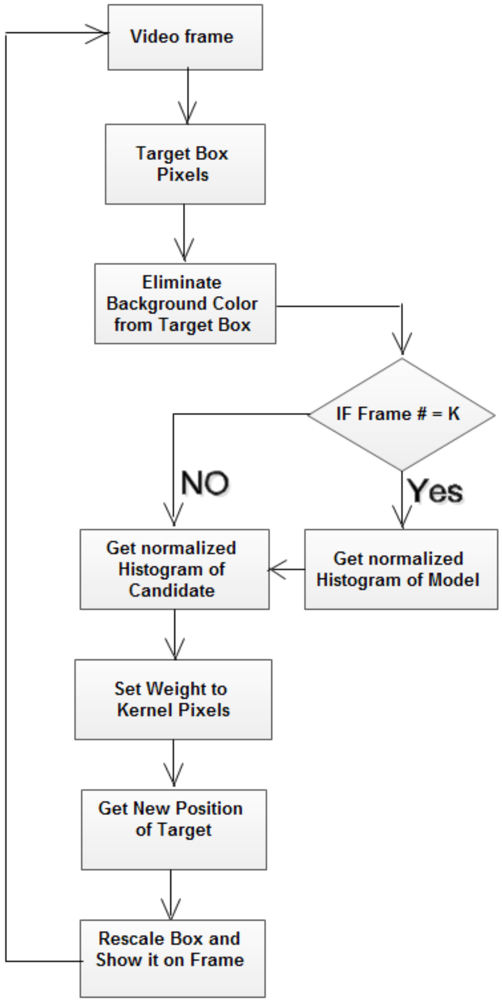


Fig. 1 Process Model of Adaptive MeanShift

3.1 Adaptive MeanShift Algorithm

1. Initialize model box.
2. Filter Background pixels.
3. Compute {qu(y)}, u = 1, . . . ,m Get weighted color Histogram of Target Object.
4. update the model every k frame. the parameter K is relative to application and we propose it can be 10 to 20.
5. Compute {pu(y)}, u = 1, . . . ,m (Get weighted color Histogram of candidate Object)
6. Compute weights wi, i = 1…. n.
7. Apply MeanShift: Compute new location z as

$$Z=\frac{\sum_{i=1}^{n} xi\ wi\ g\ (||\hat{}y0 - xi\ ||^{2}\ )}{\sum_{i=1}^{n}\ wi\ g\ \ (||\hat{}y0 - xi\ ||^{2}\ )}$$

8. While _(p(z), q) < _(p(y), q), do z ← 1/2 (y + z).
9. If kz − yk is small enough, stop and restart algorithm for next frame from step2.
10. Else, set y ← z and go to Step 2.

As we see in this approach to eliminate back color from model box we need to apply a filter to eliminate back ground pixels in achieved histograms.
In any MeanShift loop, the Histogram of Target model has been updated just like Candidate Histogram.

3.2 Filter BackGround pixels

there is many ways to Filter back ground colors from object box .one of them is comparing box pixels with primary/early background pixels at the same region. So the similar pixels assumed as background and other assumed as object.

4. Experimental Result

To compare the result of original MeanShift and proposed algorithm We apply the both methods to one real soccer match. Input image sequences are obtained by HD camera with 640*480 frame size, RGB 24 bit and frame rate 29 frame/sec. fig. 2 shows experimental result in 3 shot. In fig. 2 the first column is the result of the original MeanShift and the second column is the result of the suggested algorithm.

To compare the accuracy rate of the both methods (original MeanShift and proposed algorithm) we calculated the original MeanShift accuracy versus proposed method in small scale soccer players tracking for a long time (22500 frames) and got significant results . The test process of the both method implemented in same situation.

The average accuracy rate of original MeanShift in soccer players tracking was 81.50 and for our proposed method was 98.2 %. This accuracy difference between the both methods is because of the fact that the proposed method was missing the target objects less than original MeanShift.

Table 1 shows the accuracy rate for each target that has been tracked by original MeanShift Approach and Table 2 shows the accuracy rate for each target that has been tracked by CamShift [7] Approach (a visual object tracking method based Hue histogram).Table 3 consists the accuracy rate for each target that has been tracked by Adaptive MeanShift Approach (proposed Approach).

ACSIJ Advances in Computer Science: an International Journal, Vol. 3, Issue 3, No.9 , May 2014
ISSN : 2322-5157
www.ACSIJ.org

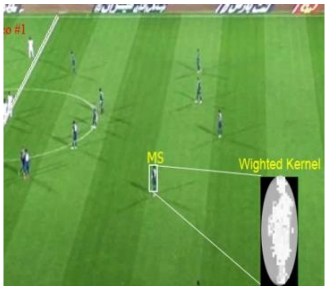
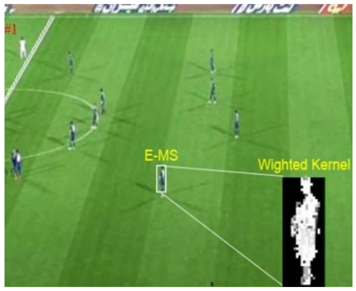
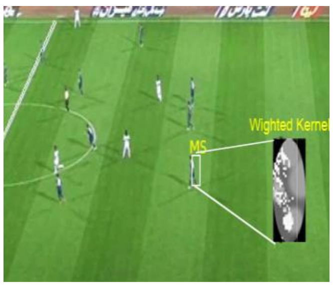
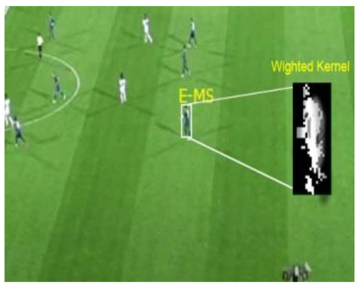
Original MeanShift	Adaptive MeanShift
 (Frame 10) Kernel get back ground pixels	 (Frame 10) Kernel is isolated
 (Frame 220) The Probablity of missing target is high	 (Frame 220) The Probablity of missing target is low
 (Frame 1600) Target is Missed because of changing target shape	 (Frame 1600) AMS continues locates target without noise

Fig. 2 comparative result of original MeanShift and proposed method

Table 1: Accuracy rate in targets tracking by Original Mean Shift

Player Box	***Correct Tracked Frames***	***Incorrect Tracked Frames***	***Tested Sequence Frames Count***	***Accuracy Rate***
1	18688	3812	22500	83 %
2	17222	5278	22500	76 %
3	18303	4197	22500	81 %
4	18700	3800	22500	83 %
5	17556	4944	22500	78 %
6	17980	4520	22500	79 %
7	19109	3391	22500	84 %
8	18925	3575	22500	84 %
9	18780	3770	22500	83 %
10	19003	3497	22500	84 %
Average accuracy				81.50 %

Table 2: Accuracy rate in targets tracking by CamShift Algorithm

Player Box	***Correct Tracked Frames***	***Incorrect Tracked Frames***	***Tested Sequence Frames Count***	***Accuracy Rate***
1	16550	5950	22500	73 %
2	17122	5378	22500	76 %
3	17052	5448	22500	75 %
4	16256	6244	22500	72 %
5	16888	5612	22500	75 %
6	16903	5597	22500	75 %
7	16289	6211	22500	72 %
8	17045	5455	22500	75 %
9	16699	5801	22500	74 %
10	17134	5366	22500	76 %
Average accuracy				74.35 %

Table 3: Accuracy rate in targets tracking by proposed method

Player Box	*Correct Tracked Frames*	*Incorrect Tracked Frames*	*Tested Sequence Frames Count*	*Accuracy Rate*
1	22336	164	22500	99 %
2	22250	250	22500	98 %
3	22196	304	22500	98 %
4	22352	148	22500	99 %
5	22259	241	22500	98 %
6	22346	154	22500	99 %
7	22334	166	22500	98 %
8	22093	407	22500	97 %
9	22035	465	22500	97 %
10	22334	156	22500	99 %
Average accuracy				98.20 %

5. Conclusion

MeanShift algorithm is a Kernel Tracking approach based on color histogram modeling. Because of changing of the color and shape of target model, it cannot track the object as much as possible.
In this paper an adaptive and robust object tracking method based on MeanShift is presented, compared with original MeanShift method and shown the result of comparing. This approach in soccer players tracking as small scale video objects was 98.2 % whereas the average accuracy rate of the original MeanShift algorithm was 81.5 % and result for CamShift method was 74.35 %.

This new approach can be used in video objects tracking systems for any environment like sports, surveillance and military systems. Robustness in tracking small scale objects in video surveillance systems is one of the most important advantages of the proposed method.

References

[1]. D. Comaniciu, V. Ramesh, P. Meer, “Kernel-Based Object Tracking”, IEEE Trans. on Pattern
Analysis and Machine Intelligence, vol.25, No. 5, 2003.

[2]. C. Zhang, Y. Qiaoy, E. Fallonz, C. Xu, “An improved CamShift algorithm for target tracking in video surveillance” , in 9th. IT & T Conference, 2009, Paper 12.

[3]. Y. Cui, S. Samarasekera, Q. Huang, M Greienhagen,"Indoor Monitoring Via the Collaboration Between a Peripheral Senson and a Foveal Sensor," IEEE Work-shop on Visual Surveillance, Bombay, India, 1998.

[4]. Y. Bar-Shalom and T. Fortmann, Tracking and Data Association. Academic Press, 1988

[5]. N. OULD-DRIS ,A. GANOUN ,R. CANALS ," IMPROVED OBJECT TRACKING WITH CAMSHIFT ALGORITHM" , in IEEE International Conference, 2006, Vol 2.

[6]. V. Karavasilis, C. Nikou, A. Likas "Visual Tracking by Adaptive Kalman Filtering and Mean Shift", in Computer Science, Volume 6040, 2010, pp 153-162.

[7]. N. Ould-Dris, A. Ganoun, R. Canal,"Improved Object Tracking With Camshift Algorithm", IEEE ICASSP 2006, Volume 2, 14-19 May 2006 .