International Journal of Research in Engineering and Science (IJRES)
ISSN (Online): 2320-9364, ISSN (Print): 2320-9356
www.ijres.org Volume 6 Issue 2 ǁ 2018 ǁ PP. 16-22

Object Tracking with Similar Background and Color Histogram

Yuyang Zhang;Xiankun Sun; Meiyun Hao
(Shanghai University Of Engineering Science,College of electronic and electrical engineering ，Shanghai Songjiang 201600 ）
Corresponding Author: Yuyang Zhang

ABSTRACT: In view of the boundary effect of the object tracking algorithm based on the correlation filter and the low tracking success rate in complicated cases, a object tracking algorithm based on similar background and color histogram is proposed. First of all, through the cosine similarity, selecting the background area with high similarity to the object as the negative sample and training the correlation filter template and reduce the boundary effect. At the same time，we employ a color histogram basedBayes classifier to distinguish between the object and the background in complex cases. Finally, weprovide both quantitative and qualitative comparison of our approach with state-of-thearttrackers on the OTB-50 and OTB-100 video sets and the result improvesthat our tracker can effectively improve the accuracy and success rate of tracking, superior state-of-the-art tracking algorithms .

KEYWORDS: correlation filter; background information; Bayes classifier; color histogram;

Date of Submission: 09-02-2018 Date of acceptance: 26-02-2018

I. Introduction

In recent years, object tracking has become an attracting area of computer vision. Many excellent object tracking algorithms are proposed, especially some object tracking algorithms based on correlation filters such as KCF[1]、DSST[2]、Staple[3]、SRDCF[4]、C-COT[5]. The excellent tracking results of these correlation filter-based tracking algorithms in OTB-50[6]and OTB-100[7] demonstrate that the correlation filter can improve the performance of object tracking under complex conditions such as illumination changes, fast motion and scale changes. However, there are two problems with most object tracking algorithms based on correlation filters. First, CF trackersare affected by the boundary effect; Secondly, most of them can not effectively use the object's color information to track.

Most existing object tracking algorithms based on correlation filters train the correlation filters based on periodic hypotheses. Taking the initial object as a positive sample, a negative sample generated by the shift operation of the positive sample by a cyclic matrix is used. Finally, the correlation filter template is trained by using positive and negative samples in the frequency domain. This method based on the assumption of extended negative samples to train the correlation filter can effectively train and detect using Fast Fourier Transform (FFT). However, since the negative samples are implicitly generated by the cyclic matrix, the correlation filters will be affected by the boundary effect of the cyclic matrix and the tracking performance of the correlation filter will be worse. In addition, since the negative sample is a false sample constructed by the positive sample, it is not a true negative sample collected from the video, so that the correlation filter can hardly distinguish between background and goal.

In order to solve the boundary effect caused by the circular matrix, Galoogahi proposed the CFLB[8]algorithm, which uses the augmented matrix to construct a relatively negative sample to solve the boundary effect. Although this method can effectively reduce the influence of the boundary effect, the correlation filter tracking effect is still poor when the object and the background are similar because it is still constructed by negative samples. Henriques proposed in the SRDCF[4] algorithm to reduce the effect of boundary effects by reducing the response of non-object areas through spatially regularized response areas. The main disadvantage of this approach is that even in the Fourier domain, the regularized goal requires a large amount of computation. In addition, careful tuning of a set of hyper-parameters in order to form normalized weights can have an impact on tracking performance. In addition, existing object tracking algorithms based on correlation filters, such as KCF and DSST algorithms, use a single HOG feature to track the object, increase the risk of tracking drift specifically when background is similar toobject. Others such as SRDCF, Staple Although the C-COT algorithm proposed to use color feature, color histogram and HOG[11] feature fusion for object tracking, it does not consider the robustness and scale invariance of color feature and color histogram feature in scale change are poor, This can lead to failure when tracking a object with a large scale change.

Therefore, in order to solving the solving the boundary effect caused by the circular matrix of correlation filter and improve the accuracy and success rate of object tracking under complex conditions by the effective employing of the color information of the object.This paper presents a object tracking with a similar background and color histogram. First, selecting the background areas close to the object or object as the negative samples by cosine similarity.Using the real background information instead of negative samples from circular matrixto solve the boundary effect of the correction filter, improving the accuracy and success of tracking the object in the complex background. Second, adding a color tracker. We employ the color histogram and Baye classifier to tracking object. Finally, Using result form correlation filter and result from Baye classifer and color histogram tofind the final position of the object.

II. Based On The Correlation Filter Object Tracking

2.1 Position tracking

The existing object tracking algorithm based on the correlation filter is generally as follows: Firstly, the initial position of the object in the image is obtained, positive and negative samples are obtained by shifting the initial region of the object by the cyclic matrix, and the two-dimensional Gaussian model is constructed by taking the geometric center of the initial region as the peak, and simultaneously extracting the features of the object. Then, Gaussian Newton method is used to calculate the correction filter template based on the two-dimensional Gaussian model and the object features. The correlation filter template to meet an output minimum error function：

$$E(H)=\left\|Y-\sum_{i=1}^{d}H^{i}*F^{i}\right\|_{2}^{2}+\sum_{i=1}^{d}\left\|H^{i}\right\|_{2}^{2} \quad (1)$$

$F^{d}\in\mathbf{R}^{m*n}$ is object characteristic, $H^{d}\in\mathbf{R}^{m*n}$ is correlation filter template, $E(H)$ is the desired output of the goal. Eq. 1 can be identically expressed as a ridge regression objective in the spatial domain:

$$E(H)=\left\|Y(\mathrm{j})-\sum_{i=1}^{d}H^{i}F^{i}\left[\Delta\tau_{j}\right]\right\|_{2}^{2}+\lambda\sum_{i=1}^{d}\left\|H^{i}\right\|_{2}^{2} \quad (2)$$

F^{i} represents the object candidate area. $\Delta\tau_{j}$ is cyclic matrix, by $F^{i}\Delta\tau_{j}$,positive and negative samples are obtained. When lower the $E(H)$, the more accurate the object tracking.

The Gaussian model is established for the positive sample area and the features are selected to train the correlation filters:

$$H^{i}=\frac{A_{t}}{B_{t}}=\frac{\sum_{i=1}^{d}GF^{i}}{\sum_{i=1}^{d}FF^{i}+\varepsilon} \quad (3)$$

H_{t}^{i} is the most suitable correlation filter. When H_{t}^{i} obtained, in the next frame of picture, the characteristics F_{t+1}^{i} of the candidate sample areas are introduced into the filter H_{t}^{i} to obtain the response set $\mathrm{y}_{i}(i=0,1\cdots,n)$:

$$y_{i}=F^{-1}(H_{t}F_{t+1}^{i}) \quad (4)$$

F^{-1} is inverse Fourier transform. Transforming the space domain to the frequency domain improves the speed of computation. The maximum response $\max(\mathrm{y}_{i})$ is the new center point of the object movement.

2.2 Scale estimation

According to the size of the previous frame Z_{t-1}^{s}, multiplied by the scale series n, to obtain different sizes of the scale set Z_{t}^{s}, into the correlation filter H , get the response set $\mathrm{y}_{s}(i=0,1\cdots,n)$, select the most suitable scale.

The principle of size selection is:

$$s^{n}P\times s^{n}R \quad (5)$$

The s is a scale factor fixed for the initial value, P and R represent the width and height of the previous frame, and the n is the scale series range.

Optimal scale selection formula:

$$y_{\mathrm{s}}=F^{-1}\left\{\frac{A_{i}^{scale}z_{t}^{s}}{B_{i}^{scale}+\lambda}\right\} \quad (6)$$

2.3 Template update

In the new frame, according to the maximum response point y_{i+1}, y_s, the new center location P_{i+1} and scale Z_{s+1} are obtained. According to the new central location eq(1), the filter template is updated. The formula is as follows.

$$A_t = \eta G_t F_t + (1-l)A_{t-1} \qquad (7)$$

$$B_t = \eta F_t F_t^* + (1-l)B_{t-1} \qquad (8)$$

η is the learning rate, A_t and B_t are the molecules and denominator of the filter template of the current frame, A_{t-1} and B_{t-1} are the molecules and denominator of the filter template of the previous frame.

III. Object Tracking Based On Similar Background And Color Histogram

3.1 Similar backgrounds

The existing object tracking algorithm based on correlation filter selects the object area as the positive sample, and then obtains the negative samples through the cyclic matrix, and trains the correlation filter template. In order to reduce the influence of the boundary effect caused by the cyclic matrix, we select the object area. Then, according to the formula (nine), we use cosine similarity to get the background area with high similarity with the object area as the negative sample of training correlation filter template.

$$\omega = \frac{F \bullet P}{\|F\| \bullet \|P\|} \qquad (9)$$

F is the object positive sample, P is the negative sample from the background region. The related filter templates based on similar backgrounds also satisfy the output minimum error function:

$$E(H) = \left\| Y(\mathrm{j}) - \sum_{i=1}^{d} H^i F^i \left[\Delta\tau_j\right] P^i \right\|_2^2 + \lambda \sum_{i=1}^{d} \left\| H^i \right\|_2^2 \qquad (10)$$

Where P^i is the background negative sample, $F^i\left[\Delta\tau_j\right]P^i$ is the set of positive and negative samples for training, H is the relevant filter template, and the training in the frequency domain obeys the formula (11):

$$H^i = \frac{A_t}{B_t} = \frac{\sum_{i=1}^{d} G F^i P^{d-i}}{\sum_{i=1}^{d} F F^i + \varepsilon} \qquad (11)$$

The formula for calculating position response and scale response is

$$y_{\text{center}} = F^{-1}(\mathrm{H}_{new} F_{t+1}^i) \qquad (12)$$

$$y_{scale} = F^{-1}(\mathrm{H}_{new} Z_{t+1}^{\mathrm{s}}) \qquad (13)$$

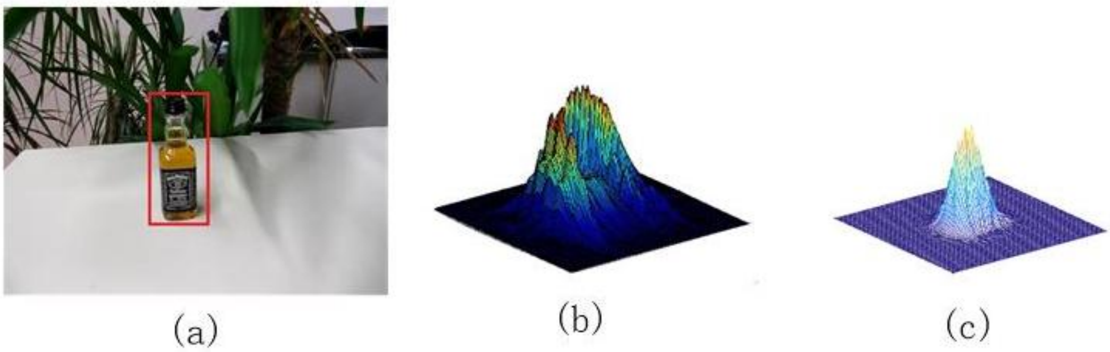

Fig 1 (a)object initial area(b)object response affected by the boundary effect (c) our response by using similar background

3.2 Object tracking based on color histgram and baye classifIer

To distinguish object pixels x ∈ O from surrounding background pixels, we employ a color histogram based Bayes classifier on the input image Z. Given a rectangular object region O (i.e. initial bounding box annotation or current tracker hypothesis) and its surrounding region S , we apply Bayes rule to obtain the object likelihood at location x as:

$$p(X \in O|\mathrm{O,S,b}) \approx \frac{P(b|X \in \mathrm{O})\mathrm{P}(\mathrm{X} \in \mathrm{O})}{\sum_{\Omega\in\{O,S\}} P(X|(\mathrm{X} \in Z)\mathrm{P}(\mathrm{X} \in Z)} \quad (14)$$

O is object area ; S is background area; Z is input image。Using Baye rule compute each pix likehoood for object.
The probability of the pixel value in the candidate area is counted, and the area with the highest probability value is selected as the position of the object in the next frame. The formula is as follows:

$$P(\mathrm{O}_{(t,i)}) = \sum_{\mathrm{x}\in O} p(X \in O|\mathrm{O,S}) \quad (15)$$

Finally we employfollowing formula to update the dimension of the object hypothesis.

$$P(\mathrm{O}_{(t,i)})=\eta P(\mathrm{O}_{(t,i)}) + (1-\eta)P(\mathrm{O}_{(1:t-1,i)}) \quad (16)$$


(a) (b) (c) (d)

Fig2 (a) The red box is the candidate area. (b),(c),(d) are candidate area as the object probability.The probability that blue box(b) is the object is 0.73 .The probability that blue box(c) is the object is 0.51.The probability that blue box(d) is the object is 0.34.

3.3 Joint tracking by correlation filter and Bayes classifier

First, tracking by correlation filter, according to the initial position of the object, the feature is extracted to build the Gaussian model G_{new}, and the correlation filter template H_{new} is calculated. According to the formula (11) to calculate the position tracking candidate area in the response of each point y_{center}.
Second, tracking by Bayes classifier, Bayes classifier should be learnt from a set of examples taken from each image, including the correct position as a positive example. Employing eq15,to get the probability $P_{(\mathrm{i,j})}$ for each pixel as the target.
Finally, the final location center is:

$$Y_p = \lambda_1 y_{\text{center}} + \lambda_2 P_{(\mathrm{i,j})} \quad (17)$$

Y_p is that the point p is the probability of the final target location center computed by correlation filter. $P_{(\mathrm{i,j})}$ is tht the point (i,j) is the probability of the final target location center computed by Baye classifier . λ_1 and λ_2 are parameters chosen on a validation set.

IV. Evalution

4.1parameter settings and test video set

The experiment used by the computer CPU i5-6500, 3.2GHZ, GTX1050. The program parameters are set as follows: The candidate box is 2.5 times the initial object size. Learning rate of 0.01. The OTB-50[6]and OTB-100[7]test sets were used for testing, which consisted of complex images such as blurred photography, fast motion, large scale changes, obvious changes of illumination, complex background environment, Challenge conditions. Moreover, OTB-50[6]and OTB-100[7] not only have dozens of short-term object tracking but also long-term object tracking. The Comparison algorithm are DSST, Staple, SRDCF, MDnet[9], C-COT and TCNN[10]. The former tracking algorithms not only runs fast, but also have high tracking accuracy and good robustness for scale changes. In addition, the evaluation index selected in this experiment is overlap and tracking success rate (Success rate).

Table 1 Comparison algorithms

algorithm name	source
DSST	Cvpr2014
Staple	Cvpr2015
SRDCF	Cvpr2015
TCNN	Cvpr2016
MDnet	Cvpr2016

C-COT	Cvpr2016

4.2Results on OTB-50 and OTB-100

In order to evaluate the impact of the correlation filter (SF) modeled on the similar backgroundon the object tracking, this section modifies the DSST, Staple and SRDCF which also use the correlation filter and HOG features, adds the background modeling, The video of OTB-50 and OTB-100 is selected for testing and comparative analysis of the results. The test results are shown in Table 2:

Table 2 Comparison of the success rate (overlap threshold 0.5) and overlap precision (OP) of the four algorithms with and without similar background

	DSST	SF-DSST	Staple	SF-Staple	SRDCF	SF-SRDCF
OTB-50	48.0	48.2	74.5	76.9	77.4	80.9
OTB-100	47.3	47.8	70.5	73.5	74.2	75.8
Mean OP	32.1	35.5	44.6	46.3	45.7	48.1
FPS	35.7	30.4	33.3	27.4	23.5	15.7

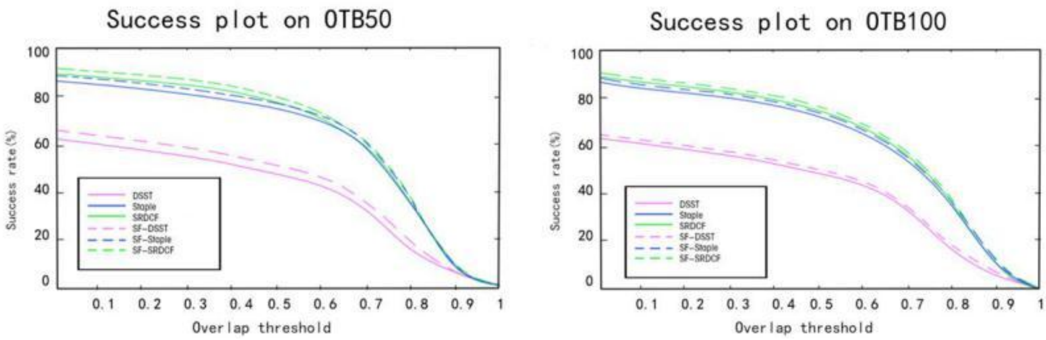

Figure 7 (a) and (b) show the success rate curves of the three algorithms for adding background information and the original algorithm on OTB-50 and OTB-100

From table (b) and Figure 7, we can see that the background information is increased. Although the FPS of DSST, Staple and SRDCF algorithms are reduced, the success rate and overlap accuracy of the success rate and the overlap accuracy are improved compared with the original algorithm. Especially for Staple and SRDCF, the success rate and coverage rate of tracking are improved greatly due to the border effect of these two algorithms using circular matrix combined with ridge regression training template. Therefore, it can be concluded that by training the negative samples for selecting the real background area, the influence of the boundary effect on the object tracking can be suppressed, and the success rate and the overlapping accuracy of the object tracking can be improved.

4.3Algorithm Comparison and Analysis

Fifteen representative videos were selected based on OTB-50 and OTB-100 for testing. The 15 videos include eleven complex cases: light change (IV), scale change (SV), occlusion (OCC), DEF Motion Blur, MB, IPR, OPR, OV, BC and LR. The proposed method is compared with the six tracking algorithms DSST, Staple, SRDCF, TCNN, MDnet and C-COT.

Algorithm Comparison Test video and test results are as follows:

Table 3 videos from OTB-50 and OTB-100

	IV	SV	OCC	DEF	MB	FM	IPR	OPR	OV	BC	LR
Bird1				√		√			√		
Blurbody		√		√	√	√	√				
Board		√			√	√		√	√	√	
Bolt			√	√			√	√			
Bolt2				√						√	

CarScale		√	√			√	√	√			
Couple		√									
Coke	√			√		√		√		√	
Crowds	√			√						√	
David3			√	√				√		√	
Dudek		√	√	√		√	√	√	√	√	
FaceOcc2	√		√				√	√			
Liquor	√	√	√		√	√		√	√	√	
Panda		√	√	√			√	√	√		√
Skiing	√	√		√			√	√			√

Table 4 success rate for each trackers

complex cases	DSST	Staple	SRDCF	TCNN	MDnet	C-COT	Ours
IV	0.438	0.515	0.542	0.679	0.672	0.674	0.692
SV	0.471	0.534	0.532	0.646	0.665	0.662	0.661
OCC	0.541	0.567	0.572	0.678	0.672	0.682	0.632
DEF	0.466	0.511	0.522	0.614	0.623	0.685	0.719
MB	0.421	0.498	0.532	0.612	0.711	0.699	0.673
FM	0.432	0.535	0.511	0.658	0.681	0.679	0.683
IPR	0.401	0.478	0.503	0.655	0.681	0.692	0.654
OPR	0.405	0.492	0.527	0.652	0.659	0.634	0.646
OV	0.389	0.502	0.519	0.621	0.675	0.688	0.671
BC	0.391	0.478	0.495	0.625	0.637	0.653	0.695
LR	0.328	0.411	0.312	0.487	0.491	0.457	0.472

From the table (4), we can see that the success rate of tracking is higher than that of the other six in the four test videos of light change (IV), object deformation (DEF), fast motion (FM) and background complexity (BC) Kind of algorithm And the success rate of the four algorithms is the best one in the aspects of scale transformation (SV), out of plane rotation (OPR), object evanescence (OV) and low resolution object (LR). However, in the three cases of occlusion (OCC), motion blur (MB) and in-plane rotation (IPR), the tracking success rate is low.

Table (5) FPS for each trackers

	Staple	SRDCF	TCNN	MDnet	C-COT	Our
Avg FPS	30.4	31.5	14.5	12.2	10.1	30.3

Table (5) shows the average tracking rate under 15 groups of videos of this algorithm and other algorithms. Since the proposed algorithm only uses HOG features and does not use convolutional neural networks and convolution features with TCNN, MDnet and C-COT, the real-time performance is better than these three algorithms. However, compared with Staple and SRDCF algorithms, this paper improves the correlation filter.

Adding background modeling and color histogram-based Bayes classifier increases the computational complexity of tracking and is therefore lower than Staple and SRDCF in real-time.

V. Conclusion

This paper presents a object tracking algorithm based on background information and color histogram. The algorithm selects the background similar to the object cosine as the negative sample to train the correlation filter to reduce the boundary effect and improve the performance of the correlation filter. And the experimental results on OTB-50 and OTB-100 show that the use of real can effectively improve the relative accuracy and success rate of tracking the object in complex cases. The algorithm also combines color histogram with Bayes classifier to track the color of object. In the OTB-50 and OTB-100, 15 types of tracking video were selected and compared with the current six tracking algorithms DSST, Staple, SRDCF, TCNN, MDnet and C-COT. The successful tracking rate of model and color histogram is higher than those of the six algorithms for illumination change, object deformation, fast motion and background complicated object tracking.

References

[1]. Henriques J F, Rui C, Martins P, et al. High-Speed Tracking with Kernelized Correlation Filters[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2015, 37(3):583-596.
[2]. Danelljan M, Häger G, Khan F S, et al. Accurate Scale Estimation for Robust Visual Tracking[C]// British Machine Vision Conference. 2014:65.1-65.11.
[3]. Bertinetto L, Valmadre J, Golodetz S, et al. Staple: Complementary Learners for Real-Time Tracking[C]// Computer Vision and Pattern Recognition. IEEE, 2016:1401-1409.
[4]. M. Danelljan, G. Hager, F. Shahbaz Khan, and M. Felsberg. Learning spatially regularized correlation filters for visual tracking. In ICCV, pages 4310–4318, 2015.
[5]. Danelljan M, Robinson A, Khan F S, et al. Beyond Correlation Filters: Learning Continuous Convolution Operators for Visual Tracking[C]// European Conference on Computer Vision. Springer, Cham, 2016:472-488.
[6]. Wu Y，Lim J，Yang M H.　Online object tracking: a benchmark［C］．　IEEE Conference on Computer Vision and Pattern Ｒecognition.　IEEE Computer Society，2013: 2411-2418.
[7]. Wu Y, Lim J, Yang M H. Object Tracking Benchmark[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2015, 37(9):1834.
[8]. Galoogahi H K, Sim T, Lucey S. Correlation filters with limited boundaries[C]// Computer Vision and Pattern Recognition. IEEE，2015:4630-4638.
[9]. Nam Hyeonseob, Han B. Learning Multi-domain Convolutional Neural Networks for Visual Tracking[C]// Computer Vision and Pattern Recognition. IEEE, 2016:4293-4302.
[10]. Nam H, Baek M, Han B. Modeling and Propagating CNNs in a Tree Structure for Visual Tracking[J]. 2016.
[11]. N. Dalal and B. Triggs. Histograms of oriented gradients for human detection. In CVPR, volume 1, pages 886–893. IEEE,2005.

Yuyang Zhang "Object Tracking with Similar Background and Color Histogram" International Journal of Research in Engineering and Science (IJRES), vol. 06, no. 02, 2018, pp. 16–22.