Adaptive tracking algorithm based on target moving speed

K W Liu*, D W Wu, Q S Xie, Z Y An

1 College of Information and Electronic Engineering, Shandong Technology and Business University, Yantai, Shandong, 264005, China
2 College of Traffic, Northeast Forestry University, Harbin Heilongjiang 150040, China
3 College of Computer Science and Technology, Shandong Technology and Business University, Yantai 264005, China
4 Shandong Co-Innovation Center of Future Intelligent Computing, Yantai 264005, China

*Corresponding author’s e-mail: 1107273630@qq.com

Abstract. To solve the problem of tracking failure caused by fast object movement, we proposed an adaptive tracking algorithm based on object movement speed on the basis of ECO-HC (ECO for hand-crafted feature version). First, correlation filtering method was used to obtain the distance of the maximum response point of adjacent frames. Then the speed of the object was judged by the distance of the maximum response point of the adjacent frame. Then the number of scale filter factors and interval frames updated by the template were adjusted on the basis of the object’s speed. The comparative experiment on the VOT2016 benchmarks showed that our tracker improved the accuracy and robustness of object tracking.

1. Introduction

Object tracking is an important research direction in the field of computer vision. It is widely used in video monitoring, intelligent transportation, precise guidance, unmanned driving, human-computer interaction and other fields[1-3].

In recent years, due to the application of correlation filtering and fast Fourier transform in the object tracking field, the performance of the object tracking algorithm has been greatly improved. Therefore, the object tracking method based on correlation filter has become one of the research hotspots in the object tracking field[4-5]. The Minimum Output Sum of Squared Error filter (MOSSE) proposed by Bolme et al.[6] applied the relevant filter to the tracking field for the first time. The calculation amount of MOSSE algorithm is small, so it is fast. However, due to the limited number of samples used by MOSSE and the single features used, MOSSE has a poor tracking effect when the object appearance changes. Kernelized Correlation Filters (KCF) proposed by Henriques et al.[7] introduced cyclic matrix and multi-channel HOG features. KCF can effectively enrich the number of samples and improve the discriminating ability of the filter, but it has poor robustness in dealing with the change of target scale. The Discriminative Scale Space Tracking Filter (DSST) proposed by Danelljan et al.[8] effectively improves the robustness of target Tracking by introducing adaptive target Scale adjustment by feature pyramid and 3d correlation Filter. Efficient Convolution Operators for Tracking (ECO) proposed by Danelljan et al.[9] solved the problems of complex calculation and over-fitting of previous Tracking algorithms in terms of model size, training set size and model update.
However, ECO's model was updated at fixed intervals. When the target moved rapidly, it was easy to cause inaccurate tracking of template drift target. Therefore, how to solve the problem of poor robustness when the object is moving fast is still one of the research emplacements of object tracking.

On the basis of ECO-HC, we studied the sparse update strategy and proposed an adaptive tracking algorithm based on the object moving speed. We adjust the interval of sparse update adaptively according to the moving speed of the object, and at the same time, we estimate the number of scale filter factors in the adaptive adjusting scale according to the moving speed of the object. Our tracker can effectively reduce the possibility of drift of tracking template and improve the robustness of object tracking through adaptive adjustment.

2. ECO-HC

The EOC algorithm analyzes the previous correlation filter tracking algorithm and summarizes several reasons that affect the performance of correlation filter algorithm. For the problem of excessive model parameters, ECO used Factorized convolution operator to reduce the parameters of the model. Aiming at the problem of redundant training set samples of ECO using the Generative Sample Space Model to classify samples and control samples of class number. In order to improve the robustness property of object tracking and prevent filter overfitting, ECO adopted the sparse model update method, which updated the model once every Ns frame. Ns=5 set in ECO’s paper. Ultimate objective function:

$$E(f) = \sum_{l=1}^{L} \pi_l \| S_f \{ \mu_l \} - y_l \|^2 + \sum_{d=1}^{D} \| \phi f d \|^2$$

(1)

Where, L is the number of sample components, each component has a weight $\pi_m$ and a feature $\mu_m$, $S_f$ is the Score of the filter, and D is the filter dimension.

3. Our Approach

3.1. Adaptive update interval

When the moving speed of the object is too fast, both the amplitude of the object's feature change and the probability of the object's feature change will be increased. If the object tracking model cannot be updated in time, it is likely to cause tracking failure. When the speed of the object is too small, if the update frequency is too fast, it is easy to cause filter overfitting, which will lead to the decrease of the robustness of the object tracking.

Therefore, in this paper, the velocity $v$ of current frame $t$ is determined by calculating the distance $x_t$ between the center point of current frame $t$ and $t-1$, the distance $x_{t-1}$ between $t-1$ and $t-2$, and the average distance $x_{t-2}$ between $t-2$ and $t-3$. When the speed of frame $t$ is too fast, the update interval of the model is reduced. When the speed of frame $t$ is too small, the update interval of the model is increased. The relationship between model update interval $Ns'$ and speed $v$ in frame $t$ is shown in equation (2):

$$Ns' = \begin{cases} 
N_0 + \varepsilon (\ln(\frac{v}{7 * Ns^{t-1}})), & v > 1 \\
N_0 + \varepsilon (\ln(\frac{v}{7 * Ns^{t}})), & v \leq 1 
\end{cases}$$

(2)

In equation (2), $N_0$ is the update interval of the initial model, $\varepsilon$ is the step function, $Ns^{t-1}$ is the update interval of the model at frame $t-1$. The velocity $v$ of the current frame $t$ is $v = \frac{x_t + x_{t-1} + x_{t-2}}{3}$. 

2
3.2. Number of adaptive scale filter factors
In the process of object tracking, if the scale estimation is not accurate, it is easy to pollute the tracking model with the object background. When the object is moving quickly, the possibility of increasing the range of the object size increases. At this time, if the number of scale filter factor \( q \) in the scale estimation is fixed, it is easy to lose the object. If the scale filter factor is too much, the computational complexity will be increased. In this paper, according to the speed \( v \) of frame \( t \), the number of adaptive adjusting scale filtering factors \( q \) effectively improves the accuracy of object tracking:

\[
q = \begin{cases} 
\tau_1, & v > 30 \\
\tau_2, & v \leq 30 
\end{cases}
\]  

(3)

In equation (3), \( \tau_1 = 17 \) and \( \tau_2 = 20 \).

4. Experiments
The experimental hardware platform is Intel(R) Core(TM) i7-6700 CPU @ 3.40 GHz, 64-bit Windows10 professional system, matlab2016a. The algorithm in this paper is tested on the VOT2016\[10\] dataset and compared with the current 6 more advanced tracking algorithms(ECO-HC\[9\], Staple\[11\], SiameseFC-AlexNet\[12-13\], MDNet\[16\], SRDCF\[15\], DSST\[8\]).

4.1. Experimental evaluation indexes
In VOT2016, the main evaluation criteria include Accuracy, Robustness and EAO (Expected Average Overlap). The greater the Accuracy and EAO values, the better the tracking algorithm performance. The smaller the Robustness, the more robust the tracking. Since \( R_s \) is a minus function of mean-time-between-failures\( (M) \)[16], the larger the value of \( R_s \) here means that the tracer is more robust. The EAO measures the expected no-reset overlap of a tracker run on a short-term sequence.

4.2. Experimental parameter setting and analysis
In our tracking, HOG\[17\] and CN are adopted as the main tracking features, where HOG cell size=6, CN cell size=4, sample learning rate =0.012, sample component \( L \)=50, and initial model update interval \( N_0 \)=5.

Table 1 shows the overall evaluation of 7 advanced trackers on the VOT2016 dataset. The data in bold in the table is the best data.

| EAO  | Ours | ECO-HC | Staple | SiameseFC-AlexNet | MDNet | SRDCF | DSST |
|------|------|--------|--------|-------------------|-------|-------|------|
| Robustness (Weighted mean) | 0.33 | 0.32 | 0.30 | 0.24 | 0.26 | 0.25 | 0.18 |
| Accuracy (Weighted mean)   | 0.54 | 0.53 | 0.54 | 0.53 | 0.54 | 0.53 | 0.53 |

It can be seen from table 1 that the EAO of the tracking results of the algorithm in this paper in the VOT2016 dataset is 0.33, which is 1% higher than ECO-HC, 3% higher than Staple, and 9%, 7%, 8% and 15% higher than SiameseFC-AlexNet, MDNet, SRDCF and DSST, respectively. The Robustness mean value of the tracking results in the VOT2016 dataset is 19.13, which is 2.27, 4.76, 10.67, 1.95, 9.19 and 25.68 higher than that of ECO-HC, Staple, SiameseFC-AlexNet, MDNet, SRDCF and DSST, respectively. As can be seen from table 1, the Robustness of the algorithm in this paper is better than that of the other six algorithms. The accuracy of the algorithm in this paper also has an excellent performance.
It can be seen from Figure 1, 2 and 3 that under baseline test mode, the overlap rate of each object tracking algorithm decreases as the number of frames increases. The overlap rate of the algorithm in this paper is higher than that of the other 6 algorithms at each stage of tracking.

5. Conclusions
We propose an adaptive tracking algorithm based on the moving speed of the object. The experimental results prove that the method of adaptively adjusting the model update interval according to the moving speed of the object adopted by our tracker effectively reduces the possibility of template drift in the tracking template. We compare our approach with state-of-the-art trackers on VOT2016 benchmarks. It is also proved that our tracker can effectively reduce the probability of contamination of the tracking template by adjusting the number of scale filter factors according to the object's moving speed.

Acknowledgments
This work was supported by from technology innovation fund of Shandong Technology and Business University (Nos. 2018yc038).

References
[1] Yin, H P. et al. (2016) A Review of Target Detection and Tracking Based on Vision. Acta Automatica Sinica, 42: 1466-1489.
[2] Lu, H. C. et al. (2018) Visual Object Tracking: a Survey. Pattern Recognition and Artificial Intelligence, 31: 61-76.
[3] Meng, L. et al. (2019) Pattern Recognition and Artificial Intelligence. Acta Automatica Sinica, 45: 1244-1260.
[4] Zhu, Z., Wu, W., Zou, W., et al., 2018, End-to-end Flow Correlation Tracking with Spatial-temporal Attention. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Wellington, 548-557.

[5] Dai, K., Wang, D., Lu, H., et al., 2019, Visual Tracking Via Adaptive Spatially-regularized Correlation Filters. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Los Angeles, 4670-4679.

[6] Bolme, D.S., Beveridge, J.R., Draper, B.A., et al., 2010, Visual Object Tracking Using Adaptive Correlation Filters. 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. San Francisco, 2544-2550.

[7] Henriques, J.F., Caseiro, R., Martins, P., et al. (2015) High-speed Tracking with Kernelized Correlation Filters. IEEE Transactions on Pattern Analysis & Machine Intelligence., 37: 583-596.

[8] Danelljan, M., Häger, G., Khan, F.S., et al., 2014, Accurate Scale Estimation for Robust Visual Tracking. British Machine Vision Conference. Nottingham, 1-5.

[9] Danelljan, M., Robinson, A., Khan, F.S. et al., 2016, ECO: Efficient Convolution Operators for Tracking. European Conference on Computer Vision. Cham, 6638-6646.

[10] Battistone, F., Santopietro, V., Petrosino, A., 2016, The Visual Object Tracking Vot2016 Challenge Results. European Conference on Computer Vision. Cham, 777–823.

[11] Bertinetto, L., Valmadre, J., Golodetz, S., et al., 2016, Staple: Complementary Learners for Real-time Tracking. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, 1401-1409.

[12] Bertinetto, L., Valmadre, J., Henriques, J.F., et al., 2016, Fully-convolutional Siamese Networks for Object Tracking. European Conference on Computer Vision. Cham, 850-865.

[13] Krizhevsky, A., Sutskever, I., Hinton, G.E., 2012, Imagenet Classification with Deep Convolutional Neural Networks. Advances in Neural Information Processing Systems. 1097-1105.

[14] Nam, H., Han, B., 2016, Learning Multi-domain Convolutional Neural Networks for Visual Tracking. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, 4293-4302.

[15] Danelljan, M., Hager, G., Khan, F.S., et al., 2015, Learning Spatially Regularized Correlation Filters for Visual Tracking. 2015 IEEE International Conference on Computer Vision. Santiago, 1-5.

[16] Kristan, M., Matas, J., Leonardis, A., et al., 2015, The Visual Object Tracking Vot2015 Challenge Results. IEEE International Conference on Computer Vision. Santiago, 227-251.

[17] Dalal, N., Triggs, B., 2005, Histograms of Oriented Gradients for Human Detection. 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. San Diego, 886-893.