Adaptive Model Updating Correlation Filter Tracker with Feature Fusion

Jingjing Shao[1], Lei Xiao[1], and Zhongyi Hu[2](✉)

[1] College of Computer Science and Artificial Intelligence, Wenzhou University, Wenzhou, China
194511981406@stu.wzu.edu.cn, xiaolei@wzu.edu.cn
[2] Intelligent Information Systems Institute, Wenzhou University, Wenzhou, China
hujunyi@163.com

Abstract. Aiming at the poor accuracy of a single feature in the challenging scenarios, as well as the failure of tracking caused by partial or complete occlusion and background clutter, a correlation filter tracking algorithm based on feature fusion and model adaptive updating is proposed. On the basis of the background-aware correlation filter, the proposed algorithm firstly introduces the CN feature and integrates with the HOG feature to improve the accuracy of tracking. Then, the Average Peak-to-Correlation Energy (APCE) is introduced, and the results of object tracking are fed back to the tracker through the ratio changes. The tracker is adaptively updated, which improves the robustness of the algorithm to occlusion and background clutter. Finally, the proposed algorithm is experimented on the self-build ship dataset. The experimental results show that the algorithm can adapt well to complex scenes, such as object occlusion and background clutter. Compared to the state-of-the-art trackers, the average precision of the proposed tracker is improved by 2.3%, the average success rate is improved by 2.9%, and the average speed is about 18 frames per second.

Keywords: Object tracking · Correlation filter · Feature fusion · Model updating · APCE

1 Introduction

Visual tracking is one of the key technologies in the computer vision field, and it has wide application prospects in video surveillance, human-computer interaction, medical diagnosis and so on. With the continuous deepening of research [1–4], visual tracking has made some progress in stages, but it is difficult to accurately locate the tracked object due to the interference factors such as partial

J. Shao—Student.

This study was financially supported by the Natural Science Foundation of Zhejiang Province Major Project (LZ20F020004), the Science and Technology Plan Major Science and Technology Projects of Wenzhou (ZY2019020), the Natural Science Foundation of Zhejiang Province (LY16F020022), and Wenzhou Science and Technology Planning Project (S20180017) of China.

Y. Peng et al. (Eds.): PRCV 2020, LNCS 12306, pp. 255–265, 2020.
https://doi.org/10.1007/978-3-030-60639-8_22

or complete occlusion, rotation, motion blur and so on. Therefore, there are still great challenges in building a robust tracker.

In recent years, mainstream tracking algorithms are divided into two categories, namely correlation filter and deep learning [5–8]. Among them, correlation filters (CFs) [9–12] algorithm is a classical algorithm for object tracking, which is favored by researchers because of its fast speed [13]. Bolme et al. [14] proposed MOSSE filter, which was the first time to introduce CF into object tracking with extremely fast speed. Based on that, Henriques et al. [15] introduced a Gaussian kernel function for acceleration, and extended the single-channel grayscale feature to the multi-channel Histogram of Oriented Gradient (HOG) to improve the tracking accuracy. Li et al. [16] proposed a scale adaptive multi-feature fusion tracker, adding the HOG feature and CN feature [17] on the basis of gray feature to improve the overall performance of the tracker. Besides, Danelljan et al. [18] proposed three-dimensional filter, one-dimensional scale filter and two-dimensional translation filter. This precise scale estimation method can be combined with any other tracking algorithm without scale estimation, and won the first place in the VOT2014 [19] competition. Since the methods based on CF are affected by the boundary effect, in order to overcome this problem, Danelljan et al. [20] added spatial regularization to suppress it, so that the search area can be expanded, and Gauss-Seidel was used to solve the filter to simplify the calculation. The models of the above algorithms are not effective for tracking targets with deformation and motion blur. Bertinetto et al. [21] complemented the HOG feature and color histogram feature, which was robust to motion blur, illumination and deformation, and added scale to the HOG to improve the accuracy of the tracker. Galoogahi et al. [22] used the negative samples generated by real shifts to include a larger search area and real background, and proposed an ADMM-based optimization method to reduce the computation. In recent years, deep learning-based methods have become more and more popular. Wang et al. [23] proposed a lightweight end-to-end training network, DCFNet, which simultaneously learns deep features and performs filtering processes. Wu et al. [24] learned the multi-level same-resolution compressed (MSC) features, which effectively incorporate both deep and shallow features for efficient online tracking, in an end-to-end offline manner.

The above methods have achieved good tracking effects in terms of accuracy and robustness. However, in the case of complex scenes, such as partial or complete occlusion, background clutter, etc., the problem of tracking loss will still occur. For this reason, in the framework of background-aware correlation filters (BACF), the following improvements have been made: (1) The use of excellent features is the basis for accurate tracking. A single feature have defects in accuracy. Considering the method of feature fusion to improve the accuracy of tracking, and adding the CN feature on the basis of HOG feature. (2) In order to better solve the problem of tracking failure caused by occlusion, the APCE method is introduced in the online update stage to adaptively update model to improve the robustness of the tracker.

2 The Tracker

The proposed algorithm is based on the background-aware correlation filter, which combines the features of HOG and CN, and introduces APCE [25] for adaptive model update, thereby improving the tracking algorithm's robustness to occlusion and background clutter. The framework of the proposed algorithm is shown in Fig. 1.

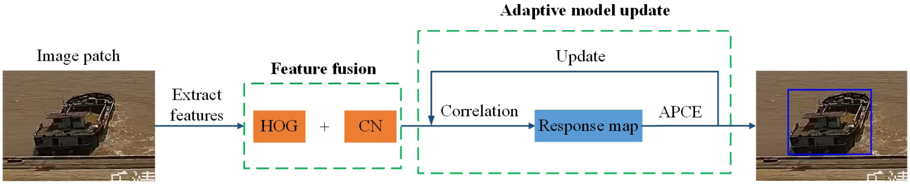


Fig. 1. Framework of the Proposed Algorithm. For each input image patch, first extract the HOG and CN features from the prediction area, fuse the two, and then obtain the corresponding response map through correlation filtering. APCE is introduced to adaptively update the model to determine whether to update the model at the current frame. Finally, update the model at the appropriate frame.

2.1 Background-Aware Correlation Filters

The background-aware correlation filters significantly increases the number of samples based on the traditional CF method and improves the sample quality through cropping operator, and has good real-time tracking. Therefore, We make improvements on the basis of background perception related filters in order to improve the accuracy of the algorithm. The basic objective function [26] of CF is:

$$E(\mathrm{h}) = \frac{1}{2}\|\mathrm{y} - \sum_{k=1}^{K} \mathrm{h}_k \star \mathrm{x}_k\|_2^2 + \frac{\lambda}{2}\sum_{k=1}^{K}\|\mathrm{h}_k\|_2^2 \tag{1}$$

where y is the desired output response, x_k and h_k represents the kth channel of the vectorized image and filter respectively. λ is a regularization constant, and $\star$ is the spatial correlation operator. Equation 1 is the form of a single sample. When we use D cyclic samples, it becomes the following form:

$$E(\mathrm{h}) = \frac{1}{2}\sum_{j=1}^{T}\|\mathrm{y}(j) - \sum_{k=1}^{K} \mathrm{h}_k{}^{\top}\mathbf{P}\mathrm{x}_k[\triangle\tau_j]\|_2^2 + \frac{\lambda}{2}\sum_{k=1}^{K}\|\mathrm{h}_k\|_2^2 \tag{2}$$

the size of the sample x changes from D to T, which is much larger. Use the larger sample to generate a cyclic sample. $[\triangle\tau_j]$ is the circular shift operator. Then we need to extract the middle part of the size D. This step is replaced by

$\mathbf{P}$, which is a $D \times T$ binary matrix. $\mathbf{P}$ can be calculated in advance, and it is a constant matrix.

Taking advantage of the fast solution of the cyclic samples in the frequency domain, the expression is transformed into the frequency domain. The formula is as follows:

$$\begin{aligned} E(\mathrm{h}, \hat{\mathrm{g}}) &= \frac{1}{2}\|\hat{\mathrm{y}} - \hat{\mathbf{X}}\hat{\mathrm{g}}\|_2^2 + \frac{\lambda}{2}\|\mathrm{h}\|_2^2 \\ s.t. \quad \hat{\mathrm{g}} &= \sqrt{T}(\mathbf{F}\mathbf{P}^\top \otimes \mathbf{I}_K)\mathrm{h} \end{aligned} \tag{3}$$

where $\hat{\mathbf{X}} = [\mathrm{diag}(\hat{\mathrm{x}}_1)^\top, ..., \mathrm{diag}(\hat{\mathrm{x}}_K)^\top]$ and $\hat{}$ refers to the Discrete Fourier Transform (DFT) of a signal. $\hat{\mathrm{g}}$ is a $KT \times 1$ auxiliary variable and $\hat{\mathrm{g}} = [\hat{\mathrm{g}}_1^\top, ..., \hat{\mathrm{g}}_K^\top]$. h is defined as $\mathrm{h} = [\mathrm{h}_1^\top, ..., \mathrm{h}_K^\top]$ of size $KD \times 1$. The DFT of one-dimensional signal α is expressed as $\hat{\alpha} = \sqrt{T}\mathbf{F}\alpha$, $\mathbf{F}$ is an $T \times T$ orthogonal Fourier transform matrix. $\mathbf{I}_K$ is a $K \times K$ identity matrix ($\mathbf{P}\mathbf{P}^\top = \mathbf{I}$), $\otimes$ refers to the Kronecker product.

Finally, the optimization solution of Eq. 3 is mainly used to put the constraint term into the optimization function by using the Augmented Lagrangian Method (ALM) [27].

$$\begin{aligned} L(\hat{\mathrm{g}}, \mathrm{h}, \hat{\zeta}) = &\frac{1}{2}\|\hat{\mathrm{y}} - \hat{\mathbf{X}}\hat{\mathrm{g}}\|_2^2 + \frac{\lambda}{2}\|\mathrm{h}\|_2^2 \\ &+ \hat{\zeta}^\top(\hat{\mathrm{g}} - \sqrt{T}(\mathbf{F}\mathbf{P}^\top \otimes \mathbf{I}_K)\mathrm{h}) \\ &+ \frac{\mu}{2}\|\hat{\mathrm{g}} - \sqrt{T}(\mathbf{F}\mathbf{P}^\top \otimes \mathbf{I}_K)\mathrm{h}\|_2^2 \end{aligned} \tag{4}$$

where $\hat{\zeta} = [\hat{\zeta}_1^\top, ..., \hat{\zeta}_K^\top]$ and μ is a penalty factor. Equation 4 can be solved iteratively using Alternating Direction of Method of Multipliers (ADMM) [27] technology, and $\hat{\mathrm{g}}$ and h are optimized and solved separately.

2.2 Feature Fusion

The single feature has defects in accuracy. Considering the method of feature fusion to improve tracking accuracy, CN feature is added to the basis of HOG feature.

Color-Naming (CN). CN is an 11-dimensional color space feature that maps the 3-dimensional color features of the RGB space to black, blue, brown, gray, green, orange, pink, purple, red, white, and yellow. CN can separate objects of different colors, and it can distinguish objects and backgrounds with significant color difference and similar texture shapes.

The CN adopts the adaptive color attribute algorithm to map the RGB space to the 11-dimensional color space with obvious discrimination to obtain the 11-dimensional color feature vector, which is then mapped into the 10-dimensional subspace, reducing the dimension from 11 to 10 dimensions. Therefore, HOG and CN are serially combined into $\mathbf{M}$, assuming that the vectors of HOG and CN are $\mathbf{H}_i (i = 1, 2, ..., 31)$ and $\mathbf{C}_j (j = 1, 2, ..., 10)$, respectively. $\mathbf{H}_i$ and $\mathbf{C}_j$ represent the i-th channel HOG and the j-th channel CN of the image respectively, then

$\mathbf{M} = [\mathbf{H}_1\ \mathbf{H}_2\ ...\ \mathbf{H}_{31}\ \mathbf{C}_1\ \mathbf{C}_2\ ...\ \mathbf{C}_{10}]$, the 31-channel HOG and the 10-channel CN extracted from the training image patch are serially fused to obtain the 41-channel $\mathbf{M}$.

HOG emphasizes the edge information of the image, while CN focuses on color information. The two features are complementary and improve the performance of the filter. Although the idea is simple, the performance improvement is very promising.

2.3 Adaptive Model Update

In the process of model tracking, the appearance and scale of the object will change. Figure 2 shows the object occlusion during tracking. If the tracker is updated at Fig. 2(b), the model may drift or even lose the object. In order to adapt to the changes of the tracking model, the maximum response value and the APCE are introduced to determine when the model will be updated. The formula is as follows:

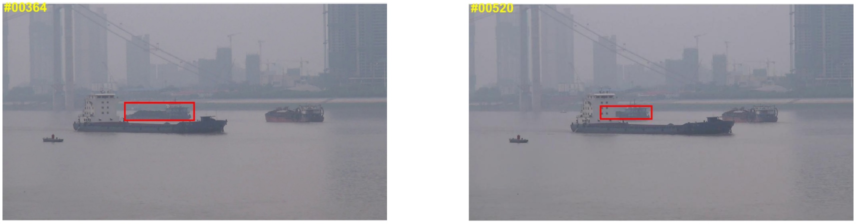
$$\mathrm{APCE}(t) = \frac{|\mathbf{F}_{\max}(t) - \mathbf{F}_{\min}(t)|^2}{\mathrm{mean}(\sum\limits_{w,h}(\mathbf{F}_{w,h}(t) - \mathbf{F}_{\min}(t))^2)} \tag{5}$$

where $\mathbf{F}_{\max}$, $\mathbf{F}_{\min}$ and $\mathbf{F}_{w,h}$ represent the maximum response, minimum response and current frame response value, respectively. When the target is occluded or lost, APCE will suddenly decrease. In this case, the model is not updated to avoid model drift. Only when APCE and $\mathbf{F}_{\max}$ are greater than the historical mean in a certain proportion, the model is updated, greatly reducing the model drift.

The online updating strategy of the model is still the same linear interpolation method as the traditional CF:

$$\hat{\mathrm{x}}^{(f)}_{model} = (1-\eta)\ \hat{\mathrm{x}}^{(f-1)}_{model} + \eta\ \hat{\mathrm{x}}^{(f)} \tag{6}$$

where η if a learning rate, $\hat{\mathrm{x}}^{(f)}_{model}$ indicates the model at frame f.



(a) No occlusion

(b) Occlusion

Fig. 2. Two frames of a ship sequence on the self-build ship dataset. (a) is the 343th frame image of a ship video sequence, with the object in the red bounding box and not occluded by other ships or objects; (b) is the 520th frame image of a ship video sequence, with the object in the red bounding box and occluded by other ships. (Color figure online)

3 Experiments

In order to verify the reliability of the proposed tracker AMUMF (Adaptive Model Updating Correlation Filter Tracker with Feature Fusion), the self-build ship dataset was used for evaluation, and compared with 6 excellent correlation filter trackers, such as KCF, SAMF, STAPLE, STAPLE_CA, SRDCF, BACF.

3.1 Experimental Setup and Methodology

The experimental environment of the algorithm is MATLAB R2016a on Windows system. All experiments are completed on a desktop computer equipped with an Intel Core i5-9400 CPU at 2.90 GHz.

Experimental Dataset. The experimental data used in this research is a self-build ship dataset, which contains 60 ship video sequences. In order to better evaluate and analyze the advantages and disadvantages of the tracking method, 11 attributes such as illumination variation (IV), scale variation (SV), occlusion (OCC), deformation (DEF), motion blur (MB), fast motion (FM), in-plane rotation (IPR), out-of-plane rotation (OPR), out-of-view (OV), background clutters (BC) and low resolution (LW) are used to annotate the sequence, so as to classify these sequences. Figure 3 shows the 60 ship video tracking sequences.

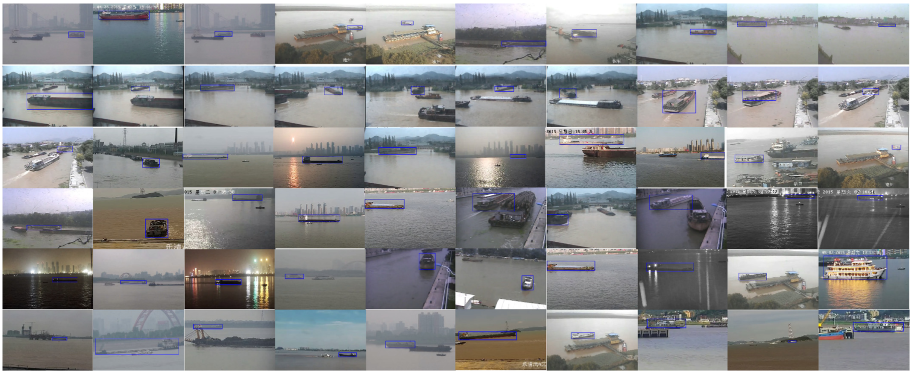


Fig. 3. Ship video tracking sequences. The blue box in the figure represents the tracked target. (Color figure online)

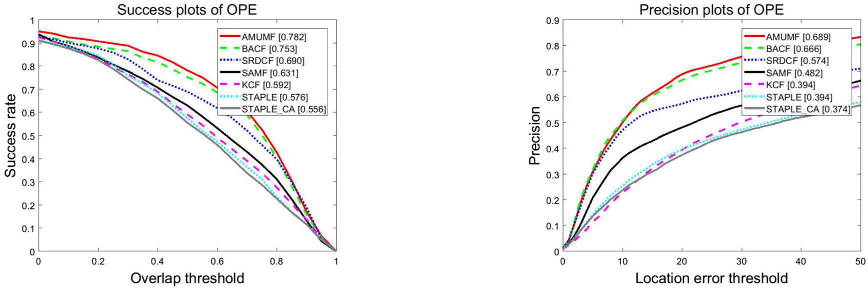
Parameter Settings. The specific parameters of the algorithm are set as: the thresholds of the maximum response and APCE in the adaptive model updating are 0.5 and 0.85, respectively. Other parameter settings are the same as the BACF.

3.2 Analysis

According to the evaluation method of the OTB [28], the one-pass evaluation (OPE) method is adopted. And there are two evaluation criteria selected, i.e. precision plot and success plot.

Table 1. Comparison of overall performance of 7 trackers on self-build ship dataset (/%). The best results are shown in bold, and the second-ranked is underlined.

	AMUMF	BACF	STAPLE_CA	SRDCF	STAPLE	SAMF	KCF
Success rate	**78.2**	<u>75.3</u>	55.6	69.0	57.6	63.1	59.2
Precision	**68.9**	<u>66.6</u>	37.4	57.4	39.4	48.2	39.4



(a) success plot

(b) precision plot

Fig. 4. Comparison of success plot and precision plot of 7 trackers on self-build ship dataset.

Quantitative Analysis. We tested AMUMF on the self-build ship dataset and compared with other 6 trackers. Table 1 shows the success rate and precision of AMUMF and other 6 trackers. It can be seen that the success rate and precision of AMUMF are 78.2% and 68.9%, respectively, and the best results are obtained. This is 2.9% and 2.3% higher than BACF without feature fusion and adaptive model update. Figure 4 shows the corresponding precision curve and success rate curve of the 7 tracker. Figure 5 shows a comparison of success plot based on video attributes. It can be seen that AMUMF performs well at low resolution (LR), background clutter (BC), scale variation (SV), in-plane rotation (IPR), occlusion (OCC), out-of-plane rotation (OPR). Especially under OCC, the success rate of AMUMF is 5.3% higher than BACF.

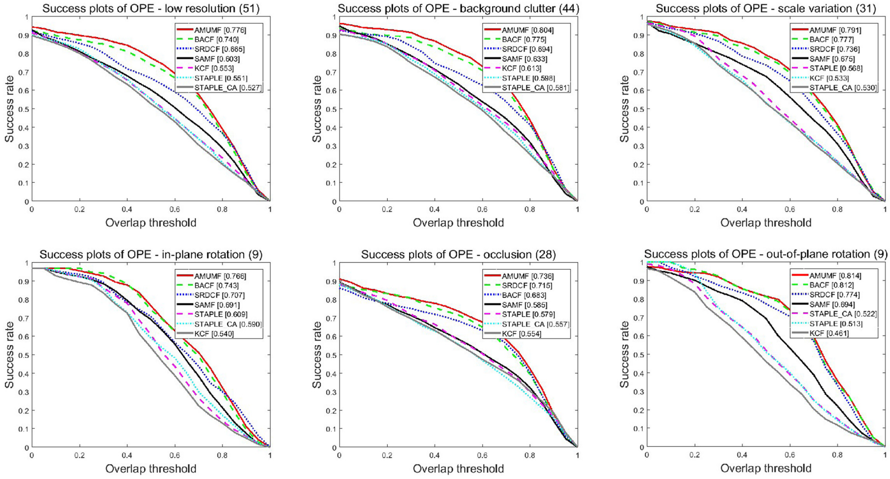


Fig. 5. Attribute-based evaluation. Comparison of success plot of 7 trackers on self-build ship dataset. AMUMF outperforms other trackers in these 6 video attributes.

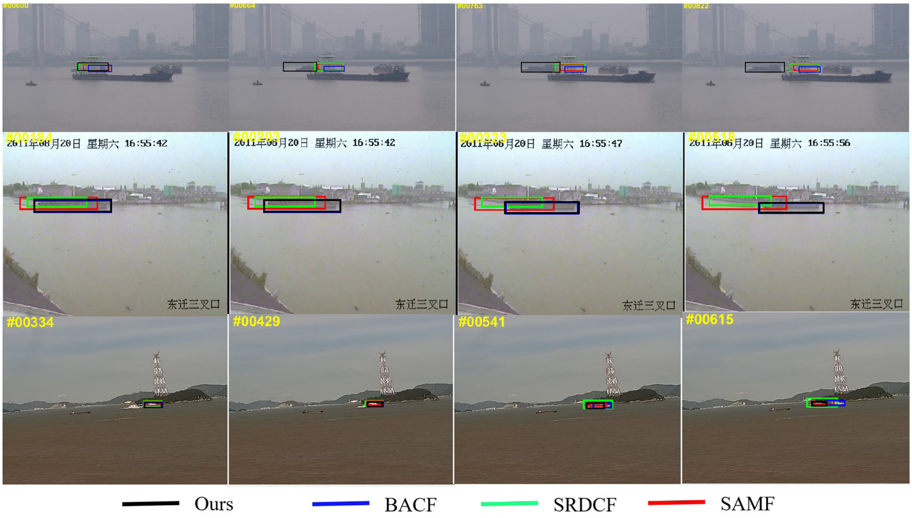

#00600
#00664
#00763
#00822
#00184
#00203
#00323
#00518
2011年08月20日 星期六 16:55:42
2011年08月20日 星期六 16:55:42
2011年08月20日 星期六 16:55:47
2011年08月20日 星期六 16:55:56
东迁三叉口
#00334
#00429
#00541
#00615
Ours
BACF
SRDCF
SAMF

Fig. 6. Comparison of tracking results of 4 types of trackers. Each row represents a sequence.

Qualitative Analysis. We selected 3 representative video sequences for qualitative analysis, and compared AMUMF with BACF, SRDCF and SAMF as shown in Fig. 6. All three sequences under LR. In addition, the 1st sequence under OCC and BC. Only AMUMF continues to track accurately, other trackers are lost. The 2nd sequence under IPR, AMUMF and BACF can continue to track accurately, other trackers produce drift. In the 3rd sequence, two ships with the same appearance intersect. In addition to BACF, other trackers can track the target, but only AMUMF can accurately locate.

4 Conclusion

In the framework of background-aware correlation filter, a correlation filter tracker based on feature fusion and model adaptive updating is proposed. Two kinds of features are extracted, HOG and CN, and they are serially fused to obtain the final response map, so that the object is accurately located. We also introduce a high-confidence model updating to adaptively update tracking model, which effectively improves the robustness of the tracker to occlusion and background clutter. Experiments on the self-build ship dataset prove that the proposed AMUMF is superior to other trackers in terms of precision and success rate. In the future, we will further research the optimization of algorithms and how to improve the real-time tracking.

References

1. Xiao, L., Xu, M., Hu, Z.: Real-time inland CCTV ship tracking. Math. Probl. Eng. **2018**(2018), 1–10 (2018)
2. Xiao, L., Wang, H., Hu, Z.: Visual tracking via adaptive random projection based on sub-regions. IEEE Access **6**, 41955–41965 (2018)
3. Hu, Z., Xiao, L., Teng, F.: An overview of compressive trackers. Int. J. Signal Process. Image Process. Pattern Recogn. **9**(9), 113–122 (2016)
4. Hu, Z., Zou, M., Chen, C., Wu, Q.: Tracking via context-aware regression correlation filter with a spatial-temporal regularization. J. Electron. Imaging **29**(2), 023029 (2020)
5. Yun, S., Choi, J., Yoo, Y., Yun, K., Young Choi, J.: Action-decision networks for visual tracking with deep reinforcement learning. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 2711–2720 (2017)
6. Valmadre, J., Bertinetto, L., Henriques, J., Vedaldi, A., Torr, P.H.: End-to-end representation learning for correlation filter based tracking. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 2805–2813 (2017)
7. Ma, C., Huang, J.B., Yang, X., Yang, M.H.: Hierarchical convolutional features for visual tracking. In: Proceedings of the IEEE International Conference on Computer Vision, pp. 3074–3082 (2015)
8. Bertinetto, L., Valmadre, J., Henriques, J.F., Vedaldi, A., Torr, P.H.S.: Fully-convolutional Siamese networks for object tracking. In: Hua, G., Jégou, H. (eds.) ECCV 2016. LNCS, vol. 9914, pp. 850–865. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-48881-3_56
9. Mueller, M., Smith, N., Ghanem, B.: Context-aware correlation filter tracking. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 1396–1404 (2017)
10. Li, F., Tian, C., Zuo, W., Zhang, L., Yang, M.H.: Learning spatial-temporal regularized correlation filters for visual tracking. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 4904–4913 (2018)

11. Lukezic, A., Vojir, T., Cehovin Zajc, L., Matas, J., Kristan, M.: Discriminative correlation filter with channel and spatial reliability. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 6309–6318 (2017)
12. Tang, M., Yu, B., Zhang, F., Wang, J.: High-speed tracking with multi-kernel correlation filters. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 4874–4883 (2018)
13. Henriques, J.F., Caseiro, R., Martins, P., Batista, J.: Exploiting the circulant structure of tracking-by-detection with kernels. In: Fitzgibbon, A., Lazebnik, S., Perona, P., Sato, Y., Schmid, C. (eds.) ECCV 2012. LNCS, vol. 7575, pp. 702–715. Springer, Heidelberg (2012). https://doi.org/10.1007/978-3-642-33765-9_50
14. Bolme, D.S., Beveridge, J.R., Draper, B.A., Lui, Y.M.: Visual object tracking using adaptive correlation filters. In: 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pp. 2544–2550 (2010)
15. Henriques, J.F., Caseiro, R., Martins, P., Batista, J.: High-speed tracking with kernelized correlation filters. IEEE Trans. Pattern Anal. Mach. Intell. **37**(3), 583–596 (2014)
16. Li, Y., Zhu, J.: A scale adaptive kernel correlation filter tracker with feature integration. In: Agapito, L., Bronstein, M.M., Rother, C. (eds.) ECCV 2014. LNCS, vol. 8926, pp. 254–265. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-16181-5_18
17. Danelljan, M., Shahbaz Khan, F., Felsberg, M., Van de Weijer, J.: Adaptive color attributes for real-time visual tracking. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 1090–1097 (2014)
18. Danelljan, M., Häger, G., Khan, F., Felsberg, M.: Accurate scale estimation for robust visual tracking. In: British Machine Vision Conference, pp. 1–11. BMVA Press (2014)
19. Kristan, M., et al.: The visual object tracking VOT2014 challenge results. In: Agapito, L., Bronstein, M.M., Rother, C. (eds.) ECCV 2014. LNCS, vol. 8926, pp. 191–217. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-16181-5_14
20. Danelljan, M., Hager, G., Shahbaz Khan, F., Felsberg, M.: Learning spatially regularized correlation filters for visual tracking. In: Proceedings of the IEEE International Conference on Computer Vision, pp. 4310–4318 (2015)
21. Bertinetto, L., Valmadre, J., Golodetz, S., Miksik, O., Torr, P.H.: Staple: complementary learners for real-time tracking. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 1401–1409 (2016)
22. Kiani Galoogahi, H., Fagg, A., Lucey, S.: Learning background-aware correlation filters for visual tracking. In: Proceedings of the IEEE International Conference on Computer Vision, pp. 1135–1143 (2017)
23. Wang, Q., Gao, J., Xing, J., Zhang, M., Hu, W.: DCFNet: discriminant correlation filters network for visual tracking. arXiv preprint arXiv:1704.04057 (2017)
24. Wu, Q., Yan, Y., Liang, Y., Liu, Y., Wang, H.: DSNet: deep and shallow feature learning for efficient visual tracking. In: Jawahar, C.V., Li, H., Mori, G., Schindler, K. (eds.) ACCV 2018. LNCS, vol. 11365, pp. 119–134. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-20873-8_8
25. Wang, M., Liu, Y., Huang, Z.: Large margin object tracking with circulant feature maps. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 4021–4029 (2017)
26. Kiani Galoogahi, H., Sim, T., Lucey, S.: Multi-channel correlation filters. In: Proceedings of the IEEE International Conference on Computer Vision, pp. 3072–3079 (2013)

27. Boyd, S., Parikh, N., Chu, E., Peleato, B., Eckstein, J.: Distributed optimization and statistical learning via the alternating direction method of multipliers. Found. Trends Mach. Learn. **3**(1), 1–122 (2011)
28. Wu, Y., Lim, J., Yang, M.H.: Object tracking benchmark. IEEE Trans. Pattern Anal. Mach. Intell. **37**(9), 1834–1848 (2015)