

Channel pruning for visual tracking

Manqiang Che, Runling Wang, Yan Lu, Yan Li,
Hui Zhi, and Changzhen Xiong*

North China University of Technology, Beijing, China
xczkiong@163.com

Abstract. Deep convolutional feature based Correlation Filter trackers have achieved record-breaking accuracy, but the huge computational complexity limits their application. In this paper, we derive the efficient convolution operators (ECO) tracker which obtains the top rank on VOT-2016. Firstly, we introduce a channel pruned VGG16 model to fast extract most representative channels for deep features. Then an Average Feature Energy Ratio method is put forward to select advantageous convolution channels, and an adaptive iterative strategy is designed to optimize object location. Finally, extensive experimental results on four benchmarks OTB-2013, OTB-2015, VOT-2016 and VOT-2017, demonstrate that our tracker performs favorably against the state-of-the-art methods.

Keywords: Correlation filter · Deep feature · Channel pruning · Iterative optimization

1 Introduction

Visual tracking is one of the fundamental problems in computer vision. Tracking of objects or feature points plays a crucial role in real-time vision applications, such as traffic control, smart surveillance, human-computer interactions, to name a few. Even though significant progress has been made in this area, it is still a challenging problem due to fast motions, occlusions, deformations, illumination variations and etc.

Correlation Filter (CF) based trackers have attracted considerable attention due to the high computational efficiency. Feature representations such as grayscale templates [24], HOG [1] and Color Names (CN) [2] have successfully been employed in CF based trackers. Deep convolutional neural networks (CNNs) are also resorted to visual tracking for robust target representation [3,4,7]. Deep features based correlation filter can effectively increase the tracking accuracy [3,15,25], but the huge computational complexity limits their application. In order to solve this problem, Wang et al. [14] propose a real time tracker via convolutional channel reduction. ECO tracker [6] applies a combination of the deep features along with HOG and CN features to tracking task and proposes a generative sample space model for higher precision. It also introduces a factorized convolution operator to dramatically decrease parameters and an efficient

model update strategy to improve the speed. Then it obtains the highest tracking accuracy at that time, but the speed is far from real-time requirement.

We can find that deep features selected in ECO are not robust for fast motion and serious changes in appearance of objects on some videos, such as the objects in Figure 1. And ECO tracker adopts the fixed channel number for selected convolutional layers, which is not suitable for all tested video sequences. Consequently, this paper proposes a channel pruning tracker (CPT) via channel pruned model and feature maps. Experiments on popular datasets display that our proposed CPT has better robustness (see Figure 1 for visualized tracking results).

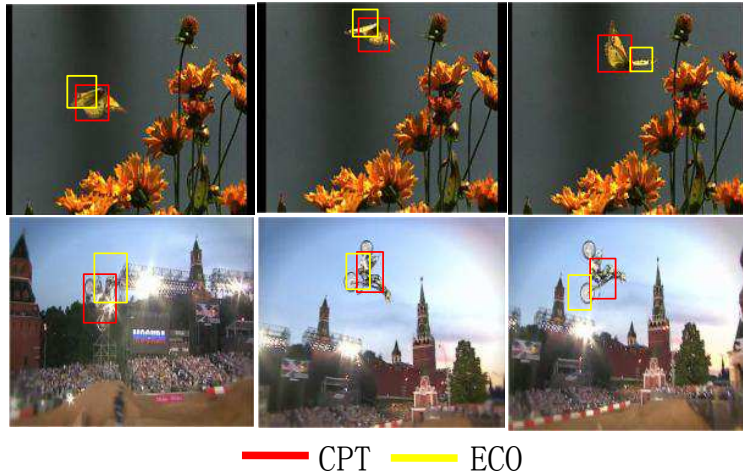


Fig. 1. Comparisons of tracking results with ECO. Example frames are from butterfly (top row) and motocross1 (bottom row) video sequences. CPT tracker with deeper features can handle such variations successfully, improving both the accuracy and robustness.

Our main contributions are four folds:

- Our work is the first attempt to apply channel pruned VGG model to visual tracking field. Thus CPT tracker can use more deep convolutional layers with rich semantic features and will not cause a decline in speed.
- An Average Feature Energy Ratio method is proposed to adaptively reduce the dimensions of convolution channels. It can effectively extract different dimensions of convolutional features for different video sequences.
- An adaptive iteration strategy is applied to adaptively terminate the optimization process of target location. It can further speed up the tracker without a precision reduction.

- We extensively validate our algorithm on four benchmarks, OTB-2013, OTB-2015, VOT-2016 and VOT-2017. Our CPT tracker performs favorably against state-of-the-art trackers.

2 Related Work

In this section, we briefly introduce trackers based on correlation filters and CNN accelerating methods related to our work.

CF based methods have shown superior performances on object tracking benchmarks [19,18,20,22]. The MOSSE tracker [24] learns a minimum output sum of squared error filter for fast tracking, making researchers fully realize the advantages of correlation filters in speed. Then several extensions have been put forward to substantially promote the tracking precisions including CSK method [26] based on intensity features, KCF approach [1] with HOG descriptors and CN tracker [2] using colour attributes, showing a remarkable tracking speed. Bertinetto et al. [9] propose a tracker based on HOG and colour histograms integration for targets appearance representation. Danelljan et al. [30] introduce a spatial regularization component to penalize the filter coefficients near the boundary regions to suppress the boundary effect.

As the surge deep learning, more and more state-of-the-art visual trackers have benefited from deep CNN model owing to its powerfulness in feature extraction. Ma et al. [4,13] extract hierarchical convolutional features from the VGG19 network [21] and combine three feature maps to correlation filter tracker. Danelljan et al. [3] learn a continuous convolution filter for tracking, with multi-scale deep features and hand-crafted features as HOG and CN, to account for appearance variations and considerably improve the tracking accuracy. In order to improve the speed of deep features based trackers, Wang et al. [14] make full use of multi-resolution deep features for precise location and remove the redundancy by reducing the channel number so as to obtain a practical speed. ECO tracker [6] introduces a factorized convolution operator to simplify the multi-channel filters of C-COT [3] and achieves a satisfactory tracking accuracy and speed. However, the fixed channel number for selected layers in ECO is not suitable for all tested video sequences.

There has been numerous work on accelerating CNNs [21] using channel pruning that removes redundant channels on feature maps. [28] regularizes networks to improve accuracy. Channel-wise SSL [28] prunes first few convolutional layers to reach high compression. Some model compression based methods [29,27] focus on pruning the fully connected layers. [23] proposes an inference-time approach to prune redundancy inter channels. Combining with tensor factorization, it obtains 5 $\times$ speed-up VGG16 model while with only 0.3% increase of error. It is worth mentioning that this work has achieved considerable results in the area of detection, but unfortunately has not been introduced into the visual tracking field. Consequently, we introduce the channel pruned VGG16 into the visual tracking field.

3 Proposed Algorithm

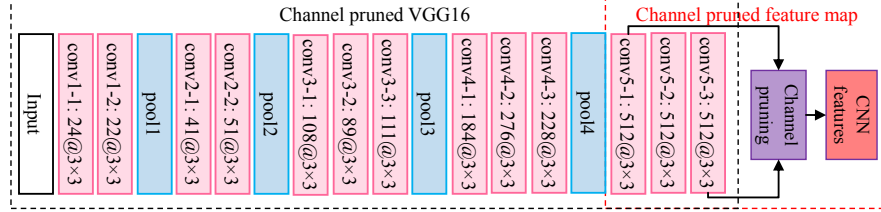


Fig. 2. Feature maps extraction framework using channel pruning.

3.1 Channel Pruned VGG Model

Earlier convolutional layers provide more spatial information, while the latter layers encode rich semantic features [4]. The ECO tracker has achieved high tracking accuracy by fusing the shadow spatial information (Conv1) and deep semantic features (Conv5) of VGG-M. Moreover, CFWCR [11] assigns larger weight for the feature map extracted from the Conv5 layer and gains better robustness in VOT-2016. DRT [25] uses the Conv4-3 layer of VGG16 and the Conv1 layer of VGG-M to obtain higher tracking accuracy. Thus, aiming for higher precision, more deep semantic features are needed. But it will inevitably cause a decline in tracking speed. Consequently, we introduce a pruned VGG16 model obtained in [23] to fast extract more deep semantic features. With an iterative two-step algorithm (LASSO regression and Least Square Reconstruction), channels of layers from the original VGG16 network are pruned to a desired number. It dramatically decreases the feature channels and accelerates the VGG16 model by $5\times$ speed-up in object detections [23]. For advantageous reason, we attempt to apply it to tracking field. As shown in Figure 2, pruned channels of VGG16 network from Conv1-1 to Conv4-3 layers are marked, e.g. channels of Conv1-1 layer have pruned from 64 to 24 dimensions. Additionally, the set part of the pruned VGG16 is Conv5 layer, whose feature maps are pruned with another novel method for more effective information. Details are described in section 3.2.

3.2 Channel Pruned Feature Map

In our work, the Conv5-1 and Conv5-3 from channel pruned VGG16 network are selected as our tracking layers for feature extraction, which are full of semantic information to handle large appearance changes. Then an Average Feature Energy Ratio method is utilized to prune the ineffective channels as the factorized convolution operator does in ECO.

For a new frame, the correlation filter based trackers acquire the search region according to the localization of the previous frame, followed by extracting the search regions features and obtaining the response map. As shown in Figure 3, the wanted feature map should have larger energy value of target and smaller that of other area in the search region. However, there exist large amount of features containing backgrounds information and make interferences to the tracking task. Moreover, a vast majority of features contain negligible energy. These latter two categories of features can hardly contribute to target localization, but cause a set number of calculations.

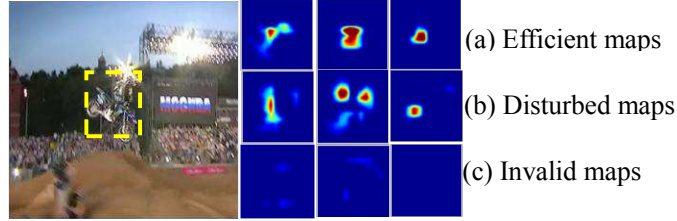


Fig. 3. Feature maps in different channels. The input search region image is from the challenging motocross1 sequence. The target is in yellow dotted bounding box. (a) Efficient features for tracking task. (b) Features with noises in search region. (c) Invalid features containing negligible energy.

Consequently, we explore a response map evaluation mechanism called the Average Feature Energy Ratio (AFER) method which is defined as

$$AFER_l^d = \frac{F_l^d(O)}{1 + F_l^d(S)} \quad (1)$$

Here, d indicates the d th dimension of features from l layer. $F_l^d(O)$, $F_l^d(S)$ denote the average feature energy of the object and the whole search region.

$$F_l^d(A) = \frac{\sum_{i,j}^{I,J} P(i,j)}{I \bullet J} \quad (2)$$

where I and J indicate the width, height of the region A . $P(i,j)$ is the value (energy) of the location (i,j) after convolution operation. AFER indicates the validity of the response maps and the confidence level of the tracking object. The larger the AFER value is, the more effective the acquired features are, while the smaller the AFER value is, the more background interferences exist. Consequently, we adaptively select convolution channels whose AFER is greater than the given threshold for target location to prune ineffective channel. On the other hand, we put forward the channel screening approach to exclude duplicate features from different layers while at the same channel. For the reason we consider that features from the neighboring hierarchical layers while at the same

channel are similar. Specifically, when there are same channels in the adjacent two-layer after feature map channel pruning, we only select the features of the lower layer and discard the features of the upper layer of this channel. Figure 4 shows channel pruned Conv5-1 feature maps from the first frame of Basketball on VOT-2016 benchmark. In this way, our tracker extracts useful information for tracking task and significantly increases the tracking speed.

Note that ours channel pruning is different from channel reliability of [8]. The latter estimates the channel reliability whose scores are used for weighting the per-channel filter responses in localization in each frame. However, our AFER is only calculated by feature energies of target and search regions in the initial frame. It focuses more on the evaluation of features by the first frame of the input video in order to prune invalid channels.

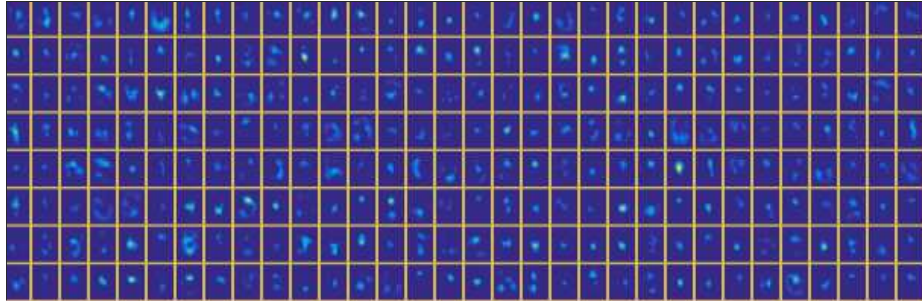


Fig. 4. Visualization of the selected features. We prune the ineffective or duplicated convolution channels, adaptively retaining the most advantageous channels for deep features.

3.3 Adaptive Iterative Optimization Strategy

The real-time performance of visual tracking mainly depends on the amount of calculations in feature extraction, filter training and position location. At the stage of the last one, C-COT and ECO tracker first perform a grid search, where the score function is evaluated at the discrete locations to obtain the initialization position p_0 . Then they employ the standard Newton iteration method to predict optimal object position $p_t(x, y)$ under fixed number of iterations. In theory, the more iteration numbers, the closer solution is to the true one, the greater the amounts of calculations are. However, when iteration reaches a certain number, minimal changes to the optimization have little or no effect on the final results. At this time, the extra number of iterations will not only significantly increase the amount of calculation but also cause a waste of resources.

In order to reduce the redundancy of iterations, an adaptive Newton iterative optimization strategy is designed to adaptively terminate the iterative

process. The main idea is to calculate the position difference between two consecutive iterations and find a suitable position error threshold τ . When it satisfies $\text{sum}(|p_t(x, y) - p_{t-1}(x, y)|) < \tau$, the iterative process stops, here t denotes the iteration numbers. This strategy can speed up our tracker with hardly no decline in precision.

4 Experiments

We evaluate our proposed tracking method on OTB-2013 [19], OTB-2015 [18], VOT-2016 [20] and VOT-2017 [22] benchmarks. The algorithm is implemented in Matlab R2015b, using MatConvNet toolbox, with an Intel Core i7-7800XCPU, 16GB RAM, and a GTX1080Ti GPU card.

We select Conv5-1 and Conv5-3 layers of channel pruned VGG16 as our feature extraction layers. The AFER thresholds are 1.1 and 1.5. The position error threshold is set to $\tau = 10^{-6}$. In addition, in order to improve the performance of the filter, the bandwidth of Gaussian labeled function for training sample is set to 0.15, the learning rate is set to 0.0115, the search region is set to 3.5 times of the target size. The model updating gap is 3 frames. Other parameters are the same as ECO tracker [6]. Code is available at <https://github.com/chemanqiang/CPT>.

4.1 Evaluation on VOT-2016

Table 1. Experimental results on VOT-2016. Our tracker achieves a substantial improvement over the baseline ECO method. The first and second best trackers are highlighted in red and blue, respectively.

Tracker	CPT	CPT_fast	CFWCR	ECO	C-COT	TCNN
EAO	0.41	0.394	0.391	0.374	0.331	0.327
Acc. Raw	0.56	0.55	0.58	0.55	0.54	0.53
Fail. Raw	0.68	0.77	0.81	0.87	0.89	0.90

We evaluate our tracker on VOT-2016 challenge that contains 60 annotated videos with substantial variations and measure the performance using Expected Average Overlap (EAO). Then compared the proposed tracking algorithm with four state-of-the-art methods, namely ECO [6], C-COT [3], CFWCR [11] and TCNN [10]. For clarity, we display the results in Table 1. CPT_fast algorithm here is a variation of our proposed. The difference between CPT and CPT_fast is that the latter regards the location and scale as two problems. It applies the location filter firstly to predict the targets position and then trains another 1D filter for scale estimation with the scale pyramid [17] based on the predicted position. Our CPT tracker outperforms all the trackers in VOT2016 challenge with an EAO score of 0.410, achieving a relative performance gain of 4.86%

compared with CFWCR. Moreover, our CPT tracker acquires an improvement over the baseline ECO with a relative gain of 9.63% in EAO. Note that our CPT_fast tracker with an EAO of 0.394, which is also competitive among the state-of-the-art trackers in the experiment.

4.2 Evaluation on VOT-2017

The VOT-2017 benchmark obtained 10 pairs of new sequences not present in other benchmarks and replaced 10 least challenging sequences in VOT-2016. Figure 5 illustrates the excellent performance of CPT tracker with four top ranked trackers including ECO [6], LSART [15], CFWCR [11] and CFCF [16]. In addition, we evaluate the compared trackers in terms of EAO, Accuracy Ranks (Ar) and Robustness Ranks (Rr). The detailed results are in Table 2. There is a large gap between other algorithms and ours, which illustrates our CPT tracker performs best against the evaluated trackers. Specially, the CPT and CPT_fast improve the ECO tracker by 24.2% and 6.05% in the metric of EAO, respectively.

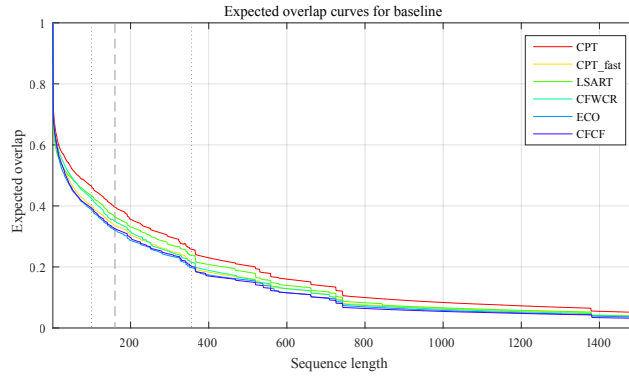


Fig. 5. Expected Overlap curves on VOT-2017 for baseline. When there are challenges of camera motion, occlusion, size change and etc., our CPT tracker has much better performance than the compared trackers.

Table 2. Experimental performances on VOT-2017. The first and second best trackers are highlighted in red and blue, respectively.

Tracker	CPT	CPT_fast	LSART	CFWCR	CFCF	ECO
EAO	0.349	0.298	0.323	0.303	0.286	0.281
Acc. Raw	0.50	0.52	0.49	0.48	0.50	0.48
Fail. Raw	1.04	1.20	0.94	1.21	1.17	1.12

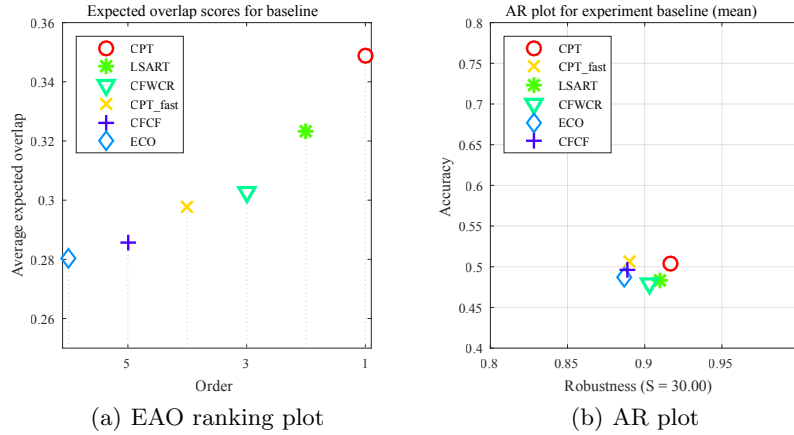


Fig. 6. A comparison for accuracy-robustness results on VOT2017 dataset. (a) The plot of EAO ranking for experiment baseline. (b) The plot of accuracy and robustness scores (AR). Apparently, our CPT gains a superior result.

Moreover, to better demonstrate the superiority of our tracker, we show the compared EAO ranking plot and accuracy-robustness results in Figure 6. Note that, the better trackers are located at the upper-right corner according to the protocol. Clearly, the proposed tracker obtains the rightmost position in the plot. Overall, our CPT achieves the appealing performance results both in accuracy and robustness on VOT-2017 dataset.

4.3 Evaluation on OTB

For completeness, we also display the evaluation results on OTB-2013 and its extensive dataset OTB-2015, which contain 11 various challenging factors such as deformation, occlusion, scale variation and etc. We employ the one-pass evaluation (OPE) with precision and success plots metrics. The precision metric measures the frame locations rate within a certain threshold distance from ground truth locations while the success plot metric measures the overlap rate between the predicted bounding boxes and the ground truth. Then compare our algorithm with another seven state-of-the-art trackers including ECO [6], VITAL [12], HCFTS [13], C-COT [3], LMCf [5], CSR-DCF [8] and Staple [9]. Figure 7 illustrates the precision and success plots based on center location error and bounding box overlap ratio, respectively. It clearly demonstrates that our CPT and CPT_fast gain the first and second top in precision on OTB-2013 and OTB-2015, outperforming the state-of-the-art trackers significantly.

We evaluate the speed and effectiveness for channel pruned model, channel pruned feature map component of our approach on OTB-2015 benchmark. The notation CPT_VGG16 donates the method using the original VGG16 model.

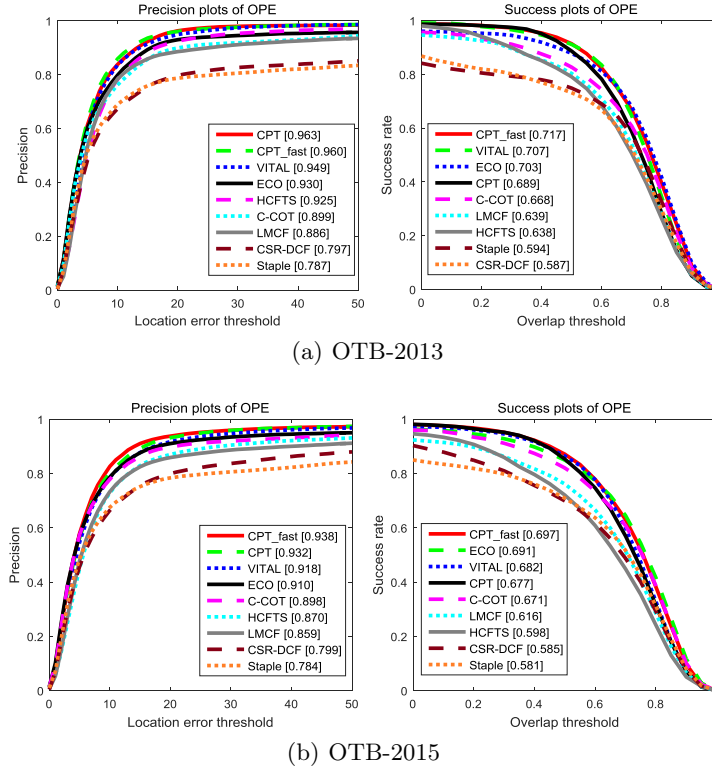


Fig. 7. Precision and success plots on OTB-2013 and OTB-2015. The numbers in the legend indicate the representative precisions at 20 pixels for precision plot and the area-under-curve scores for success plot. The proposed CPT and CPT_fast gain satisfactory results.

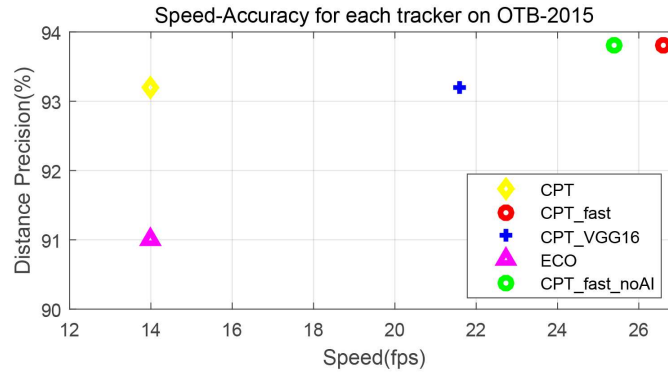


Fig. 8. Comparison on speed and accuracy of our three trackers and ECO. All our three trackers outperform the ECO tracker in speed and accuracy. Our CPT_fast tracker is the best.

Another strategy is the same as earlier introduced CPT_fast tracker using channel pruned VGG16 model. The results are shown in Figure 8. From this figure, the CPT_fast tracker wins the highest precision and fastest speed, showing the advantages of channel pruning and accurate scale estimation using hand-craft features. CPT_VGG16 runs slower than CPT_fast, which illustrates the channel pruned VGG16 model effectively improves the computational speed. We also investigate our visual tracking version without adaptive iterative optimization strategy(CPT_fast_noAI) based on CPT_fast. Consequently, we can conclude that the process of adaptive iteration improves the running speed effectively with hardly no change of precision. All our four trackers outperform the ECO tracker. Our trackers gain the best results and show the favorable performances in precision and real-time application.

4.4 Comparison of CPT and CPT_fast

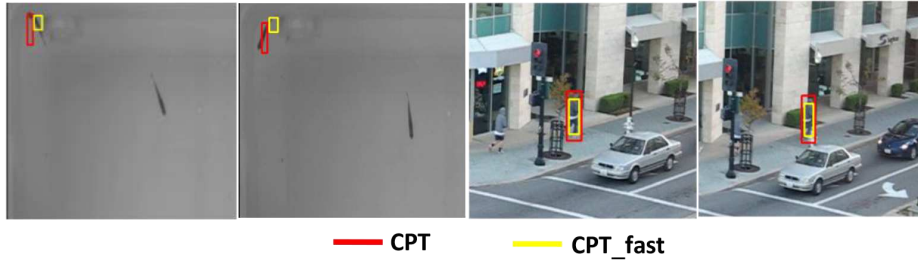


Fig. 9. Differences of CPT and CPT_fast(zebrafish1 and human4)

CPT tracker only predict 7 different scales for fast scale variations by CNN features, while CPT_fast has 33 predicted scales calculated by HOG. Therefore, when locating the target center is accurate to some extent, CPT_fast tracker can predict more accurate and faster than CPT, especially in complex scenes with multi-scale deformations. We show several different cases between CPT and CPT_fast in Figure 9. When subject to dramatic scale variations, CPT_fast tracker with more predicted scales can quickly find the appropriate scale to mark object in a larger scale range. Besides, violent non-rigid deformation leads to serious changes in target's appearance. However, CPT_fast constantly learns new features and loses more original features, which can make CPT_fast fail easily. CPT tracker does not have notable scale-predicted interference so that it can track object more robust. As a result, CPT has better stability than CPT_fast, while CPT_fast has better scale adaptability than CPT. That is also the reason why the OTB and VOT datasets exhibit different tracking performances.

5 Conclusions

We present a novel and robust channel pruning tracker (CPT) in this paper. Firstly, channel pruned VGG model is applied to fast extract deeper convolutional features with rich semantic information. Then we utilize the Average Feature Energy Ratio to further prune the redundant convolution channels, which are from the feature extraction layers and adaptive iterative strategy to optimize target location. Finally, we evaluate our CPT method on the OTB-2013, OTB-2015, VOT-2016 and VOT-2017 datasets. Extensive experiments demonstrate that the proposed CPT tracker outperforms the state-of-the-art trackers over all four benchmarks. The tracking of speed of CPT_fast tracker achieves 26 fps on OTB-2015. Our trackers gain the best results and show the favorable performances in precision and real-time application.

References

1. Henriques J. F., Caseiro R., Martins P., Batista J.: High-Speed Tracking with Kernelized Correlation Filters. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **37**(3), 583–596 (2015)
2. Danelljan M., Shahbaz Khan F., Felsberg M., 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)
3. Danelljan M., Robinson A., Shahbaz Khan F., Felsberg M.: Beyond Correlation Filters: Learning Continuous Convolution Operators for Visual Tracking. In: *Proceedings of European Conference on Computer Vision*, pp. 472–488 (2016)
4. Ma C., Huang J.B., Yang X., Yang M.H.: Hierarchical Convolutional Features for Visual Tracking. In: *Proceedings of the 2015 IEEE International Conference on Computer Vision*, pp. 3074–3082 (2015)
5. 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. 4800–4808 (2017)
6. Danelljan M., Bhat G., Shahbaz Khan F., Felsberg M.: ECO: Efficient Convolution Operators for Tracking. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 6931–6939 (2017)
7. Wang L., Ouyang W., Wang X., Lu H.: Visual Tracking with Fully Convolutional Networks. In: *Proceedings of the IEEE International Conference on Computer Vision*, pp. 3119–3127 (2015)
8. Lukezic A., Vojr 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. 4847–4856 (2017)
9. Bertinetto L., Valmadre J., Golodetz S., Miksik O., Torr P.H.S.: Staple: Complementary Learners for Real-Time Tracking. In: *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1401–1409 (2016)
10. Kang K., Li H., Yan J., Zeng X., Yang B., Xiao T., Zhang C., Wang Z., Wang R., Wang X., Ouyang W.: T-CNN: Tubelets with Convolutional Neural Networks for Object Detection from Videos. In: *IEEE Transactions on Circuits and Systems for Video Technology*, pp. 1–1 (2017)

11. He Z., Fan Y., Zhuang J., Dong Y., Bai H.: Correlation Filters with Weighted Convolution Responses. In: Proceedings of the IEEE International Conference on Computer Vision Workshops, pp. 1992–2000 (2017)
12. Song Y., Ma C., Wu X., Gong L., Bao L., Zuo W., Shen C., Lau R., Yang M.H.: VITAL: Visual Tracking via Adversarial Learning. In: Computer Vision and Pattern Recognition (2018 Spotlight)
13. Ma C., Huang J.B., Yang X., Yang M.H.: Robust Visual Tracking via Hierarchical Convolutional Features. In: arXiv prePreprint(2017)
14. Wang X., Li H., Li Y., Shen F., Porikli F.: Robust and Real-time Deep Tracking Via Multi-Scale Domain Adaptation. In: International Conference on Multimedia and Expo, pp. 1338–1343 (2017)
15. Sun C., Wang D., Lu H., Yang M.H.: Learning Spatial-Aware Regressions for Visual Tracking. In: Computer Vision and Pattern Recognition(2018 Spotlight)
16. Gundogdu E., Alatan A.A.: Learning Attentions: Good Features to Correlate for Visual Tracking. In: arXiv prePreprint(2017)
17. Danelljan M., Hager G., Shahbaz Khan F., Felsberg M.: Accurate Scale Estimation for Robust Visual Tracking. In: British Machine Vision Conference (BMVC). Durham: British Machine Vision Association, pp. 1–11 (2014)
18. Wu Y., Lim J., Yang M.: Object tracking benchmark. TPAMI **37**(9), 1834–1848 (2015)
19. Wu Y., Lim J., Yang M.: Online object tracking: A benchmark. In: Computer Vision and Pattern RecognitionR (2013)
20. Kristan M., Matas J., Leonardis A., et al.: The Visual Object Tracking VOT2016 challenge results. In: ECCV Workshops (2016)
21. Simonyan K., Zisserman A.: Very deep convolutional networks for large-scale image recognition. Computer Science **1**(2), 3–3 (2014)
22. Kristan M., Leonardis A., Matas J., et al.: The Visual Object Tracking VOT2017 challenge results. In: ICCV Workshops (2017)
23. He Y., Zhang X., Sun J.: Channel Pruning for Accelerating Very Deep Neural Networks. In: International Conference on Computer Vision (ICCV), pp. 1398–1406 (2017)
24. Bolme D.S., Beveridge J.R., Draper B.A., et al.: Visual object tracking using adaptive correlation filters. In: Proceedings of European Conference on Computer Vision, pp. 2544–2550 (2010)
25. Sun C., Wang D., Lu H., Yang M.: Correlation Tracking via Joint Discrimination and Reliability Learning. In: Proceedings of European Conference on Computer Vision (2018)
26. Henriques J.F., Rui C., Martins P., Batista J.: Exploiting the Circulant Structure of Tracking-by-detection with Kernels. In: Proceedings of European Conference on Computer Vision, pp. 702–715 (2012)
27. Hu H., Peng R., Tai Y.W., et al.: Network Trimming: A Data-Driven Neuron Pruning Approach towards Efficient Deep Architectures. In: arXiv preprint arXiv:1607.03250 (2016)
28. Wen W., Wu C., Wang Y., et al.: Learning Structured Sparsity in Deep Neural Networks. In: Proceedings of Advances In Neural Information Processing Systems, pp. 2074–2082 (2016)
29. Zhou H., Alvarez J.M., Porikli F.: Less is more: Towards compact cnns. In: Proceedings of European Conference on Computer Vision, pp. 662–667 (2016)
30. Danelljan M., Hger G., Khan F.S., Felsberg M.: Learning Spatially Regularized Correlation Filters for Visual Tracking. In: Proceedings of International Conference on Computer Vision, pp. 4310–4318 (2015)