Integrating Boundary and Center Correlation Filters for Visual Tracking with Aspect Ratio Variation (Supplementary Material)

Feng Li[1], Yingjie Yao[1], Peihua Li[2], David Zhang[3], Wangmeng Zuo [*,1], and Ming-Hsuan Yang[4]

[1]School of Computer Science and Technology, Harbin Institute of Technology, China
[2]School of Information and Communication Engineering, Dalian University of Technology, China
[3]Department of Computing, The Hong Kong Polytechnic University, China
[4]School of Engineering, University of California, Merced, USA
fengli_hit@hotmail.com, yyjhit@sina.com, peihuali@dlut.edu.cn
csdzhang@comp.polyu.edu.hk, wmzuo@hit.edu.cn, mhyang@ucmerced.edu

In the supplementary material, we first provide the results under all 11 video attributes on OTB-2015, then analyze the effects of each component on IBCCF and different layer weights in boundary CFs on the tracking performance.

1. Evaluation under different video attributes

To investigate the performance under different attributes, we compare IBCCF with 12 state-of-the-art trackers: HCF [9], C-COT [5], SRDCF [4], DeepSRDCF [3], RPT [7], SCF [8], SAMF [6], Scale-DLSSVM [10], Staple [1], DSST [2], MEEM [12] and SKSCF [13]. Fig. 1 shows the overlap success plots of all competing trackers. We can see that our IBCCF is among the top three trackers on all video attributes and achieves the second best performance on 7 of total 11 attributes. Note that SINT+ [11] is not used for comparison since we cannot find its code on OTB-2015.

2. Impacts of BCFs and near-orthogonality

Here, we investigate the impact of boundary CFs and near-orthogonality on the proposed IBCCF approach. To achieve this, we make four different variants of IBCCF: the tracker only with 1D boundary CFs (BCFs), the tracker only with the center CF (i.e., HCF), the IBCCF tracker without orthogonality constraints denoted by IBCCF (w/o constraint) and full IBCCF model. Table 1 summarizes the mean OP and AUC score of four methods on OTB-2015.

From Table 1, one can observe that BCFs underperforms the center CF. It is not difficult to explain the reason that the searching windows of BCFs contain less context information than the center CF for localizing the targets, so it drifts earlier as the target with large appearance variations. However, incorporating BCFs tracker into the center CF can bring significant improvements even without orthogonality constraints. It indicates BCFs tracker and center CF tracker are complementary. Moreover, orthogonality constraint can further improve 6.4% Mean OP and 4.1% AUC score over IBCCF (w/o constraint). These results validate the effectiveness of BCFs tracker and orthogonality constraint, and they are key parts of the proposed IBCCF tracker.

3. Effects of Feature Weights in Boundary CFs

In this paper, we introduce a family of boundary CFs, which are different from the center CF and aim at detecting the subtle changes in boundary positions. Therefore, we would like to study the impact of weights of different layers in boundary CFs on tracking performance. For center CF, the weights for layer *conv*3-4, *conv*4-4 and *conv*5-4 are set to 0.02, 0.5 and 1, respectively. We empirically find it is not optimal for the boundary CFs as a better performance can be obtained when all weights are set to 1.

To investigate this issue, we conduct experiments on OTB-2015 dataset, and and the results are illustrated in Fig. 2. In Fig. 2(a), we fix the weights of *conv*4-4 and *conv*5-4 as 1, and change the layer *conv*3-4 weights. The results show that it has little effect on the performance, and the best result is obtained at weight being 0. Hence, we discard the features from *conv*3-4, and only employ layer *conv*4-4 and *conv*5-4 in boundary CFs for improving efficiency. As shown in Fig. 2(b), the best performance is achieved when the weights of both layer *conv*4-4 and *conv*5-4 equal to 1. We use this weights setting throughout all the experiments.

*Corresponding Author.

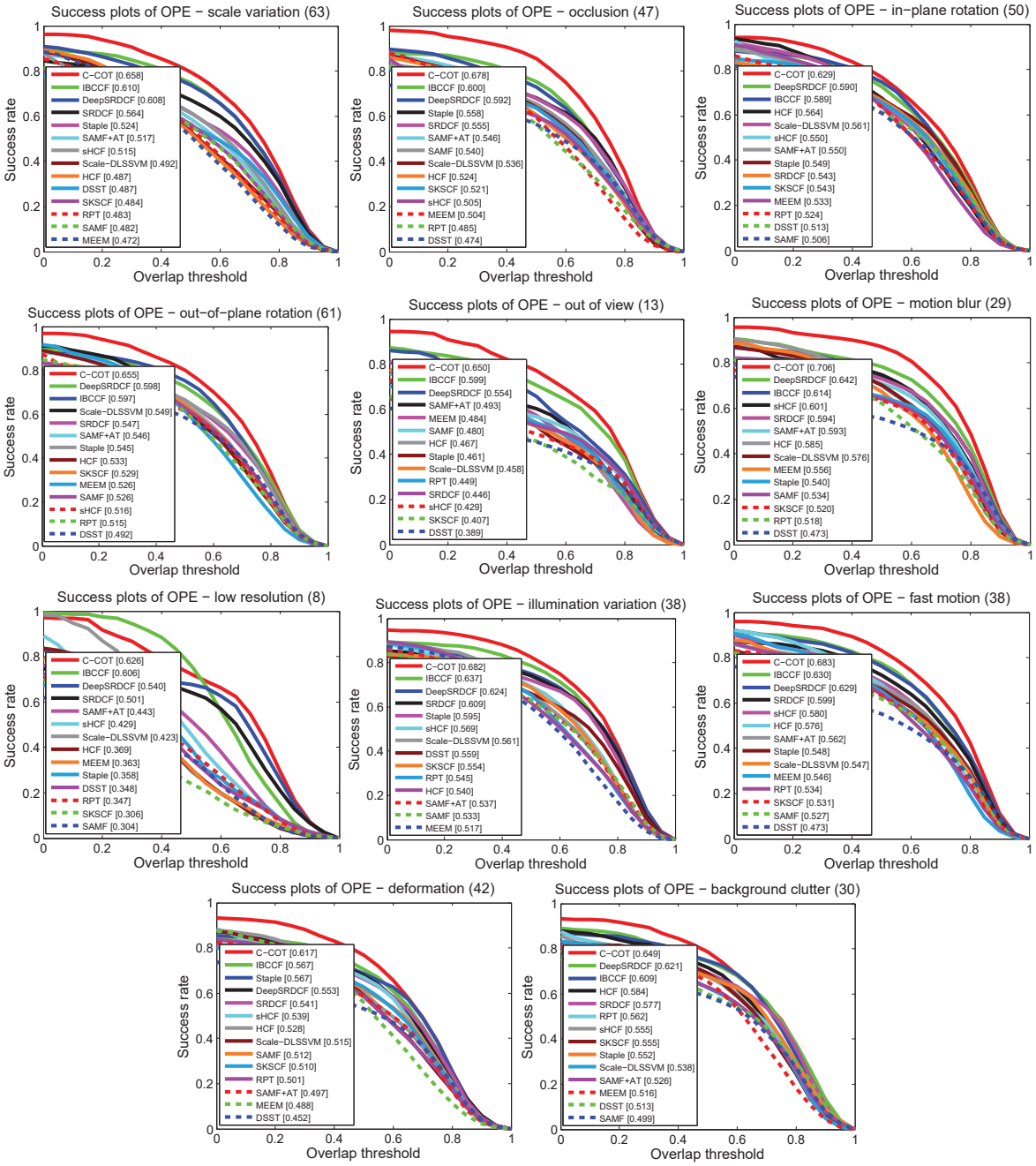
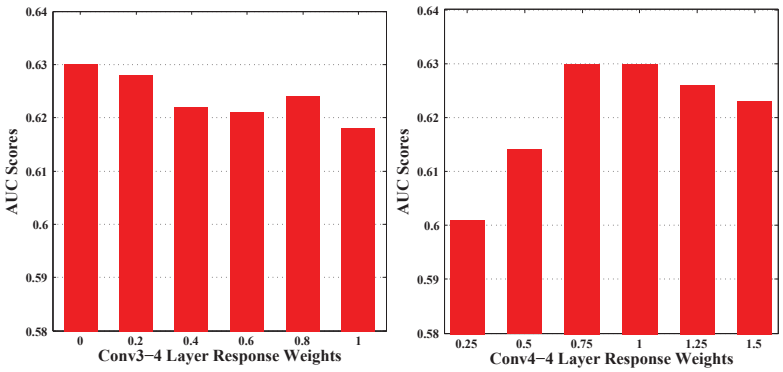


Figure 1: Overlap success plots of the the competing trackers under all eleven attributes on the OTB-2015 dataset.

	Mean OP	AUC Score
BCFs	50.2	39
Center CF	65.6	56.2
IBCCF (w/o constraints)	72	58.9
IBCCF	78.4	63

Table 1: Evaluation results of component experiments on both mean OP and AUC score metrics (in %).



(a) (b)

Figure 2: Effect of layer weights in the boundary CFs on OTB-2015 dataset. (a) Fix the weights of layer *conv*4-4 and *conv*5-4 equal to 1 and change layer *conv*3-4 response weights. (b) Fix the weights of layer *conv*3-4 and *conv*5-4 equal to 0 and 1, then change layer *conv*4-4 response weights.

References

[1] L. Bertinetto, J. Valmadre, S. Golodetz, O. Miksik, and P. Torr. Staple: Complementary learners for real-time tracking. In *CVPR*, 2016. 4321

[2] M. Danelljan, G. Hager, F. S. Khan, and M. Felsberg. Discriminative scale space tracking. *TPAMI*, PP(99):1–1, 2016. 4321

[3] M. Danelljan, G. Hager, F. Shahbaz Khan, and M. Felsberg. Convolutional features for correlation filter based visual tracking. In *ICCVW*, 2015. 4321

[4] M. Danelljan, G. Hager, F. Shahbaz Khan, and M. Felsberg. Learning spatially regularized correlation filters for visual tracking. In *ICCV*, 2015. 4321

[5] M. Danelljan, A. Robinson, F. Shahbaz Khan, and M. Felsberg. Beyond correlation filters: Learning continuous convolution operators for visual tracking. In *ECCV*, 2016. 4321

[6] Y. Li and J. Zhu. A scale adaptive kernel correlation filter tracker with feature integration. In *ECCVW*, 2014. 4321

[7] Y. Li, J. Zhu, and S. C. H. Hoi. Reliable patch trackers: Robust visual tracking by exploiting reliable patches. In *CVPR*, 2015. 4321

[8] S. Liu, T. Zhang, X. Cao, and C. Xu. Structural correlation filter for robust visual tracking. In *CVPR*, 2016. 4321

[9] C. Ma, J.-B. Huang, X. Yang, and M.-H. Yang. Hierarchical convolutional features for visual tracking. In *ICCV*, 2015. 4321

[10] J. Ning, J. Yang, S. Jiang, L. Zhang, and M. H. Yang. Object tracking via dual linear structured svm and explicit feature map. In *CVPR*, 2016. 4321

[11] R. Tao, E. Gavves, and A. W. Smeulders. Siamese instance search for tracking. In *CVPR*, 2016. 4321

[12] J. Zhang, S. Ma, and S. Sclaroff. Meem: robust tracking via multiple experts using entropy minimization. In *ECCV*, 2014. 4321

[13] W. Zuo, X. Wu, L. Lin, L. Zhang, and M.-H. Yang. Learning support correlation filters for visual tracking. *arXiv preprint arXiv:1601.06032*, 2016. 4321