

Learning Background-Aware Correlation Filters for Visual Tracking - Supplementary Material

Hamed Kiani Galoogahi¹, Ashton Fagg² and Simon Lucey¹

¹Robotics Institute, Carnegie Mellon University

²SAIVT Lab, Queensland University of Technology

hamedk@andrew.cmu.edu ashton@fagg.id.au slucey@cs.cmu.edu

1. Introduction

In this supplementary material we provide additional evaluations of the BACF tracker as following.

Online adaptation rate: We evaluated our tracker on OTB50 [2] using different adaptation rates (η) varying from 0.01 to 0.02 with an step of 0.001. We set the image region area of training samples to 5^2 times the object area. The result in Fig. 1 shows the best success rate is achieved by $\eta = 0.0125$. We fix the adaptation rate to 0.0125 over all evaluations in this paper.

Spatial size of training samples: We evaluated the performance of our tracker over a range of different spatial support sizes on the OTB50 dataset, as shown in Table 1. We set the spatial size of training samples to be N^2 times bigger than the target, where $N \in [2, \dots, 5]$. This experiment shows that increasing the support size improves the overlap precision, since more background patches are used for learning the tracker. However, since the tracking speed is linearly related to the support size, runtime performance suffers. We set the spatial support of training samples to be five times bigger than that of the target, $N = 5$ to trade-off between the accuracy and speed of our tracker.

Full attribute based evaluation: The complete comparison on all 11 attributes of OTB100 [3] is illustrated in Fig. 2, showing the superior performance of our method compared to all other HOG based trackers for all attributes.

Detailed comparison of CCOT [1] and BACF: CCOT and BACF are compared in Table 2 and Table 3, in terms of accuracy (success rate, IoU > 0.50), and time to track (in minutes). This result details the comparison provided by Fig. 6 in the main manuscript of the paper. This evaluation shows that in terms of number of videos one tracker outperforms another one, both trackers show very competitive accuracy. In terms of time to track, however, this comparison shows a big difference between BACF and CCOT. In average, BACF needs ~ 1.3 minute to track each video (varying respect to

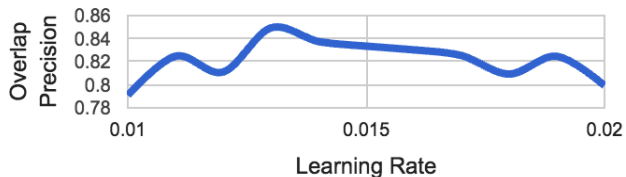


Figure 1. Learning rate selection. The best result is achieved by $\eta = 0.0125$, which is used for all experiments in this paper.

Table 1. Evaluating our method respect to different spatial size of training samples on OTB50 dataset. Results are reported as success rate at IoU > 0.50. We set $N = 5$ for all experiments.

	N	2	3	4	5
Succ. rate (%)		59.92	75.01	79.51	85.4
FPS		60.2	51.3	43.7	34.1

the videos' length), while for CCOT, the average time is significantly higher, around 230 minutes for each video. Please refer to the Tables and Fig. 6 in the main paper for more details.

Code and results: Some tracking videos, the MATLAB implementation of BACF and tracking results (predicted bounding box on all four datasets- mat files) can be found on www.hamedkiani.com.

References

- [1] M. Danelljan, A. Robinson, F. S. Khan, and M. Felsberg. Beyond correlation filters: Learning continuous convolution operators for visual tracking. In *ECCV*, pages 472–488, 2016. 1
- [2] Y. Wu, J. Lim, and M.-H. Yang. Online object tracking: A benchmark. In *CVPR*, pages 2411–2418, 2013. 1
- [3] Y. Wu, J. Lim, and M.-H. Yang. Object tracking benchmark. *PAMI*, 37(9):1834–1848, 2015. 1

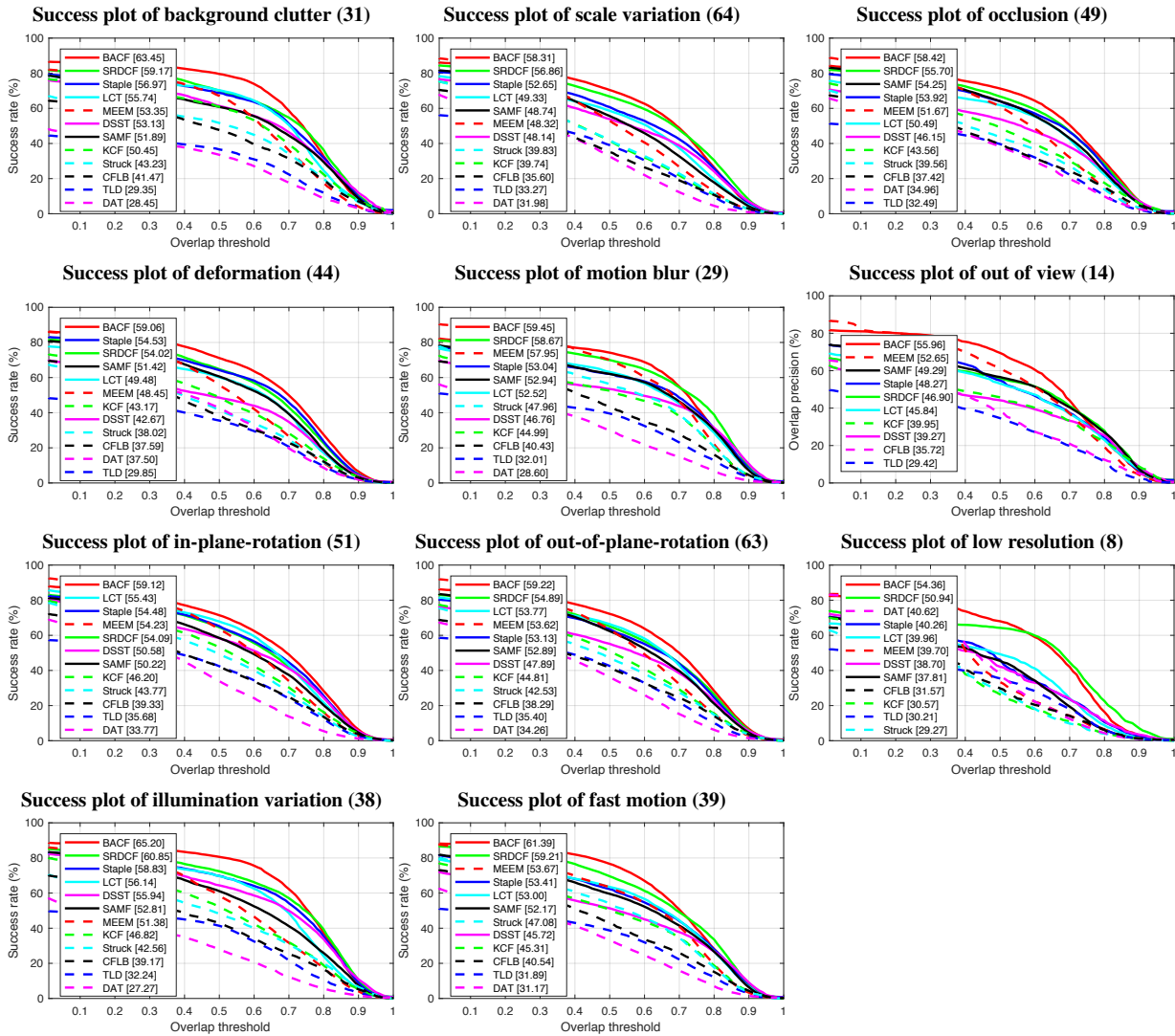


Figure 2. Attribute based evaluation. Success plots compare BACF with state-of-the-art HOG based trackers on OTB100. BACF outperformed all the trackers over all attributes. AUCs are reported in brackets. The number of videos for each attribute is shown in parenthesis.

Table 2. Detailed comparison of BACF and CCOT on OTB100 videos, accuracy (success rate at IoU > 0.50) and time to track (in minutes, amount of time each tracker needs to track each video). Red rows: BACF outperforms CCOT, Blue rows: CCOT and BACF perform equally, Green rows: CCOT outperforms BACF.

	Accuracy		Time to track	
	BACF	CCOT	BACF	CCOT
Basketball	81.4	98.9	1.6	319.9
Biker	48.6	53.3	0.1	25.7
Bird1	36.1	5.6	0.5	84.0
Bird2	47.5	100.0	0.2	35.4
BlurBody	100.0	98.8	1.7	175.8
BlurCar1	99.6	99.9	2.7	374.7
BlurCar2	100.0	100.0	2.5	215.0
BlurCar3	100.0	100.0	1.0	187.0
BlurCar4	100.0	100.0	2.3	180.7
BlurFace	100.0	100.0	1.7	233.4
BlurOwl	24.3	100.0	2.5	337.7
Board	92.0	87.2	3.8	346.1
Bolt	100.0	90.6	0.8	88.3
Bolt2	96.6	65.5	0.7	85.8
Box	41.5	89.5	3.7	589.8
Boy	99.8	100.0	1.2	136.0
Car1	100.0	100.0	1.5	424.0
Car2	100.0	100.0	1.5	441.7
Car24	100.0	100.0	3.7	733.9
Car4	100.0	100.0	1.1	320.9
CarDark	99.5	100.0	0.4	98.0
CarScale	100.0	80.2	0.6	54.3
ClifBar	67.4	71.2	0.7	131.8
Coke	80.1	54.0	0.7	153.1
Couple	87.2	72.1	0.2	22.9
Coupon	100.0	100.0	0.5	167.4
Crossing	100.0	100.0	0.1	18.2
Crowds	99.8	97.4	0.4	54.8
Dancer	99.6	100.0	0.4	88.8
Dancer2	100.0	100.0	0.2	49.6
David	96.4	100.0	0.9	242.4
David2	100.0	100.0	0.6	114.6
David3	100.0	99.6	0.7	99.5
Deer	92.9	100.0	0.2	18.8
Diving	18.6	18.6	0.4	67.8
Dog	44.1	83.5	0.3	35.0
Dog1	100.0	100.0	2.2	345.3
Doll	99.7	99.7	8.4	990.5
DragonBaby	51.4	96.5	0.3	37.1
Dudek	100.0	97.4	2.8	576.6
FaceOcc1	100.0	100.0	3.4	484.7
FaceOcc2	99.1	95.3	1.2	459.8
Fish	100.0	100.0	0.7	244.4
FleetFace	79.7	74.5	1.5	418.6
Football	98.1	79.0	0.5	127.4
Football1	97.3	37.8	0.1	10.3
Freeman1	94.8	43.6	0.2	82.5
Freeman3	55.0	100.0	0.3	119.8
Freeman4	42.4	74.6	0.2	60.4

Table 3. Detailed comparison of BACF and CCOT on OTB100 videos (cont.)

	Accuracy		Time to track	
	BACF	CCOT	BACF	CCOT
Girl	89.8	99.6	0.9	121.1
Girl2	7.1	97.5	4.5	732.3
Gym	34.5	6.1	1.5	353.8
Human2	96.7	94.2	5.6	590.4
Human3	2.9	87.8	3.1	615.4
Human4	97.9	75.3	1.2	234.1
Human5	96.9	99.6	0.9	181.7
Human6	97.8	96.2	1.6	213.8
Human7	100.0	100.0	0.5	114.1
Human8	100.0	100.0	0.2	35.6
Human9	99.7	100.0	0.5	134.1
Ironman	12.7	88.0	0.3	51.8
Jogging_1	97.1	97.4	0.7	106.9
Jogging_2	15.0	100.0	1.0	138.2
Jump	4.9	9.8	0.3	41.5
Jumping	97.5	99.0	0.3	85.7
KiteSurf	53.6	100.0	0.1	12.0
Lemming	94.9	95.8	4.1	654.0
Man	100.0	100.0	0.1	21.7
Matrix	40.0	62.0	0.2	16.4
Mhyang	91.6	100.0	2.7	701.7
MotorRolling	9.2	7.3	0.6	57.1
MountainBike	100.0	97.8	0.5	93.4
Panda	41.8	31.4	0.9	261.3
RedTeam	98.5	93.1	1.8	497.3
Rubik	11.8	99.4	9.0	924.6
Shaking	96.2	4.9	0.9	208.9
Singer1	100.0	100.0	0.9	131.1
Singer2	100.0	3.5	0.9	145.0
Skater	35.6	75.0	0.2	58.5
Skater2	77.7	76.8	0.7	236.1
Skating1	18.3	38.5	0.8	142.9
Skating2_1	31.5	59.4	1.5	232.2
Skating2_2	62.8	29.8	1.7	248.4
Skiing	6.2	30.9	0.1	13.0
Soccer	44.4	77.3	0.9	195.6
Subway	100.0	100.0	0.2	31.3
Surfer	94.4	96.0	0.4	88.8
Suv	98.6	97.0	1.4	425.7
Sylvester	84.2	95.0	1.9	531.7
Tiger1	97.5	99.7	0.9	192.2
Tiger2	92.3	89.3	0.9	222.2
Toy	89.0	84.1	0.4	105.8
Trans	34.7	48.4	0.5	44.2
Trellis	100.0	99.1	1.1	300.7
Twinnings	98.1	99.6	0.7	233.9
Vase	44.2	42.1	0.4	99.9
Walking	99.8	95.4	0.6	107.1
Walking2	100.0	100.0	0.8	300.4
Woman	99.4	98.0	1.1	193.7
Liquor	98.2	92.8	6.4	770.3