Parallel Tracking and Verifying: A Framework for Real-Time and High Accuracy Visual Tracking

Heng Fan Haibin Ling
Meitu HiScene Lab, HiScene Information Technologies, Shanghai, China
Department of Computer and Information Sciences, Temple University, Philadelphia, PA USA
{hengfan,hbling}@temple.edu

1. Siamese networks for verification

1.1. Architecture

Figure 1 shows the detailed architecture of siamese networks. Note that there are only two max pooling layers after conv1-2 and conv2-2 because too many max pooling layers may reduce the spatial information in features. Table 1 demonstrates the parameters of each layer. The source code of PTAV will be released for reproducible research.

1.2. Network training

We use the ALOV dataset [14] to train the siamese networks. Note that we exclude all videos that appear in OTB2013 [17], OTB2015 [18] and TC128 [9]. After removing these sequences, the training dataset and the evaluation datasets have no common objects. As in [15], we generate multiple pairs using every two frames in a video. One element in the pair is the groundtruth bounding box in one frame and the other one is a box sampled in the other frame. The pair is considered to be positive if the sampled box has a intersection-over-union overlap larger than 0.7 with the corresponding groundtruth box and considered to be negative if the overlap is smaller than 0.5. The training pairs and validation pairs are generated from different videos, and therefore from different objects. For training, in total we have sampled 60, 000 pairs of frames from ALOV dataset and each pair has 128 pairs of boxes. For validation, we have gathered 2, 000 pairs of frames and the same as for training each pair of frames contains 128 pairs of boxes.

We use the pre-trained network parameters from VGGNet [13] to initialize the networks. The initial learning rate is 0.001 and the weight decay parameter is 0.001. The learning rate is decreased by a factor of 10 after every 2 epochs. Training stops when the validation loss does not decrease any more.

2. Detailed results on OTB2015 [18]

In this supplementary material, detailed results on OTB2015 [18] with 100 videos are provided. Table 2 shows the per-video distance precisions for all trackers in comparison. Note that table 2 also contains the results on OTB2013 [17]. Figure 2 and Figure 3 show the distance precision plots and overlap successful plots for all 11 attributes, respectively.

3. Detailed results on TC128 [9]

We also present detailed results on the TC128 [9] with 128 videos. The per-video distance precision for all trackers in our comparison are reported in table 3. Figure 4 and Figure 5 show the distance precision plots and overlap successful plots for all 11 attributes, respectively.

4. Detailed results on UAV20L [12]

We present detailed results on the UAV20L [12] with 20 long videos. The shortest video contains 1717 frames, and the longest video contains 5527 frames. The per-video distance precision for all trackers in our comparison are reported in table 4. Figure 6 and Figure 7 show the distance precision plots and overlap successful plots for all 12 attributes, respectively.

References

[1] L. Bertinetto, J. Valmadre, S. Golodetz, O. Miksik, and P. H. Torr. Staple: Complementary learners for real-time tracking. In *CVPR*, 2016. 3, 6

[2] M. Danelljan, G. Hager, F. S. Khan, and M. Felsberg. Discriminative scale space tracking. *IEEE TPAMI*, 2016. 3, 6, 9

[3] M. Danelljan, G. Hager, F. Shahbaz Khan, and M. Felsberg. Learning spatially regularized correlation filters for visual tracking. In *ICCV*, 2015. 3, 6, 9

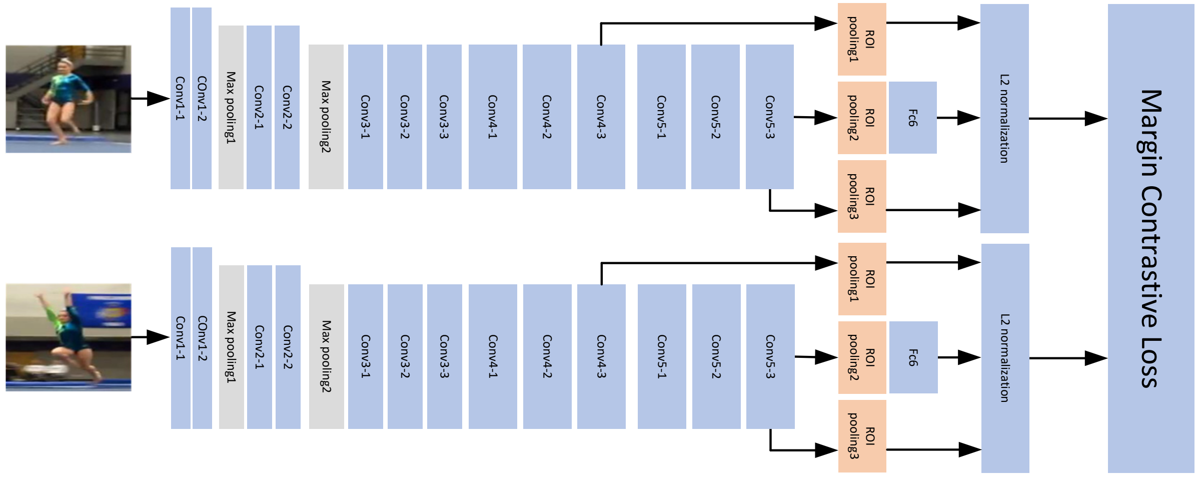

Figure 1. The detailed architecture of siamese networks.

Table 1. Parameters of the siamese networks.

Layer name	Conv1-1	Conv1-2	Max pooling1	Conv2-1	Conv2-2	Max pooling2
Parameters	3x3x64	3x3x64	2x2	3x3x128	3x3x128	2x2
Layer name	Conv3-1	Conv3-2	Conv3-3	Conv4-1	Conv4-2	Conv4-3
Parameters	3x3x256	3x3x256	3x3x256	3x3x512	3x3x512	3x3x512
Layer name	Conv5-1	Conv5-2	Conv5-3	ROI pooling1	ROI pooling2	ROI pooling3
Parameters	3x3x512	3x3x512	3x3x512	7x7	7x7	7x7

[4] J. Gao, H. Ling, W. Hu, and J. Xing. Transfer learning based visual tracking with gaussian processes regression. In *ECCV*, 2014. 3

[5] S. Hare, S. Golodetz, A. Saffari, V. Vineet, M.-M. Cheng, S. L. Hicks, and P. H. Torr. Struck: Structured output tracking with kernels. *IEEE TPAMI*, 38(10):2096–2109, 2016. 3, 6, 9

[6] J. F. Henriques, R. Caseiro, P. Martins, and J. Batista. High-speed tracking with kernelized correlation filters. *IEEE TPAMI*, 37(3):583–596, 2015. 3, 6, 9

[7] Z. Hong, Z. Chen, C. Wang, X. Mei, D. Prokhorov, and D. Tao. Multi-store tracker (muster): A cognitive psychology inspired approach to object tracking. In *CVPR*, 2015. 9

[8] Y. Li and J. Zhu. A scale adaptive kernel correlation filter tracker with feature integration. In *ECCV Workshop*, 2014. 9

[9] P. Liang, E. Blasch, and H. Ling. Encoding color information for visual tracking: Algorithms and benchmark. *IEEE TIP*, 24(12):5630–5644, 2015. 1, 6, 7, 8

[10] C. Ma, J.-B. Huang, X. Yang, and M.-H. Yang. Hierarchical convolutional features for visual tracking. In *ICCV*, 2015. 3, 9

[11] C. Ma, X. Yang, C. Zhang, and M.-H. Yang. Long-term correlation tracking. In *CVPR*, 2015. 6, 9

[12] M. Mueller, N. Smith, and B. Ghanem. A benchmark and simulator for uav tracking. In *ECCV*, 2016. 1, 9, 10, 11

[13] K. Simonyan and A. Zisserman. Very deep convolutional networks for large-scale image recognition. In *ICLR*, 2014. 1

[14] A. W. Smeulders, D. M. Chu, R. Cucchiara, S. Calderara, A. Dehghan, and M. Shah. Visual tracking: An experimental survey. *IEEE TPAMI*, 36(7):1442–1468, 2014. 1

[15] R. Tao, E. Gavves, and A. W. Smeulders. Siamese instance search for tracking. In *CVPR*, 2016. 1, 3

[16] N. Wang and D.-Y. Yeung. Learning a deep compact image representation for visual tracking. In *NIPS*, 2013. 3

[17] Y. Wu, J. Lim, and M.-H. Yang. Online object tracking: A benchmark. In *CVPR*, 2013. 1

[18] Y. Wu, J. Lim, and M.-H. Yang. Object tracking benchmark. *IEEE TPAMI*, 37(9):1834–1848, 2015. 1, 3, 4, 5

[19] J. Zhang, S. Ma, and S. Sclaroff. Meem: robust tracking via multiple experts using entropy minimization. In *ECCV*, 2014. 3, 6, 9

Table 2. A per-video comparison on the OTB2015 [18]. The best two results are highlighted with **red** and **blue** fonts (last row of table).

	PTAV	HCF [10]	SRDCF [3]	Staple [1]	MEEM [19]	SINT [15]	LCT [10]	fDSST [2]	KCF [6]	TGPR [4]	Struck [5]	DLT [16]
CarDark	1.000	1.000	1.000	1.000	1.000	0.964	1.000	1.000	1.000	1.000	1.000	0.715
Car4	1.000	0.997	1.000	1.000	0.686	1.000	0.989	1.000	0.953	1.000	0.992	1.000
David	0.994	1.000	1.000	1.000	0.904	1.000	1.000	1.000	1.000	0.987	0.329	0.321
David2	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.711
Sylvester	0.914	0.852	0.845	0.850	0.954	0.977	0.975	0.837	0.843	0.946	0.995	0.770
Trellis	1.000	1.000	1.000	0.996	0.968	1.000	1.000	1.000	1.000	0.981	0.877	0.339
Fish	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.401
Mhyang	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Soccer	0.941	0.816	0.934	0.296	0.314	0.531	0.151	0.946	0.791	0.143	0.253	0.138
Matrix	0.390	0.620	0.370	0.150	0.640	0.760	0.360	0.390	0.170	0.110	0.120	0.010
Ironman	0.596	0.645	0.030	0.145	0.506	0.614	0.145	0.078	0.217	0.096	0.114	0.127
Deer	1.000	1.000	1.000	1.000	1.000	1.000	0.817	1.000	0.817	1.000	1.000	0.042
Skating1	1.000	1.000	0.898	1.000	0.693	0.433	1.000	1.000	1.000	0.700	0.465	0.763
Shaking	0.978	0.868	0.014	0.019	0.995	0.981	0.984	0.953	0.019	0.641	0.192	0.926
Singer1	1.000	1.000	1.000	1.000	0.470	0.521	1.000	1.000	0.843	0.219	0.641	1.000
Singer2	0.926	0.041	0.973	1.000	0.038	0.732	0.973	0.904	0.948	0.954	0.036	0.036
Coke	0.966	0.962	0.818	0.897	0.945	0.969	0.914	0.890	0.838	0.942	0.948	0.340
Bolt	0.954	1.000	0.017	1.000	0.966	0.974	1.000	0.017	0.989	0.026	0.020	0.026
Boy	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Dudek	0.876	0.905	0.833	0.822	0.792	0.930	0.907	0.876	0.877	0.681	0.897	0.918
Crossing	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.950	1.000	1.000
Couple	0.914	0.921	1.000	0.679	1.000	0.871	0.571	0.600	0.257	0.107	0.736	0.307
Football1	1.000	1.000	0.784	1.000	1.000	1.000	0.973	1.000	0.959	0.986	1.000	0.608
Jogging-1	0.974	0.974	0.974	0.228	0.964	0.980	0.971	0.231	0.235	0.225	0.241	0.228
Jogging-2	0.935	1.000	0.997	0.192	0.971	0.977	0.974	0.163	0.163	0.997	0.254	0.173
Doll	0.968	0.978	0.993	0.993	0.985	0.959	0.981	0.994	0.968	0.971	0.919	0.957
Girl	0.918	1.000	0.994	0.868	1.000	1.000	1.000	0.916	0.864	0.904	1.000	0.776
Walking2	0.798	1.000	1.000	1.000	0.392	0.982	0.404	0.774	0.434	0.996	0.982	1.000
Walking	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.748
Fleetface	0.710	0.590	0.597	0.628	0.591	0.714	0.549	0.625	0.460	0.393	0.639	0.434
Freeman1	0.951	0.979	0.948	1.000	0.997	0.402	0.972	0.951	0.402	0.985	0.801	0.380
Freeman3	0.913	0.811	0.996	0.915	0.985	0.935	0.783	0.917	0.911	0.122	0.789	1.000
Freeman4	0.799	0.943	0.996	0.703	0.565	0.339	0.951	0.194	0.534	0.519	0.375	0.346
David3	1.000	1.000	1.000	1.000	0.996	1.000	1.000	1.000	1.000	1.000	0.337	0.698
Jumping	0.974	1.000	1.000	0.307	1.000	0.978	0.978	0.946	0.342	0.109	1.000	0.962
CarScale	0.813	0.627	0.778	0.853	0.651	0.627	0.730	0.813	0.806	0.806	0.647	0.714
Skiing	0.086	0.988	0.074	0.160	1.000	1.000	0.136	0.086	0.074	0.111	0.037	0.123
Dog1	1.000	1.000	1.000	1.000	0.982	0.850	1.000	1.000	1.000	1.000	0.996	0.996
Suv	0.978	0.979	0.975	0.978	0.743	0.957	0.980	0.978	0.979	0.531	0.572	1.000
MotorRolling	0.067	0.945	0.043	0.055	0.061	0.610	0.043	0.043	0.049	0.061	0.085	0.043
MountainBike	1.000	1.000	1.000	1.000	0.917	0.921	0.996	1.000	1.000	1.000	0.921	0.811
Lemming	0.939	0.258	0.323	0.273	0.911	0.875	0.856	0.275	0.487	0.275	0.628	0.298
Liquor	0.852	0.816	0.982	0.982	0.925	0.856	0.789	0.975	0.976	0.657	0.390	0.357
Woman	0.938	0.940	0.988	0.998	0.963	0.936	0.940	0.938	0.938	0.940	1.000	0.938
Faceocc1	0.882	0.600	0.831	0.918	0.683	0.858	0.906	0.882	0.730	0.831	0.575	0.462
Faceocc2	0.998	0.994	0.841	0.988	0.986	0.814	0.998	1.000	0.972	0.979	1.000	0.850
Basketball	0.930	1.000	0.996	0.879	0.892	0.866	1.000	0.913	0.923	0.993	0.120	0.086
Football	1.000	1.000	1.000	0.801	0.992	0.166	1.000	1.000	0.796	1.000	0.751	0.296
Subway	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.983	0.023
Tiger1	0.828	0.811	0.957	0.974	0.822	0.837	0.865	0.097	0.851	0.269	0.175	0.433
Tiger2	0.879	0.567	0.940	0.874	0.488	0.685	0.693	0.921	0.356	0.792	0.630	0.329
Biker	0.514	0.521	0.514	0.514	0.535	0.549	0.507	0.514	0.507	0.514	0.556	0.958
Bird1	0.373	0.392	0.071	0.363	0.289	0.453	0.346	0.373	0.069	0.811	0.150	0.588
Bird2	0.848	0.980	0.535	0.960	1.000	0.970	0.778	0.848	0.475	0.737	0.545	0.202
Blurbody	0.997	0.991	0.997	0.988	0.886	0.491	0.967	0.997	0.584	0.790	0.814	0.045
Blurcar1	0.997	0.995	0.999	0.695	0.993	0.410	0.997	0.997	0.995	0.035	0.996	0.026
Blurcar2	1.000	0.961	1.000	1.000	0.959	0.856	0.998	1.000	0.938	0.962	0.916	0.749
Blurcar3	1.000	1.000	1.000	0.986	1.000	0.683	1.000	1.000	0.994	0.042	1.000	0.252
Blurcar4	0.997	1.000	1.000	1.000	0.976	0.800	1.000	0.997	0.997	0.937	0.997	0.563
Blurface	1.000	1.000	1.000	0.998	0.990	0.937	1.000	1.000	1.000	0.990	0.436	0.191
Blurowl	0.930	0.962	0.984	0.472	0.995	0.678	0.891	0.937	0.228	0.512	0.989	0.070
Board	0.090	0.870	0.761	0.774	0.605	0.529	0.734	0.090	0.656	0.053	0.752	0.570
Bolt2	0.696	0.952	0.017	0.997	0.017	0.014	0.017	0.017	0.017	0.020	0.109	0.973
Box	0.924	0.394	0.415	0.414	0.370	0.960	0.068	0.396	0.415	0.294	0.239	0.396
Car1	1.000	0.391	1.000	1.000	0.196	0.350	0.438	1.000	0.739	0.337	0.334	1.000
Car2	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.074	1.000	1.000
Car24	1.000	1.000	1.000	1.000	1.000	0.942	0.853	1.000	1.000	0.992	0.170	1.000
ClifBar	1.000	0.915	0.945	0.697	0.915	0.572	0.939	1.000	0.445	0.146	0.581	0.464
Coupon	1.000	1.000	1.000	1.000	0.394	0.388	1.000	1.000	1.000	0.388	1.000	0.382
Crowds	1.000	1.000	1.000	1.000	1.000	0.991	1.000	1.000	1.000	1.000	0.911	0.916
Dancer	1.000	1.000	1.000	1.000	0.916	0.969	1.000	1.000	1.000	0.964	0.987	0.964
Dancer2	1.000	1.000	1.000	1.000	0.980	1.000	1.000	1.000	1.000	0.993	1.000	1.000
Diving	0.405	0.753	0.391	0.363	0.209	0.433	0.753	0.400	0.535	0.214	0.521	0.256
Dog	0.992	1.000	1.000	0.992	1.000	1.000	0.756	0.992	0.992	0.992	0.945	0.961
DragonBaby	0.442	0.867	0.336	0.858	0.823	0.850	0.549	0.389	0.336	0.752	0.106	0.372
Girl2	0.955	0.076	0.075	0.087	0.801	0.746	0.076	0.080	0.071	0.577	0.272	0.074
Gym	0.934	0.988	0.983	0.977	0.913	0.952	0.986	0.553	0.795	0.858	0.597	0.146
Human2	0.441	0.540	0.848	0.894	0.180	0.797	0.595	0.160	0.171	0.738	0.432	0.556
Human3	0.975	0.035	0.034	0.034	0.866	0.068	0.006	0.006	0.006	0.010	0.010	0.009
Human4	0.828	0.852	1.000	0.958	0.504	0.597	0.852	0.204	0.534	0.508	0.211	0.205
Human5	0.997	0.245	0.997	1.000	0.997	0.443	0.245	0.997	0.265	0.993	0.990	0.891
Human6	0.855	0.381	0.924	0.980	0.663	0.328	0.274	0.855	0.290	0.295	0.255	0.446
Human7	0.964	1.000	1.000	1.000	1.000	0.472	0.308	0.472	0.472	0.856	1.000	0.436
Human8	1.000	1.000	1.000	1.000	1.000	0.883	1.000	1.000	1.000	0.188	0.195	0.219
Human9	1.000	1.000	0.862	0.803	0.498	0.685	0.774	0.521	0.725	0.639	0.282	0.321
Jump	0.082	0.082	0.025	0.074	0.066	0.451	0.041	0.074	0.057	0.066	0.082	0.066
KiteSurf	1.000	0.464	0.679	1.000	1.000	0.786	0.464	0.464	0.333	0.440	0.905	0.286
Man	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.970	1.000	1.000
Panda	0.990	0.973	0.310	0.587	1.000	0.986	0.500	0.363	0.364	1.000	1.000	0.996
RedTeam	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Rubik	0.489	0.897	0.397	1.000	0.536	0.821	0.982	0.457	0.969	0.156	0.307	0.359
Skater	0.619	0.994	0.863	0.888	0.925	0.944	1.000	0.388	0.938	0.963	0.994	0.981
Skater2	0.591	0.903	0.600	0.839	0.913	0.823	0.729	0.543	0.694	0.352	0.726	0.237
Skating2-1	0.205	0.725	0.630	0.068	0.273	0.600	0.093	0.057	0.383	0.268	0.190	0.040
Skating2-2	0.228	0.444	0.467	0.049	0.178	0.429	0.019	0.015	0.490	0.262	0.292	0.114
Surfer	1.000	1.000	0.997	0.375	0.987	0.830	0.979	1.000	0.910	0.992	0.971	0.585
Toy	0.915	0.830	0.963	0.952	0.745	0.878	0.893	0.959	0.985	0.690	0.897	0.214
Trans	0.315	0.298	0.274	0.476	0.210	0.339	0.210	0.218	0.306	0.250	0.226	0.185
Twinnings	1.000	0.989	0.513	0.998	0.987	0.506	0.850	0.765	0.907	0.994	1.000	0.998
Vase	0.694	0.624	0.808	0.871	0.476	0.742	0.476	0.756	0.793	0.779	0.513	0.413
Average	**0.849**	**0.837**	0.789	0.784	0.781	0.773	0.762	0.720	0.692	0.643	0.639	0.526

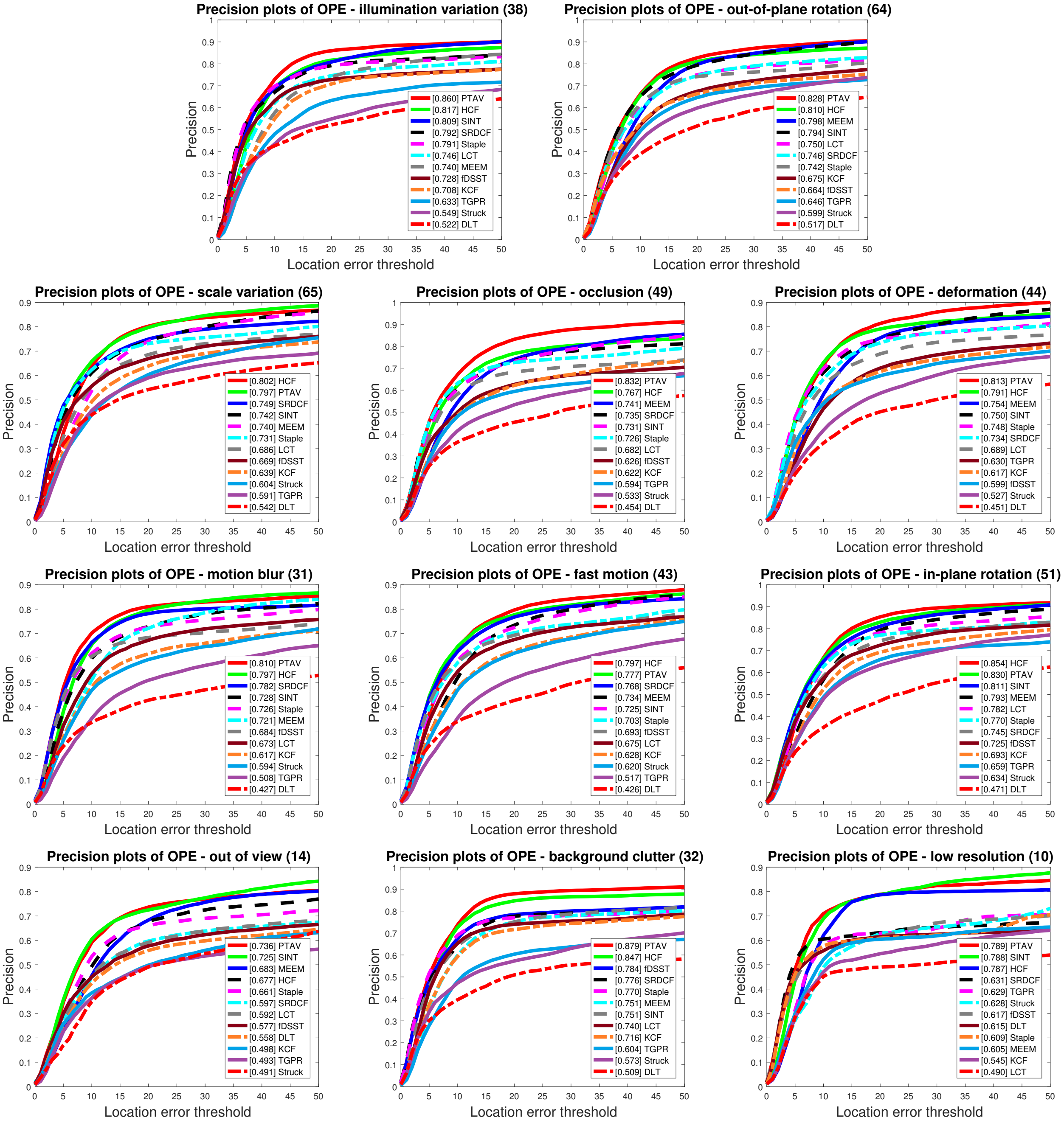

Figure 2. Distance precision plots on OTB2015 [18] for 11 attributes, which are background cluttered (BC), deformation (DEF), fast motion (FM), in-plane rotation (IPR), illumination variation (IV), low resolution (LR), motion blur (MB), occlusion (OCC), out-of-plane rotation (OPR), out-of-view (OV) and scale variation (SV). The distance precision (DP) score is shown in the legend. Our PTAV ranks top 2 on all 11 attributes.

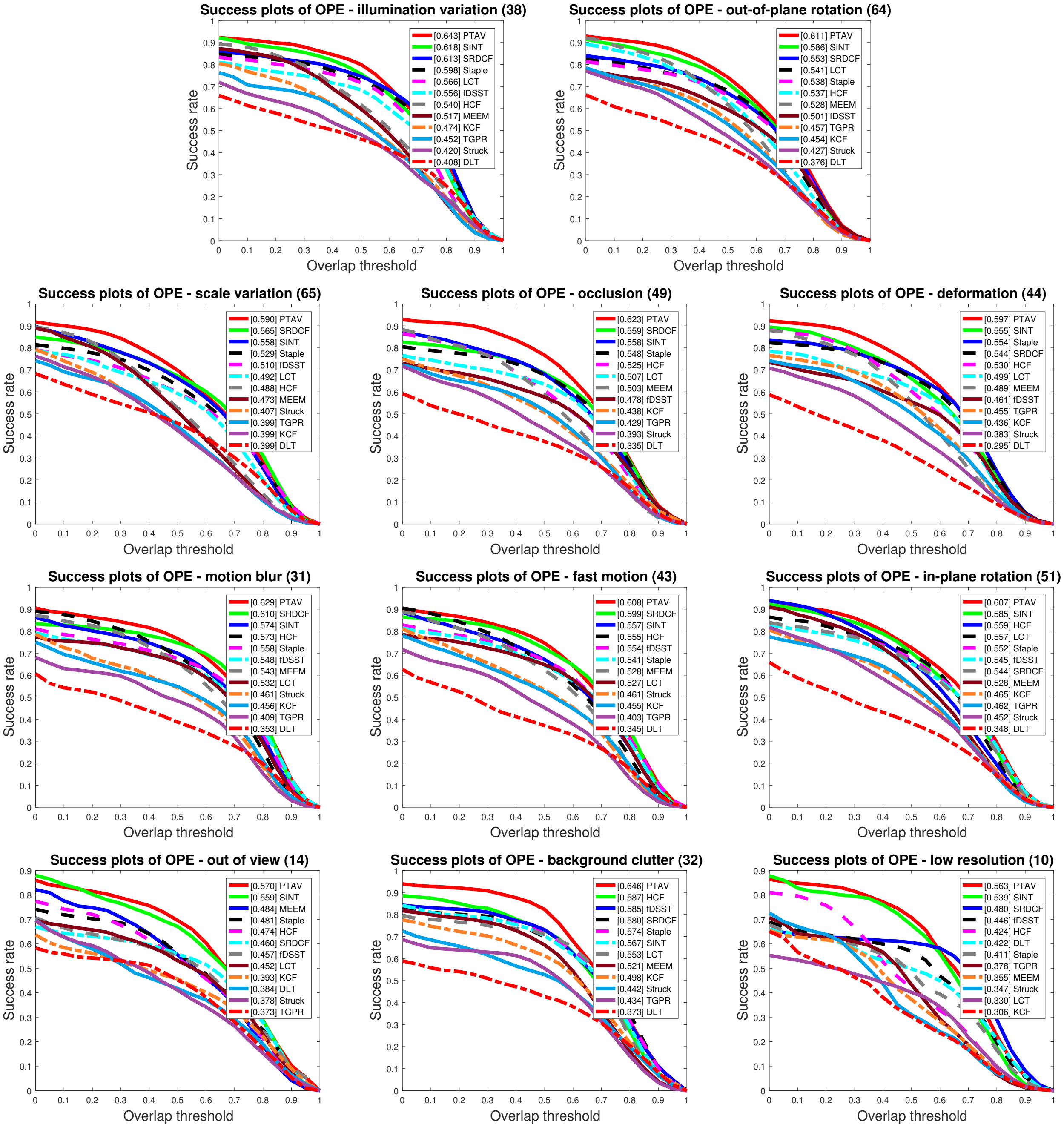

Figure 3. Overlap success plots on OTB2015 [18] for 11 attributes, which are background cluttered (BC), deformation (DEF), fast motion (FM), in-plane rotation (IPR), illumination variation (IV), low resolution (LR), motion blur (MB), occlusion (OCC), out-of-plane rotation (OPR), out-of-view (OV) and scale variation (SV). The title of each attribute plot contains the name of the attribute and the number of videos associated with it. The overlap success (OS) score is shown in the legend. Our PTAV ranks top 2 on all 11 attributes.

Table 3. A per-video comparison on the TC128 [9]. The best two results are highlighted with **red** and **blue** fonts (last row of table).

	PTAV	HCF [11]	Staple [1]	SRDCF [3]	MEEM [19]	Struck [5]	LCT [11]	fDSST [2]	KCF [6]
David	0.992	1.000	1.000	1.000	0.985	0.975	1.000	1.000	1.000
Airport_ce	0.757	0.392	0.493	0.446	0.378	0.405	0.412	0.459	0.392
Baby_ce	1.000	1.000	1.000	1.000	0.956	0.892	0.841	1.000	0.662
Badminton_ce1	0.997	0.997	0.915	0.988	0.986	0.998	0.993	0.352	1.000
Badminton_ce2	0.918	0.973	0.969	0.787	0.631	0.915	0.929	0.868	0.081
Ball_ce1	0.026	0.368	0.031	0.023	0.031	0.192	0.018	0.023	0.028
Ball_ce2	0.758	0.917	0.481	0.483	0.771	0.599	0.569	0.458	0.479
Ball_ce3	0.919	0.762	0.762	0.762	0.762	0.762	0.762	0.762	0.762
Ball_ce4	0.457	0.054	0.058	0.058	0.059	0.058	0.035	0.058	0.015
Basketball	0.930	1.000	0.879	0.996	0.877	0.259	1.000	0.913	0.921
Basketball_ce1	0.522	0.938	0.649	0.643	0.649	0.135	0.978	0.377	0.480
Basketball_ce2	0.549	0.358	0.587	0.798	0.112	0.481	0.582	0.145	0.244
Basketball_ce3	0.683	0.955	1.000	0.728	0.760	0.776	0.735	0.726	0.726
Bee_ce	0.922	1.000	1.000	0.422	1.000	1.000	0.389	0.400	0.289
Bicycle	1.000	1.000	1.000	1.000	0.646	1.000	0.524	1.000	0.524
Bike_ce1	1.000	1.000	1.000	1.000	0.999	0.984	1.000	1.000	1.000
Bike_ce2	0.180	0.988	0.138	1.000	0.248	0.164	0.472	0.167	0.548
Biker	0.650	0.594	0.956	0.461	0.589	0.683	0.456	0.467	0.467
Bikeshow_ce	0.213	0.753	0.050	0.152	0.102	0.579	0.055	0.288	0.258
Bird	0.838	0.980	0.909	0.515	0.980	0.485	0.768	0.515	0.566
Board	0.186	0.834	0.860	0.855	0.023	0.263	0.831	0.100	0.803
Boat_ce1	0.061	0.080	0.066	0.531	0.151	0.074	0.265	0.061	0.130
Boat_ce2	0.701	0.743	0.745	0.733	0.743	0.745	0.697	0.748	0.745
Bolt	0.954	1.000	1.000	0.017	0.983	0.054	1.000	0.017	0.997
Boy	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Busstation_ce1	0.972	0.107	0.107	0.105	0.124	0.116	0.113	0.127	0.113
Busstation_ce2	1.000	0.929	0.899	1.000	0.876	0.878	0.896	1.000	0.899
CarDark	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
CarScale	0.813	0.627	0.841	0.778	0.714	0.690	0.730	0.813	0.806
Carchasing_ce1	0.920	0.287	0.283	0.930	0.281	0.279	0.289	0.285	0.285
Carchasing_ce3	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Carchasing_ce4	1.000	0.973	1.000	1.000	0.541	0.394	0.229	1.000	0.837
Charger_ce	0.117	0.091	0.658	0.577	0.027	0.084	0.087	0.124	0.101
Coke	0.966	0.962	0.887	0.818	0.945	0.811	0.914	0.897	0.852
Couple	0.729	0.921	0.679	1.000	1.000	0.714	0.571	0.586	0.250
Crossing	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Cup	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Cup_ce	0.038	0.012	0.021	0.018	0.015	0.015	0.018	0.038	0.041
David3	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Deer	1.000	1.000	0.831	1.000	1.000	0.901	0.817	1.000	0.887
Diving	0.472	0.801	0.342	0.645	0.182	0.242	0.641	0.394	0.368
Doll	0.969	0.978	0.992	0.993	0.984	0.977	0.981	0.994	0.973
Eagle_ce	1.000	1.000	0.536	0.464	1.000	1.000	0.080	1.000	0.080
Electricalbike_ce	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
FaceOcc1	0.882	0.600	0.918	0.831	0.720	0.706	0.906	0.882	0.731
Face_ce	0.913	0.037	0.044	0.042	0.037	0.037	0.037	0.042	0.044
Face_ce2	0.338	0.682	0.716	0.588	0.520	0.473	0.088	0.101	0.101
Fish_ce1	0.858	0.698	0.728	0.090	0.085	0.072	0.085	0.077	0.082
Fish_ce2	0.545	0.194	0.147	0.159	0.215	0.387	0.148	0.148	0.161
Football1	1.000	0.973	1.000	0.784	1.000	1.000	0.973	1.000	0.986
Girl	0.882	1.000	0.864	0.994	1.000	1.000	1.000	0.894	0.866
Girlmov	0.962	0.075	0.075	0.075	0.983	0.227	0.075	0.075	0.075
Guitar_ce1	0.985	0.963	0.933	1.000	1.000	0.985	0.985	0.985	0.985
Guitar_ce2	0.543	0.840	0.514	0.508	0.454	0.719	0.524	0.543	0.575
Gym	0.918	0.974	0.956	0.954	0.936	0.931	0.948	0.555	0.966
Hand	0.443	0.197	0.176	0.148	0.172	0.156	0.996	0.197	0.189
Hand_ce1	0.938	0.825	0.898	0.958	0.479	0.065	0.918	0.940	0.895
Hand_ce2	0.853	0.964	1.000	0.896	1.000	0.988	0.996	0.853	0.637
Hurdle_ce1	0.957	0.993	0.707	0.973	0.730	0.993	0.700	0.967	0.983
Hurdle_ce2	0.977	1.000	0.984	0.980	0.947	0.928	0.967	0.977	0.970
Iceskater	0.622	0.762	0.822	0.800	0.684	0.744	0.590	0.616	0.714
Ironman	0.428	0.645	0.133	0.030	0.193	0.163	0.145	0.084	0.102
Jogging1	0.974	0.974	0.228	0.974	0.980	0.231	0.971	0.231	0.235
Juice	1.000	1.000	1.000	1.000	0.567	1.000	1.000	1.000	1.000
Kite_ce1	0.998	0.469	0.469	0.469	0.450	0.469	0.469	0.471	0.469
Kite_ce2	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.600
Kite_ce3	0.994	1.000	1.000	1.000	1.000	1.000	1.000	0.994	0.994
Kobe_ce	0.588	0.211	0.215	0.246	0.414	0.275	0.211	0.270	0.220
Lemming	0.939	0.258	0.273	0.323	0.908	0.776	0.856	0.275	0.275
Liquor	0.967	0.816	0.928	0.982	0.551	0.258	0.789	0.974	0.430
Logo_ce	0.916	0.290	1.000	1.000	0.464	0.989	0.393	0.993	1.000
Matrix	0.390	0.620	0.360	0.370	0.090	0.300	0.360	0.360	0.280
Messi_ce	0.717	0.636	0.768	0.989	1.000	0.967	0.533	0.533	0.283
Michaeljackson_ce	0.573	0.316	0.575	0.623	0.565	0.656	0.455	0.575	0.776
Microphone_ce1	0.863	0.995	0.618	1.000	0.990	1.000	1.000	0.853	1.000
Microphone_ce2	1.000	1.000	1.000	1.000	1.000	1.000	0.602	1.000	1.000
MotorRolling	0.104	0.945	0.055	0.043	0.043	0.159	0.043	0.043	0.049
Motorbike_ce	0.947	0.645	0.236	0.238	0.698	0.236	0.238	0.238	0.238
MountainBike	1.000	1.000	1.000	1.000	0.886	1.000	0.996	1.000	1.000
Panda	0.033	0.535	1.000	1.000	0.523	1.000	0.382	0.033	0.033
Plane_ce2	0.714	0.940	0.980	0.681	0.098	0.300	0.931	0.714	0.914
Plate_ce1	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Plate_ce2	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Pool_ce1	0.952	0.054	0.054	0.072	1.000	1.000	0.048	0.048	0.048
Pool_ce2	0.940	0.023	0.030	0.030	1.000	1.000	0.023	0.023	0.023
Pool_ce3	0.065	0.065	0.056	0.056	1.000	0.032	0.065	0.065	0.056
Railwaystation_ce	0.717	0.785	0.092	0.964	0.111	0.036	0.036	0.036	0.036
Ring_ce	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Sailor_ce	0.988	1.000	1.000	0.396	1.000	1.000	0.413	0.582	0.403
Shaking	0.964	0.868	0.019	0.014	0.975	0.770	0.984	0.953	0.022
Singer1	1.000	1.000	1.000	1.000	0.567	1.000	1.000	1.000	1.000
Singer2	0.929	0.041	1.000	0.973	0.036	0.036	0.973	0.904	0.973
Singer_ce1	0.897	0.963	0.949	0.953	0.813	0.972	1.000	0.897	0.939
Singer_ce2	0.257	0.057	0.790	0.053	0.802	0.007	0.393	0.133	0.275
Skating1	1.000	1.000	0.525	0.898	0.805	0.990	1.000	1.000	1.000
Skating2	0.446	0.871	0.823	0.774	0.226	0.298	0.798	0.038	0.588
Skating_ce1	0.394	0.337	0.682	0.054	0.570	0.323	0.401	0.059	0.296
Skating_ce2	0.087	0.284	0.433	0.052	0.151	0.052	0.087	0.099	0.072
Skiing	0.086	0.988	0.160	0.074	0.123	0.062	0.136	0.086	0.074
Skiing_ce	0.411	0.581	0.785	0.836	0.755	0.687	0.652	0.528	0.405
Skyjumping_ce	0.862	0.733	0.934	0.074	0.724	0.638	0.087	0.292	0.361
Soccer	0.941	0.816	0.306	0.934	0.288	0.288	0.151	0.946	0.161
Spiderman_ce	0.083	0.490	0.353	0.450	0.063	0.014	0.254	0.071	0.060
Subway	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Suitcase_ce	0.886	0.804	0.815	0.804	0.793	0.391	0.793	0.821	0.799
Sunshade	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
SuperMario_ce	0.884	0.815	1.000	1.000	1.000	0.993	0.322	0.404	0.233
Surf_ce1	0.153	0.438	0.205	0.062	0.178	0.193	0.248	0.040	0.035
Surf_ce2	0.136	0.020	0.049	0.049	0.430	0.069	0.041	0.028	0.013
Surf_ce3	0.158	0.538	0.616	0.237	0.444	0.093	0.398	0.140	0.136
Surf_ce4	0.081	0.044	0.333	0.081	0.193	0.074	0.111	0.170	0.081
TableTennis_ce	0.864	0.763	0.394	0.106	0.157	0.626	0.657	0.833	0.657
TennisBall_ce	0.396	0.038	0.024	0.031	0.024	0.132	0.021	0.021	0.017
Tennis_ce1	0.991	1.000	1.000	1.000	1.000	1.000	1.000	0.989	0.993
Tennis_ce2	0.993	0.990	1.000	1.000	1.000	1.000	1.000	0.993	1.000
Tennis_ce3	0.951	0.113	1.000	1.000	1.000	1.000	0.108	0.108	0.574
Thunder_ce	1.000	1.000	1.000	1.000	1.000	1.000	0.704	1.000	1.000
Tiger1	0.946	0.856	0.975	0.975	0.935	0.856	0.890	0.949	0.873
Tiger2	0.918	0.567	0.874	0.940	0.729	0.332	0.688	0.921	0.397
Torus	1.000	1.000	1.000	1.000	1.000	1.000	0.913	1.000	1.000
Toyplane_ce	0.501	0.279	0.079	0.669	0.084	0.104	0.084	0.086	0.084
Trellis	1.000	1.000	0.995	1.000	0.968	0.898	1.000	1.000	1.000
Walking	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Walking2	0.796	1.000	1.000	1.000	0.450	0.730	0.404	0.868	0.440
Woman	0.938	0.940	0.998	0.988	0.218	1.000	0.940	0.938	0.938
Yo-yos_ce1	0.298	0.102	0.115	0.140	0.217	0.111	0.153	0.094	0.136
Yo-yos_ce2	0.716	0.216	0.357	0.216	0.361	0.733	0.222	0.216	0.222
Yo-yos_ce3	0.025	0.030	0.030	0.060	0.060	0.786	0.030	0.025	0.025
Jogging2	0.935	1.000	0.192	0.997	0.971	0.173	0.974	0.163	0.163
Average	**0.741**	**0.705**	0.667	0.663	0.641	0.614	0.606	0.575	0.551

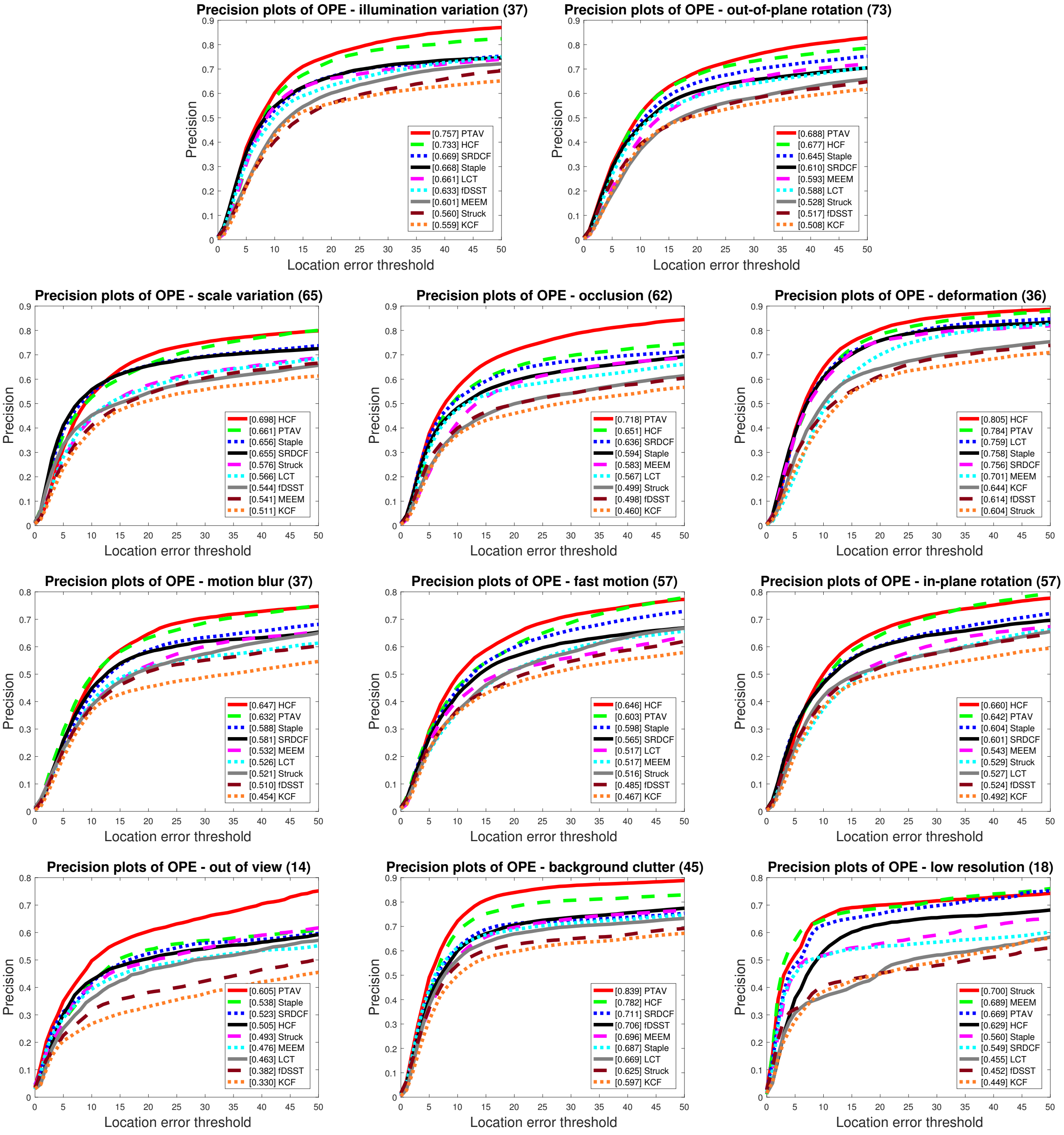

Figure 4. Distance precision plots on TC128 [9] for 11 attributes, which are background cluttered (BC), deformation (DEF), fast motion (FM), in-plane rotation (IPR), illumination variation (IV), low resolution (LR), motion blur (MB), occlusion (OCC), out-of-plane rotation (OPR), out-of-view (OV) and scale variation (SV). The distance precision (DP) score is shown in the legend. Our PTAV ranks top 2 on all 10 of 11 attributes.

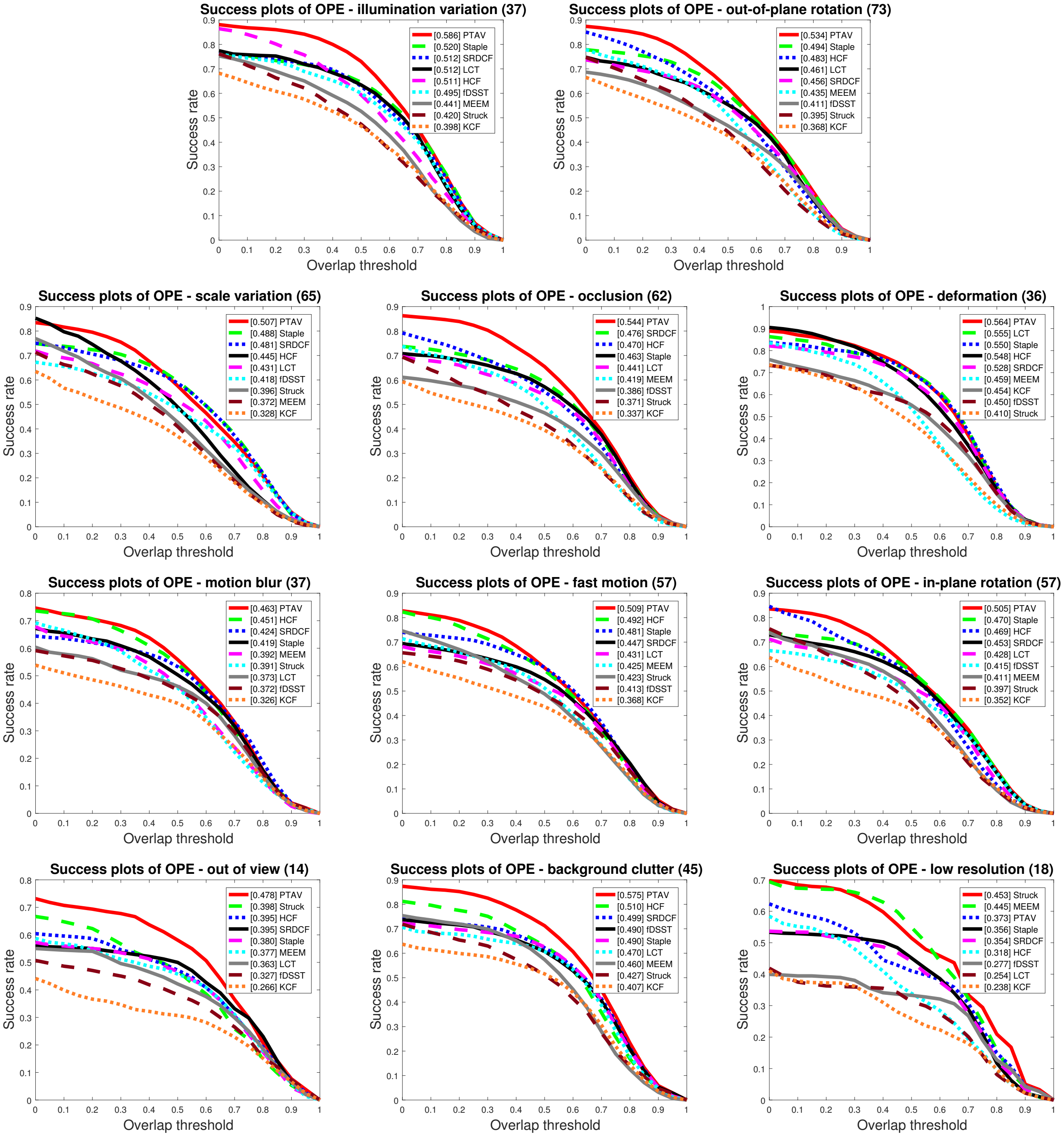

Figure 5. Overlap success plots on TC128 [9] for 11 attributes, which are background cluttered (BC), deformation (DEF), fast motion (FM), in-plane rotation (IPR), illumination variation (IV), low resolution (LR), motion blur (MB), occlusion (OCC), out-of-plane rotation (OPR), out-of-view (OV) and scale variation (SV). The title of each attribute plot contains the name of the attribute and the number of videos associated with it. The overlap success (OS) score is shown in the legend. Our PTAV ranks top 2 on all 10 of 11 attributes.

Table 4. A per-video comparison on the UAV20L [12]. The best two results are highlighted with **red** and **blue** fonts (last row of table).

	PTAV	MUSTer [7]	SRDCF [3]	HCF [10]	MEEM [19]	SAMF [8]	Struck [5]	fDSST [2]	LCT [11]	KCF [6]
Bike1	0.456	0.406	0.330	0.579	0.576	0.395	0.156	0.449	0.349	0.138
Bird1	0.439	0.436	0.436	0.435	0.441	0.436	0.461	0.439	0.436	0.435
Car1	0.830	0.657	0.657	0.618	0.385	0.615	0.569	0.657	0.618	0.382
Car3	1.000	0.981	1.000	1.000	1.000	1.000	0.912	1.000	1.000	0.837
Car6	0.369	0.350	0.117	0.118	0.098	0.175	0.158	0.148	0.124	0.125
Car8	0.434	0.419	0.750	0.087	0.407	0.250	0.388	0.434	0.106	0.078
Car9	0.988	0.979	0.424	0.109	0.396	0.423	0.115	0.423	0.151	0.225
Car16	0.712	0.118	0.356	0.666	0.349	0.125	0.408	0.440	0.093	0.112
Group1	0.356	0.548	0.895	0.939	0.930	0.798	0.672	0.356	0.490	0.212
Group2	0.899	0.350	0.114	0.114	0.114	0.114	0.143	0.114	0.114	0.114
Group3	0.399	0.130	0.362	0.363	0.366	0.367	0.365	0.352	0.349	0.365
Person2	0.881	1.000	1.000	1.000	1.000	1.000	1.000	0.881	1.000	0.996
Person4	0.298	0.266	0.937	0.877	0.457	0.294	0.450	0.299	0.298	0.299
Person5	0.742	1.000	0.980	1.000	1.000	1.000	0.709	0.683	0.660	0.508
Person7	0.885	0.321	0.330	0.352	0.357	0.196	0.505	0.203	0.186	0.143
Person14	0.976	0.801	0.063	0.064	0.063	0.064	0.064	0.064	0.064	0.064
Person17	0.994	0.406	0.605	0.999	0.993	0.996	0.894	0.994	0.605	0.605
Person19	0.138	0.233	0.297	0.303	0.221	0.301	0.201	0.138	0.173	0.156
Person20	0.236	0.474	0.376	0.228	0.173	0.480	0.217	0.236	0.445	0.317
Uav1	0.443	0.411	0.115	0.154	0.319	0.107	0.361	0.122	0.108	0.110
Average	**0.624**	**0.514**	0.507	0.500	0.482	0.457	0.437	0.422	0.368	0.311

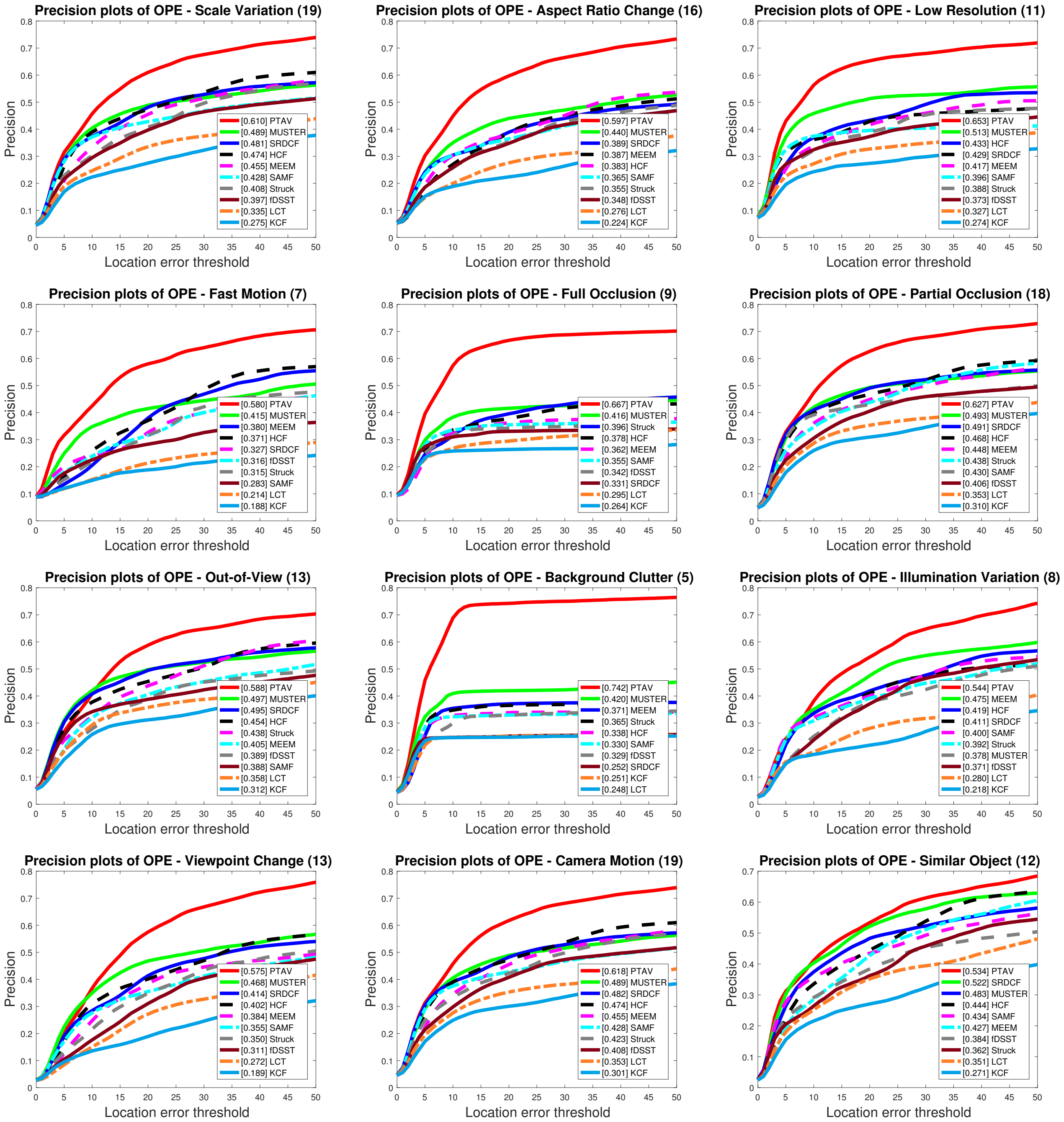

Figure 6. Distance precision plots on UAV20L [12] for 12 attributes, which are scale variation (SV), aspect ratio change (ARC), low resolution (LR), fast motion (FM), full occlusion (FOC), partial occlusion (POC), out-of-view (OV), background clutter (BC), illumination variation (IV), viewpoint change (VC), camera motion (CM) and similar object (SOB). The distance precision (DP) score is shown in the legend. Our PTAV ranks top 1 on all 12 attributes.

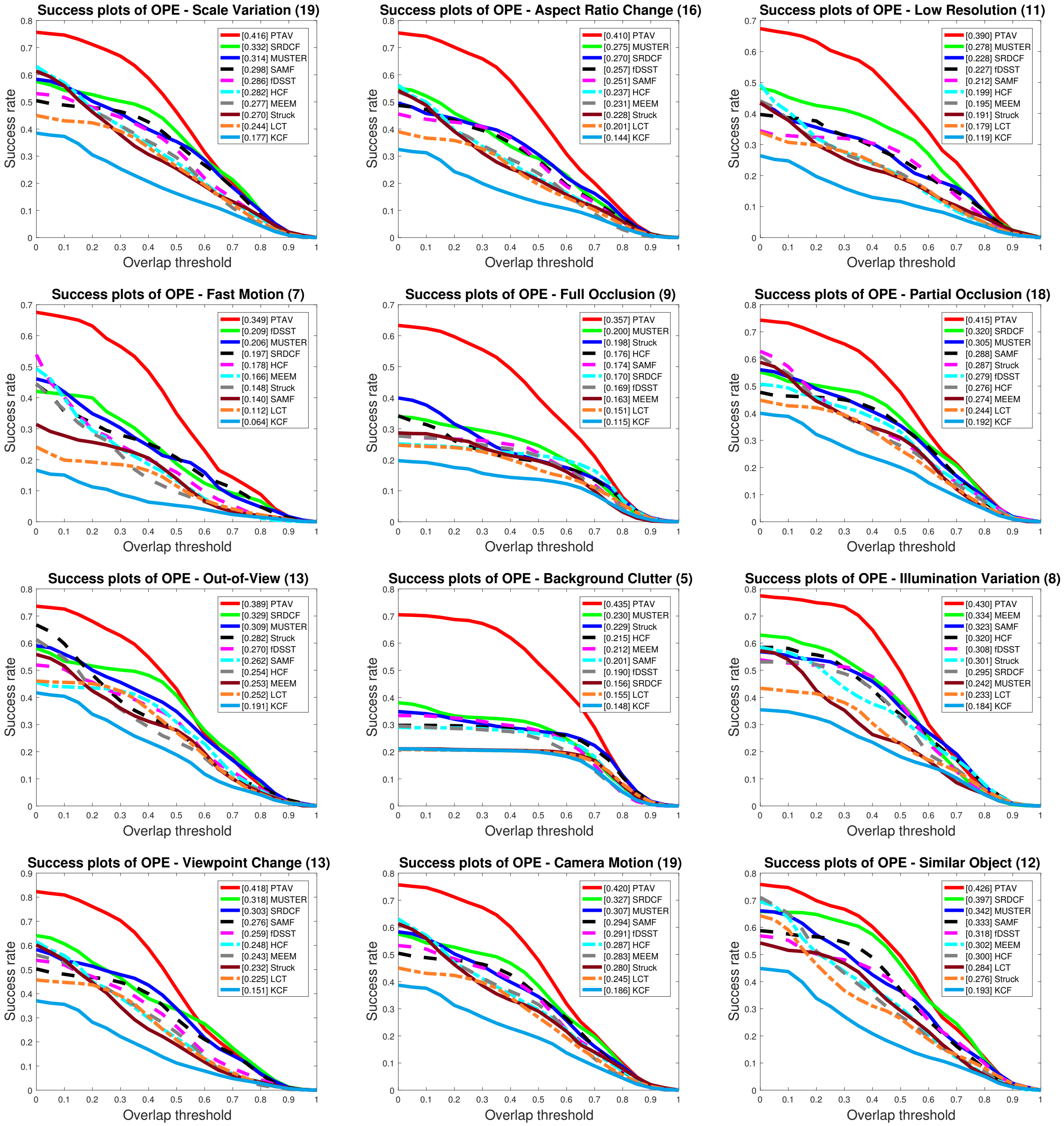

Figure 7. Overlap success plots on UAV20L [12] for 12 attributes, which are scale variation (SV), aspect ratio change (ARC), low resolution (LR), fast motion (FM), full occlusion (FOC), partial occlusion (POC), out-of-view (OV), background clutter (BC), illumination variation (IV), viewpoint change (VC), camera motion (CM) and similar object (SOB). The title of each attribute plot contains the name of the attribute and the number of videos associated with it. The overlap success (OS) score is shown in the legend. Our PTAV ranks top 1 on all 12 attributes.