Supplementary Material

Adaptive Color Attributes for Real-Time Visual Tracking

Martin Danelljan[1], Fahad Shahbaz Khan[1], Michael Felsberg[1], Joost van de Weijer[2]
[1]Computer Vision Laboratory, Linköping University, Sweden
[2]Computer Vision Center, CS Dept. Universitat Autonoma de Barcelona, Spain
{martin.danelljan, fahad.khan, michael.felsberg}@liu.se, joost@cvc.uab.es

In this supplementary material, we provide derivations and report additional results. In section 1, we show the derivation of the learning rule (equation 4 in the paper). Section 2 provides the detailed solution to the minimization problem (equation 6 in the paper), associated with the dimensionality reduction scheme. The additional results from our experiments are provided in section 3.

1 Derivation of the Learning Rule

This section proves that the cost function (3) in the paper (again stated here in (1)) is minimized by choosing the coefficients as equation (4) in the paper.

$$\epsilon = \sum_{j=1}^{p} \beta_j \left(\sum_{m,n} |\langle \phi(x^j_{m,n}), w^j \rangle - y^j(m,n)|^2 + \lambda \langle w^j, w^j \rangle \right) \tag{1a}$$

$$\text{where} \quad w^j = \sum_{k,l} a(k,l) \phi(x^j_{k,l}) \tag{1b}$$

Recall that x^j is the $M \times N$ training patch (of D-dimensional features) from frame number j, y^j is the associated label function, β_j are weights controlling the relative importance of different frames and λ is a regularization parameter. ϕ is the mapping to the Hilbert space induced by the kernel κ, which defines the inner product as $\langle \phi(f), \phi(g) \rangle = \kappa(f,g)$ for f and g in the same class as x^j. The cost function (1) is minimized over the coefficients a. $x_{m,n}$ denotes x cyclicly shifted m and n times in the first and second coordinate respectively. This means that $x_{m,n}(k,l) = x(k-m, l-n)$, if x is extended periodically. The summations in (1) are taken over all such shifts $(m,n) \in \{0,\dots,M-1\} \times \{0,\dots,N-1\}$.

Equation 1 is rewritten to (2), by inserting w^j in (1b) to (1a).

$$\epsilon = \sum_{j=1}^{p} \beta_j \left(\sum_{m,n} \left(\sum_{k,l} a(k,l) \kappa(x^j_{m,n}, x^j_{k,l}) - y^j(m,n) \right)^2 + \lambda \sum_{m,n} a(m,n) \sum_{k,l} a(k,l) \kappa(x^j_{m,n}, x^j_{k,l}) \right) \tag{2}$$

This function is convex in a, since the squared L^2-norm of an affine transformation is convex and ϵ is a sum of such functions. The global minimum can thus be obtained by finding a stationary point (where all partial derivatives are zero). The derivative with respect to $a(r,s)$ is computed in (3).

$$\begin{aligned} \frac{\partial \epsilon}{\partial (a(r,s))} &= 2 \sum_{j=1}^{p} \beta_j \sum_{m,n} \kappa(x^j_{m,n}, x^j_{r,s}) \left(\sum_{k,l} a(k,l) \kappa(x^j_{m,n}, x^j_{k,l}) - y^j(m,n) + \lambda a(m,n) \right) \\ &= 2 \sum_{j=1}^{p} \beta_j \sum_{m,n} \kappa(x^j_{r-m,s-n}, x^j) \left(\sum_{k,l} a(k,l) \kappa(x^j_{m-k,n-l}, x^j) - y^j(m,n) + \lambda a(m,n) \right) \end{aligned} \tag{3}$$

Here we have used the symmetric property and the assumed shift invariance of the kernel function. Shift invariance means that $\kappa(f_{m,n}, g_{m,n}) = \kappa(f,g)$ for all m, n, f and g. It is proved for inner product and radial basis function kernels in [8]. We define the kernel output function u^j_x in (4).

$$u^j_x(m,n) = \kappa(x^j_{m,n}, x^j), \quad m,n \in \mathbb{Z} \tag{4}$$

Note that u_x^j is periodic. Using this definition, the derivative in (3) can be expressed as (5). We use $*$ to denote circular convolution.

$$\begin{aligned}\frac{\partial \epsilon}{\partial(a(r,s))} &= 2\sum_{j=1}^{p}\beta_j \sum_{m,n} u_x^j(r-m,s-n)\left(\sum_{k,l} a(k,l)u_x^j(m-k,n-l) - y^j(m,n) + \lambda a(m,n)\right) \\ &= 2\sum_{j=1}^{p}\beta_j \sum_{m,n} u_x^j(r-m,s-n)\left(a * u_x^j(m,n) - y^j(m,n) + \lambda a(m,n)\right) \\ &= 2\sum_{j=1}^{p}\beta_j u_x^j * \left(a * u_x^j - y^j + \lambda a\right)(r,s) \end{aligned} \tag{5}$$

By setting these derivatives to zero and using the convolution property of the discrete Fourier transform (DFT) operator $\mathscr{F}$, the solution can be derived. We denote Fourier transformed functions with capital letters. Note that the products in this case are pointwise.

$$\begin{aligned} & \frac{\partial \epsilon}{\partial(a(r,s))} = 0\,, \quad \forall (r,s) \in \{0,\ldots,M-1\}\times\{0,\ldots,N-1\} \\ \iff & \sum_{j=1}^{p}\beta_j u_x^j * \left(a * u_x^j - y^j + \lambda a\right) = 0 \\ \iff & \mathscr{F}\left\{\sum_{j=1}^{p}\beta_j u_x^j * \left(a * u_x^j - y^j + \lambda a\right)\right\} = 0 \\ \iff & \sum_{j=1}^{p}\beta_j U_x^j \left(U_x^j A - Y^j + \lambda A\right) = 0 \\ \iff & A\sum_{j=1}^{p}\beta_j U_x^j \left(U_x^j + \lambda\right) - \sum_{j=1}^{p}\beta_j Y^j U_x^j = 0 \\ \iff & A = \frac{\sum_{j=1}^{p}\beta_j Y^j U_x^j}{\sum_{j=1}^{p}\beta_j U_x^j\left(U_x^j + \lambda\right)}. \end{aligned} \tag{6}$$

Here we have also exploited the fact that the DFT is linear and invertible. The last equivalence assumes that all frequency components in the denominator are non-zero, as in the original method [8].

2 Derivation of the Dimensionality Reduction Approach

To find a suitable $D_1 \times D_2$ projection matrix B_p to be used for dimensionality reduction in frame number p, we minimize the cost function

$$\eta_{\text{tot}}^p = \alpha_p \eta_{\text{data}}^p + \sum_{j=1}^{p-1}\alpha_j \eta_{\text{smooth}}^j. \tag{7}$$

under the constraint $B_p^{\mathrm{T}} B_p = I$. The constants $\alpha_1, \ldots, \alpha_p$ weight the relative importance of the terms. The data term is given by the reconstruction error of the current appearance $\hat{x}^p$, defined as:

$$\eta_{\text{data}}^p = \frac{1}{MN}\sum_{m,n}\left\|\hat{x}^p(m,n) - B_p B_p^{\mathrm{T}}\hat{x}^p(m,n)\right\|^2 = \operatorname{tr}\left(\left(I - B_p B_p^{\mathrm{T}}\right) C_p \left(I - B_p B_p^{\mathrm{T}}\right)\right). \tag{8}$$

Where $C_p = \frac{1}{MN}\sum_{m,n}\hat{x}^p(m,n)\hat{x}^p(m,n)^{\mathrm{T}}$ is the sample covariance matrix. The smoothness cost associated with frame number j is defined as:

$$\varepsilon_{\text{smooth}}^j = \sum_{k=1}^{D_2}\lambda_j^{(k)}\left\|b_j^{(k)} - B_p B_p^{\mathrm{T}} b_j^{(k)}\right\|^2 = \operatorname{tr}\left(\left(I - B_p B_p^{\mathrm{T}}\right) B_j \Lambda_j B_j^{\mathrm{T}}\left(I - B_p B_p^{\mathrm{T}}\right)\right). \tag{9}$$

Here $b_j^{(k)}$ is the k:th column in the earlier projection matrix B_j and $\lambda_j^{(k)}$ is a weight associated to this basis vector. Λ_j is defined as the $D_2 \times D_2$ diagonal matrix containing the weights $\lambda_j^{(k)}$.

Using (8) and (9), the cost function in (7) can easily be reformulated to the *equivalent maximization problem* in (10) by exploiting the properties of the trace operator.

$$V_{\text{tot}}^{p} = \text{tr}\left(B_p^{\text{T}} R_p B_p\right) \tag{10a}$$

$$\text{where} \quad R_p = \alpha_p C_p + \sum_{j=1}^{p-1} \alpha_j B_j \Lambda_j B_j^{\text{T}} \tag{10b}$$

The matrix R_p is symmetric and non-negative definite. Equation 10a can be maximized under $B_p^{\text{T}} B_p = I$ by choosing the columns of B_p as the D_2 eigenvectors of R_p that correspond to the largest eigenvalues.

3 Evaluation

In this section, we provide additional experimental results. Table 1 and 2 show the per-video results of the color feature evaluation (section 4.3 in the paper). In table 3 and 4, we report the per-video results of our robust learning scheme using different color features (section 4.4 in the paper). Table 5 and 6 shows the per-video results of our state-of-the-art comparison (section 4.6 in the paper). The best two results in all the tables are shown in red and blue fonts respectively. Finally, we provide all the precision and success plots showing the state-of-the-art comparison (section 4.6 in the paper). These include both the overall comparisons and the attribute-based evaluations.

3.1 Dataset

The 35 color sequences in the visual tracking benchmark by Wu et al. [15] used in this paper are found at:
https://sites.google.com/site/trackerbenchmark/benchmarks/v10
All six additional color sequences can be downloaded from our project web page:
https://www.cvl.isy.liu.se/en/research/objrec/colvistrack/index.html
Board, *Stone* and *Panda* was obtained from:
http://faculty.ucmerced.edu/mhyang/project/cvpr12_scm.htm
Kitesurf was obtained from:
http://www4.comp.polyu.edu.hk/~cslzhang/CT/CT.htm
Shirt was obtained from:
http://www.eng.tau.ac.il/~oron/LOT/LOT.html
Surfer was obtained from [2]:
http://vision.ucsd.edu/~bbabenko/project_miltrack.shtml

3.2 Results

Video	Int	Int+RGB	LAB	YCbCr	Int+rg	Opp	C	HSV	Int+SO	Int+AOpp	Int+HUE	Int+CN
basketball	**6.6**	*8.14*	16.4	46.2	40.5	9.08	39.7	204	45.5	145	125	89.4
board	109	111	*23.5*	28.3	66.1	111	137	108	75	**22.2**	4450	25.1
bolt	429	401	369	393	388	412	*4.48*	**3.8**	28.5	5.65	341	4.59
boy	20.1	56.2	48	65.8	60.3	26.5	*16.3*	31	66.2	85.8	68.2	**16**
carDark	2.71	2.69	**2.33**	4.45	3.15	2.7	2.71	*2.69*	23	90.3	114	2.75
carScale	83	25.4	25.8	46.4	88	25.5	25.7	*24.1*	**13.3**	48.6	25.8	68.3
coke	**13.6**	14.7	14.5	17	29.3	14.8	*14.3*	19.6	54.3	27.4	15.3	31.6
couple	144	120	153	150	151	132	114	*113*	**75.9**	156	150	120
crossing	*9.18*	49.4	24.1	121	89.7	60.9	64.4	58.6	88.9	83.7	212	**8.91**
david	17.7	18.2	7.45	13.4	13.3	13.2	10.6	8.69	14	*5.53*	1960	**4.35**
david3	56	57.1	**5.56**	56.1	*7.05*	7.7	8.58	54.7	64.1	57.7	324	8.52
deer	5.23	*4.91*	5.37	228	5.43	5.16	4.98	16.2	211	**4.21**	358	5.16
doll	44.7	19.7	*14*	1290	98.5	19.1	40.6	75.9	20.9	32.7	3860	**12.2**
faceocc1	*12.6*	12.8	14.1	12.8	13.7	12.9	13.6	128	**11.2**	54.8	41.2	14
football1	16.2	21.5	25.2	28.8	9.88	15.1	14.5	**7.77**	33	126	71.2	*9.57*
girl	19.4	21.1	9.64	*5.18*	**3.35**	12.7	18.2	12.5	132	112	1260	19.2
human	7.6	7.62	*6.97*	7.66	7.29	7.47	7.41	7.93	111	21	244	**6.96**
ironman	185	203	164	108	158	198	*53.7*	158	147	**42.8**	86	218
jogging-1	135	110	*4.64*	139	199	182	139	106	119	123	122	**4.09**
jogging-2	164	158	**3.39**	*5*	49.6	190	157	157	160	171	198	178
kitesurf	64	42.1	**4.58**	143	56.3	38	61.2	25.1	40.4	*23.8*	71.8	80.7
lemming	114	**62.8**	82	107	81.6	128	*81.1*	81.8	169	81.2	156	83.6
liquor	155	165	97.8	167	96.6	164	*72.5*	131	110	99.1	**60.6**	109
matrix	114	83.1	239	242	66.8	85.3	85.8	**45.8**	*52.7*	110	136	68.5
motorRolling	393	340	466	**201**	234	556	620	241	392	231	*227*	426
mountainBike	6.58	6.74	6.84	6.8	6.63	6.72	*6.39*	6.64	**5.64**	8.93	336	6.72
panda	94	67.3	72.1	68.7	**59**	67.9	80.8	110	*64.2*	97.9	80.7	88.2
shaking	16.7	16.6	15.8	23.8	17.5	16.7	**12.5**	*15*	19.8	29.2	282	17.3
shirt	16.8	16.6	*9.73*	18.6	20.3	15.8	24.6	17.1	15.5	17.2	13.3	**9.72**
singer1	15.4	15.4	13.4	12.2	12.9	15.7	15.3	13	27.5	*9.69*	13.2	**7.75**
singer2	185	171	183	171	14.7	176	*6.34*	**5.72**	173	24.4	15	168
skating1	**7.78**	*7.9*	8.41	8.23	8.26	8.06	13.5	7.92	21.9	9.5	40.3	8.66
skiing	248	274	303	**155**	274	267	271	208	228	*194*	260	282
soccer	70.9	200	54.1	47.5	*27.5*	79.9	146	32.5	123	107	177	**16.8**
stone	4.57	*2.3*	131	44.4	2.9	3.68	2.71	5.69	104	146	375	**2.22**
subway	164	*155*	168	**77.5**	162	169	169	183	161	170	158	157
surfer	**4.51**	4.56	20.2	26.1	1550	4.88	31.5	17	19.9	10.4	151	*4.53*
tiger1	69.5	46.9	18.5	52.6	36.4	18.1	41.6	*17.1*	86.3	130	74.2	**17.1**
tiger2	59.6	36.5	50.5	43.4	103	29.1	27.1	**18.3**	30.3	22.1	802	*18.9*
trellis	18.9	17.8	15.3	16	16.5	17.8	15	**8.04**	59.3	13.9	14.6	*9.01*
walking2	18.5	24.5	46	*17.9*	**16.5**	25.5	39.5	42.1	32	52.5	27	38.2
woman	208	1800	**9.42**	12.1	14.1	2380	12.9	*10.3*	129	162	561	14.3
Median	50.3	39.3	*19.4*	46.3	38.5	25.5	26.4	24.6	64.1	56.2	151	**16.9**

Table 1: The per-video average center location error (CLE) (in pixels) for different color features (as in Table 1 in the paper). The best results are obtained using color names with a median CLE of 16.9 pixels.

Video	Int	Int+RGB	LAB	YCbCr	Int+rg	Opp	C	HSV	Int+SO	Int+AOpp	Int+HUE	Int+CN
basketball	**100**	**100**	75.6	49.7	57.1	*96.8*	70.2	2.76	64.8	4.97	2.48	11.9
board	8.57	8.57	**64.3**	61.4	9.29	8.57	9.29	8.57	8.57	*63.6*	4.29	*63.6*
bolt	3.43	5.43	3.14	3.14	5.43	5.43	**100**	**100**	*55.7*	**100**	3.14	**100**
boy	*84.4*	66.9	67.3	66.8	66.4	63.5	**84.6**	46.2	27.7	40.4	33.1	**84.6**
carDark	**100**	**100**	**100**	*92.9*	**100**	**100**	**100**	**100**	69	25.4	43.3	**100**
carScale	65.1	71.4	71.4	63.5	65.1	71.4	71.8	*73*	**78.6**	40.5	71.4	65.1
coke	*87.3*	84.5	**88.7**	81.8	28.5	84.5	84.5	86.3	18.2	34	85.6	63.9
couple	8.57	8.57	8.57	8.57	8.57	8.57	*10.7*	8.57	**11.4**	7.86	8.57	*10.7*
crossing	**100**	40.8	64.2	15	22.5	32.5	22.5	23.3	11.7	9.17	0.833	*96.7*
david	49.9	47.6	**100**	80.5	82.8	82.4	*98.7*	**100**	83.2	**100**	33.1	**100**
david3	65.9	65.9	**92.5**	65.1	*91.3*	*91.3*	90.5	71.8	57.1	64.7	13.1	90.5
deer	**100**	**100**	**100**	14.1	**100**	**100**	**100**	*84.5*	5.63	**100**	2.82	**100**
doll	58.1	58.2	74.1	42.4	32.1	*81.4*	59.7	53	60.4	67.5	2.2	**83.3**
faceocc1	*92.5*	91.1	85.3	91.4	87.4	91	88.1	22.9	**98.3**	53.4	51.9	85.7
football1	75.7	52.7	51.4	51.4	85.1	75.7	77	**100**	44.6	12.2	13.5	*90.5*
girl	55.4	52.4	*93.2*	**100**	**100**	86	54.8	84.6	3.8	38.6	0.4	51.4
human	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**	39.1	*75*	21.4	**100**
ironman	13.3	13.3	10.8	3.61	21.1	14.5	32.5	30.7	8.43	**38.6**	*35.5*	13.3
jogging-1	22.8	23.1	*97.4*	22.8	23.1	23.1	23.1	22.8	23.1	22.8	13.7	**97.7**
jogging-2	18.6	18.6	**100**	*95.8*	18.6	18.6	18.6	18.9	18.9	17.6	16.3	19.5
kitesurf	46.4	27.4	**96.4**	4.76	46.4	31	46.4	*48.8*	14.3	42.9	25	2.38
lemming	**43.6**	38.7	34.1	17.2	40.5	*41.1*	37.2	30.2	17.1	34.6	14.8	27.6
liquor	22.3	22.3	28.4	23	35.2	22.3	*41.9*	27.9	19.5	28	**52.6**	28.1
matrix	1.00	1.00	1.00	9.00	**35.0**	13.0	2.00	**35.0**	*33.0*	9.00	8.00	16.0
motorRolling	4.88	6.1	6.1	3.05	3.05	5.49	6.1	**13.4**	4.88	4.88	3.05	*8.54*
mountainBike	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**	*99.6*	0.439	**100**
panda	**53.5**	50.6	**53.5**	51.5	*52.7*	*52.7*	*52.7*	37.3	46.1	6.22	29.9	*52.7*
shaking	59.5	60	62.5	30.1	56.7	59.5	**89.3**	*67.7*	46.8	10.4	1.1	53.2
shirt	80.8	81	*89.2*	69.8	69.5	80.8	63.7	79.6	81.3	82.3	86.6	**89.9**
singer1	59	58.4	70.7	95.4	91.5	55.8	57.5	72.1	39.6	*97.7*	77.8	**100**
singer2	3.55	3.55	3.55	3.55	88.3	3.55	*97.3*	**100**	3.55	71.6	88	3.55
skating1	98.8	**100**	**100**	**100**	96	**100**	88.8	**100**	85.3	*99*	65.5	**100**
skiing	9.88	**13.6**	**13.6**	9.88	**13.6**	**13.6**	**13.6**	8.64	**13.6**	**13.6**	*12.3*	**13.6**
soccer	13.5	13.5	18.9	*34.4*	18.6	14.8	15.6	31.9	7.65	18.6	14.5	**71.7**
stone	89.9	**100**	16.8	20.2	**100**	89.9	**100**	*98.3*	15.1	13.4	10.1	**100**
subway	**24.6**	**24.6**	7.43	7.43	21.1	**24.6**	10.9	13.1	7.43	10.3	3.43	*22.9*
surfer	**100**	*98.7*	64.5	64.5	5.26	**100**	57.9	59.2	63.2	88.2	7.89	**100**
tiger1	25.5	36.4	52.7	41.3	47.9	62.2	45	*69.6*	12.9	15.8	21.5	**76.2**
tiger2	11	38.1	23.8	20.5	11	39.5	47.4	**69.3**	31.8	*65.5*	12.9	64.7
trellis	81	82.6	88.2	85.2	84.7	82.6	86.1	**100**	30.2	88.4	86.1	*95.4*
walking2	46	44	43.2	47.6	*48.6*	44	43.2	40.4	43	40.4	**49.8**	41.6
woman	25	25	**94**	*93.8*	*93.8*	25	**94**	**94**	38.7	19.9	11.2	*93.8*
Median	54.5	49.1	*65.9*	48.6	50.6	57.6	58.8	63.4	31	38.6	14.1	**74**

Table 2: The per-video distance precision (DP) (%) for different color features (as in Table 1 in the paper). The best performance is achieved using color names with a median DP score of 74.0%.

Video	Int+RGB	LAB	YCbCr	Int+rg	Opp	C	HSV	Int+SO	Int+AOpp	Int+HUE	Int+CN
basketball	7.47	95.4	*6.83*	35.3	7.49	**6.35**	165	44.7	9.99	60.2	39.5
board	110	*23.3*	25.2	67.8	112	142	75.2	76	**22.4**	3620	24
bolt	409	397	382	393	408	*4.03*	**3.89**	30.1	7.54	390	4.43
boy	42.8	17	44.3	17	42.8	43.7	52.4	54.8	**3.8**	*3.96*	4.64
carDark	2.76	2.87	2.72	**2.68**	2.74	*2.71*	2.93	22.6	3.07	2.73	2.79
carScale	25.6	26	26	25.7	25.6	25.7	*24*	**13.4**	25.9	26.9	26
coke	15.1	*14.9*	16.4	**14.1**	15.2	17.5	30.9	67.2	28	17.4	19.2
couple	140	123	142	143	140	123	119	**76.6**	*118*	118	123
crossing	61.3	31.8	50.5	36.5	49.1	59	61.7	78.5	*5.56*	32	**4.56**
david	12.3	8.06	12	11.5	12.9	8.36	5.45	14.4	*5.18*	8.1	**4.78**
david3	7.22	**6.04**	8.16	7.08	7.07	7.01	53.8	64.6	*6.29*	8.28	7.98
deer	4.97	4.89	4.89	*4.83*	4.95	4.9	12.6	209	**4.48**	4.9	4.87
doll	41.8	17.2	31.7	23.8	34.5	35.5	72.4	17.9	**8.82**	28.5	*9.59*
faceocc1	12.5	13.5	*12.2*	13.6	12.5	13	73.6	**11.1**	72.1	50.1	13
football1	28.3	25.2	28.1	28.6	19.9	10.9	**7.72**	39.6	17.2	27.8	*10.6*
girl	12.7	13.3	13.2	*12.4*	12.8	**12.3**	18.8	133	13.4	12.9	13.1
human	7.49	**7.31**	*7.48*	7.96	7.5	7.79	15.5	111	7.58	7.89	7.9
ironman	212	**39.9**	205	184	178	204	56.9	98.7	197	*53.2*	55.5
jogging-1	111	106	111	112	116	110	*104*	123	142	105	**103**
jogging-2	158	*156*	156	157	157	156	**151**	157	157	156	158
kitesurf	59.3	59.9	58.8	57.9	59.3	58.2	*10.7*	36.3	65.3	57.5	**3.56**
lemming	**59.9**	82.1	141	80.7	108	81	81.6	156	*80.2*	96.8	83.2
liquor	166	98	152	**96.9**	165	119	137	110	*97*	134	115
matrix	71.6	*50.4*	**45.7**	70.7	65.7	71.3	53.2	57.9	53	80.3	68.8
motorRolling	553	535	527	536	533	528	*379*	387	389	520	**369**
mountainBike	6.55	*6.27*	6.53	6.48	6.64	6.49	6.89	**5.64**	6.98	6.58	6.62
panda	95.8	**61.6**	68.7	98.5	77.8	117	139	*62.9*	76.8	110	92.6
shaking	13.3	12.7	14.9	14.8	13.4	*12.1*	13.2	21.2	**8.02**	12.5	14.3
shirt	16.1	**9.76**	11.7	14.7	15.1	16.1	18.7	15.5	16.6	14.3	*9.78*
singer1	16.6	14	13.6	14.5	16.5	16.2	11.9	27.4	*10.1*	13	**7.66**
singer2	170	201	169	14.8	171	*6.55*	**5.18**	172	24.4	8.79	178
skating1	8.45	**7.62**	8.01	8.27	8.08	14.1	*7.96*	22	8.79	8.14	9.01
skiing	275	275	275	274	275	274	**210**	*220*	273	274	275
soccer	134	119	120	120	132	63.4	32.1	104	**10.7**	145	*28.6*
stone	*1.75*	2.53	1.95	**1.75**	1.97	2.8	5.62	118	2.01	2.74	2.18
subway	157	168	165	165	*157*	*157*	157	161	**153**	166	157
surfer	4.09	4.23	4.06	*4.01*	4.05	18.6	5.67	22.2	18.7	4.71	**3.88**
tiger1	48.7	19.7	22.5	29.8	21.4	21.4	*17.7*	86.6	114	21	**13.3**
tiger2	56.4	21.2	25.9	103	41.4	32.5	**18.2**	29.7	24.8	29.8	*18.8*
trellis	22.5	14.9	20.8	21.1	20.9	20.1	*9.62*	57.8	**9.45**	15.4	33.3
walking2	27.3	46.9	**18**	*19*	47.3	47.3	39.8	34.5	41.4	36.5	38.6
woman	140	**8.31**	9.33	8.93	140	140	10.3	142	9.4	*8.77*	11
Median	42.3	22.3	25.6	24.8	37.9	23.5	27.4	63.8	*17.9*	28.2	**13.8**

Table 3: The per-video average center location error (CLE) (in pixels) for different color features with our proposed learning scheme.

Video	Int+RGB	LAB	YCbCr	Int+rg	Opp	C	HSV	Int+SO	Int+AOpp	Int+HUE	Int+CN
basketball	**100**	40.4	**100**	83.7	**100**	**100**	2.76	64.8	*93.2*	43.2	51.4
board	8.57	**63.6**	*61.4*	9.29	8.57	9.29	8.57	8.57	**63.6**	4.29	**63.6**
bolt	5.43	3.14	5.43	5.43	5.43	**100**	**100**	*48.9*	**100**	5.43	**100**
boy	66.9	84.6	66.9	84.6	66.9	67.1	66.6	33.2	**100**	**100**	*99.7*
carDark	**100**	**100**	**100**	**100**	**100**	**100**	**100**	*69*	**100**	**100**	**100**
carScale	71.4	71.8	71.4	71.4	71.4	71.4	*73*	**78.6**	71.4	71	71.8
coke	*86.3*	**91.8**	85.9	84.5	*86.3*	83.5	63.9	15.5	53.3	85.2	84.9
couple	8.57	*10.7*	8.57	8.57	8.57	*10.7*	8.57	**11.4**	*10.7*	*10.7*	*10.7*
crossing	24.2	*56.7*	34.2	51.7	34.2	26.7	24.2	11.7	**100**	*56.7*	**100**
david	83.2	**100**	83.7	84.7	83	*99.8*	**100**	77.5	**100**	**100**	**100**
david3	91.3	**92.5**	91.3	91.7	91.7	91.7	74.2	57.1	*92.1*	90.9	91.3
deer	**100**	**100**	**100**	**100**	**100**	**100**	*87.3*	5.63	**100**	**100**	**100**
doll	50.1	79	56.7	57.1	47.8	57.3	55.4	60.4	**92.4**	58	*91.5*
faceocc1	92.4	87.7	*92.5*	87.8	92.4	89.5	42.6	**98.4**	54	56.5	89.1
football1	51.4	52.7	51.4	51.4	52.7	78.4	**98.6**	35.1	67.6	54.1	*81.1*
girl	88.8	85.8	83.6	86.4	**91.8**	86.4	63.6	3.6	83	86.4	*89*
human	**100**	**100**	**100**	**100**	**100**	**100**	*82.3*	39.1	**100**	**100**	**100**
ironman	14.5	*47*	14.5	14.5	14.5	16.3	**57.8**	11.4	15.1	41.6	46.4
jogging-1	22.8	*23.1*	22.8	22.8	*23.1*	*23.1*	22.8	*23.1*	*23.1*	*23.1*	**24.4**
jogging-2	*18.6*	**18.9**	**18.9**	*18.6*	**18.9**	**18.9**	*18.6*	**18.9**	*18.6*	*18.6*	*18.6*
kitesurf	45.2	*46.4*	45.2	*46.4*	45.2	*46.4*	**100**	16.7	45.2	*46.4*	**100**
lemming	**46.9**	34	31.2	*42.7*	39.9	38.1	31.7	17.1	40.5	32.1	28.5
liquor	22.3	*28.4*	26.5	**34.9**	22.3	27.5	27.3	19.5	27.7	28	28.1
matrix	12.0	34.0	34.0	*35.0*	12.0	15.0	34.0	17.0	29.0	**41.0**	1.00
motorRolling	6.71	5.49	4.88	6.71	6.71	6.71	**9.76**	4.88	4.88	4.27	*7.93*
mountainBike	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**
panda	52.3	53.1	52.3	52.3	52.3	53.1	53.1	40.7	*53.5*	**53.9**	52.7
shaking	82.2	87.4	67.7	68.8	82.7	*90.4*	81.9	47.9	**98.4**	86.8	69.6
shirt	81.2	*89.1*	84.2	81.4	81.3	80.8	75.9	81.3	81.5	83.5	**89.6**
singer1	50.7	66.7	75.2	66.1	51.3	51.3	95.2	40.2	*96.9*	88.9	**100**
singer2	3.55	3.55	3.55	88.3	3.55	*99.5*	**100**	3.55	71.6	**100**	3.55
skating1	98.8	**100**	98.5	97.8	**100**	88.5	96.8	85.3	**100**	**100**	*99.5*
skiing	**13.6**	**13.6**	**13.6**	**13.6**	**13.6**	**13.6**	*7.41*	**13.6**	**13.6**	**13.6**	**13.6**
soccer	18.9	18.6	18.6	18.6	18.9	18.9	33.2	7.65	**91.6**	18.6	*39.8*
stone	**100**	**100**	**100**	**100**	**100**	**100**	*99.2*	15.1	**100**	**100**	**100**
subway	**24.6**	**24.6**	**24.6**	**24.6**	**24.6**	**24.6**	*24*	7.43	**24.6**	**24.6**	**24.6**
surfer	**100**	**100**	**100**	**100**	**100**	63.2	*96.1*	60.5	40.8	**100**	**100**
tiger1	39	65	37	37.8	56.7	57	*66.5*	12.9	16.3	40.7	**81.7**
tiger2	24.7	55.3	42.7	11	29.9	25.8	**74**	34.5	46.6	31.2	*64.7*
trellis	54.3	86.6	63.1	61.3	62.6	64.9	**97.4**	30.6	*93.8*	85.1	75.9
walking2	43.4	43.2	**47.4**	*45.4*	43.2	43.2	40.4	42.8	40.4	43.2	41.8
woman	25	**94**	**94**	**94**	25	25	**94**	*38.7*	**94**	**94**	**94**
Median	50.4	64.3	59	59.2	51.8	60.2	66.5	31.9	*71.5*	56.6	**81.4**

Table 4: The per-video distance precision (DP) (%) for different color features using our proposed learning scheme (as in Figure 2 in the paper).

Video	CT [16]	LSST [14]	Frag [1]	L1APG [3]	LOT [11]	ASLA [9]	TLD [10]	SCM [17]	EDFT [5]	CSK [8]	DFT [13]	CXT [4]	CPF [12]	LSHT [7]	Struck [6]	CN_2	CN
basketball	171	118	11.5	138	69.2	33.7	65.2	37.9	108	**6.6**	18	177	55.2	156	159	*9.29*	39.5
board	53.3	120	91.9	220	156	66.5	131	97.2	98.5	109	98.4	114	51.6	**18.3**	28.8	28	*24*
bolt	371	380	259	408	13.7	399	88	441	355	429	367	378	13.1	122	391	**4.2**	*4.43*
boy	32.1	59.2	33.9	7.03	66	84.4	4.09	54.9	**2.34**	20.1	106	11.1	4.84	32.6	*3.35*	4.39	4.64
carDark	120	1.32	37.7	**1.04**	28.5	1.37	26.9	1.31	18.5	2.71	58.8	19.1	53.1	26.9	*1.28*	2.83	2.79
carScale	74	66.2	15	79.8	102	*12.5*	34.3	17.4	76.4	83	75.8	55.4	30.2	**11.2**	34.2	25.2	26
coke	39.1	92.7	125	50.4	101	61.5	32.1	82.2	68.9	*13.6*	70.7	25.6	54.3	32.1	**12.1**	30.8	19.2
couple	77.8	103	**9.79**	28.4	37	87.4	64.3	28.4	89.4	144	109	50.5	34.7	114	*12.7*	123	123
crossing	8.08	2.51	57.7	63.4	34.1	**1.62**	27	*1.66*	7.61	9.18	22.3	27.2	9.88	29.9	2.63	4.29	4.56
david	14.3	16.6	93	14	24	**4.32**	34.3	78.1	9.2	17.7	42.9	7.55	25.2	14.8	43.2	7.73	*4.78*
david3	68.5	16	12.9	90	9.52	87.4	136	66.7	**6.46**	56	50.9	222	18.9	53.7	107	9.11	*7.98*
deer	245	196	112	24.2	66.6	153	118	79.6	15.7	5.23	98.7	12.5	86.5	203	6.85	*5.11*	**4.87**
doll	25.8	36.6	9.49	5.89	6.46	7.48	6.75	**2.76**	96.4	44.7	59.5	*3.78*	8.63	19.3	8.68	7.2	9.59
faceocc1	26.2	14.4	**11.2**	17.3	34.6	69.4	33.5	13	40.3	*12.6*	23.6	23.5	28.8	29.7	19.1	13.5	13
football1	16.7	9.41	14.6	9.2	6.2	13.3	52.9	38.4	**1.74**	16.2	*1.97*	4.8	12.7	3.79	28.6	9.85	10.6
girl	19.2	24	20.1	*2.8*	22.9	3.52	8.32	24.7	18	19.4	24	8.51	18.1	41.9	**2.74**	12.5	13.1
human	428	*1.78*	8.76	3.31	2.38	**1.69**	110	2.69	5.77	7.6	5.87	209	4.38	6.59	5.36	7.25	7.9
ironman	175	149	256	163	*89.3*	162	93	172	231	185	240	150	97.9	235	100	173	**55.5**
jogging-1	92.4	79.2	27.6	89.5	90.9	142	7.21	138	116	135	31.4	**5.79**	19.7	*5.99*	7.86	101	103
jogging-2	140	**4.04**	33.6	146	14.4	170	*12.1*	142	143	164	33.5	135	13.7	148	138	171	158
kitesurf	89	45.6	98.1	62.4	98.4	23.8	39.9	44	*3.58*	64	5.11	30	125	21.1	25.5	3.79	**3.56**
lemming	134	171	98.8	178	20.2	182	16.8	189	80.8	114	77.8	**8.78**	*11.5*	82.4	39.7	90.7	83.2
liquor	179	145	102	213	**8.71**	42.3	40.3	110	215	155	221	136	*22*	41	133	326	115
matrix	69	104	164	61.7	83.3	61.4	*57.2*	**53.7**	415	114	106	182	127	66.5	216	79.2	68.8
motorRolling	166	138	142	207	*132*	206	**87.8**	172	183	393	174	133	151	194	143	441	369
mountainBike	193	133	209	8.25	25	7.28	213	10.1	7.35	*6.58*	155	170	212	**5.62**	8.58	6.78	6.62
panda	148	114	114	118	79.5	64.9	**15.1**	104	38.7	94	171	*36.6*	52.5	88	89.7	92.4	92.6
shaking	126	89.8	196	110	75	23.1	68.5	22.3	104	16.7	26.3	135	187	15.7	21.9	*15.1*	**14.3**
shirt	15.1	64.5	18.8	7.14	**5.73**	62.9	22.6	18.5	47.3	16.8	44.1	6.48	7.98	26.7	*6.48*	10.6	9.78
singer1	15	**2.62**	77.1	53.4	148	3.03	10.6	*2.94*	16.6	15.4	18.8	11.5	6.63	21	12.4	9.21	7.66
singer2	146	*18.5*	72.5	181	75.3	20.2	56.7	176	20.6	185	21.8	199	52.8	**9.32**	172	167	178
skating1	184	155	138	159	111	63.7	104	49.2	199	**7.78**	174	137	109	82.3	82.3	*7.95*	9.01
skiing	277	262	272	266	245	251	264	*222*	283	248	276	**94.4**	261	260	261	275	275
soccer	69.1	175	152	101	57.1	81	69.4	73.8	243	70.9	140	50.5	50.7	82.9	73.5	**8.68**	*28.6*
stone	108	36	34.8	32.5	27	**1.79**	18.6	2.37	2.17	4.57	27.7	7.8	32.3	4.82	*1.79*	2.54	2.18
subway	7.2	139	16.2	148	12.5	3.89	86.8	158	3.46	164	*3.31*	140	140	4.39	**3.03**	157	157
surfer	60	18.8	21.6	6.53	9.39	52.1	**3.73**	73.3	5.21	4.51	219	7.59	30.3	19.7	8	4.15	*3.88*
tiger1	79	107	69	58.4	107	65.9	70	73.7	62.1	69.5	41.3	75	37.6	154	*15.6*	16	**13.3**
tiger2	69.4	77.5	118	65.2	150	101	38	107	196	59.6	**12.2**	37.2	66.3	86.3	20.1	*18.6*	18.8
trellis	51.1	44.9	63.7	62.2	47.9	*7.8*	55.9	**6.84**	59.6	18.9	44.9	19.6	44.5	61.2	15.3	20.8	33.3
walking2	64.6	57.5	57.3	*5.06*	64.8	37.3	63.1	**1.58**	28.7	18.5	29.1	31.1	54.9	50.6	12.9	47.7	38.6
woman	116	146	107	129	122	*8.27*	78.9	8.71	16.7	208	8.5	116	133	8.8	**3.44**	302	11
Median	78.4	78.4	70.8	62.9	60.9	56.8	54.4	54.3	53.5	50.3	47.9	43.8	41.1	32.3	19.6	*14.3*	**13.8**

Table 5: The per-video average center location error (CLE) (in pixels) for all the trackers in the state-of-the-art comparison (as in table 3 in the paper). The best two results are obtained with the approaches proposed in this paper.

Video	CT [16]	LSST [14]	Frag [1]	L1APG [3]	LOT [11]	ASLA [9]	TLD [10]	SCM [17]	EDFT [5]	CSK [8]	DFT [13]	CXT [4]	CPF [12]	LSHT [7]	Struck [6]	CN_2	CN
basketball	4.14	12.6	82.2	30.6	63.6	83.6	51.3	56.1	30.5	**100**	89.2	6.21	74.8	5.1	11.6	*99.9*	51.4
board	15.7	15	47.9	4.29	12.1	17.1	10	26.4	8.57	8.57	8.57	10	31.4	*71.4*	**77.1**	63.6	63.6
bolt	2.57	1.71	14	1.71	90.3	1.71	32	1.43	2.57	3.43	4.29	2.57	*93.7*	37.4	2.86	**100**	**100**
boy	66.6	44	57.5	92.9	66.6	44	**100**	44	**100**	84.4	48.5	73.9	**100**	56.3	**100**	*99.8*	99.7
carDark	0.763	**100**	55	**100**	63.4	**100**	64.6	**100**	*71.5*	**100**	54.5	70.7	24.7	62.1	**100**	**100**	**100**
carScale	64.7	64.7	68.3	64.7	47.6	*73.4*	64.3	68.3	64.7	65.1	65.1	63.5	66.7	**85.3**	64.3	72.2	71.8
coke	12.7	12.4	3.44	26.5	12	16.2	66	15.5	15.1	*87.3*	8.59	69.4	35.7	65.3	**94.8**	61.5	84.9
couple	30.7	10.7	**90.7**	60.7	63.6	22.9	31.4	50.7	21.4	8.57	8.57	57.9	68.6	10.7	*83.6*	10.7	10.7
crossing	*98.3*	**100**	40	25	65.8	**100**	57.5	**100**	**100**	**100**	68.3	57.5	89.2	55.8	**100**	**100**	**100**
david	79.2	69.2	9.55	*80.5*	34.4	**100**	65.8	29.1	**100**	49.9	31.4	**100**	24.6	76	32.7	**100**	**100**
david3	43.3	54	79	46	*98.4*	58.3	35.3	34.5	**100**	65.9	74.6	15.9	54.4	75	33.7	90.5	91.3
deer	2.82	2.82	15.5	71.8	22.5	4.23	28.2	12.7	63.4	**100**	31	*85.9*	5.63	4.23	**100**	**100**	**100**
doll	56.8	73.3	90.2	97.4	95.2	92.3	98.5	**99.6**	60.9	58.1	42.7	*99.3*	94	37.4	90.9	97.6	91.5
faceocc1	23.5	86.9	**98.1**	68.3	26.2	21.2	17.3	*93.6*	62.8	92.5	62.2	38.1	30.6	63.9	57	87.7	89.1
football1	75.7	86.5	60.8	83.8	**100**	83.8	51.4	51.4	**100**	75.7	**100**	**100**	83.8	*98.6*	50	87.8	81.1
girl	50.4	58.8	65.2	**100**	64	**100**	*94*	44.6	73.2	55.4	29.6	84.6	77.2	21.2	**100**	86.4	89
human	0.485	**100**	*98.8*	**100**	**100**	**100**	42.2	**100**	**100**	**100**	**100**	23.5	**100**	**100**	**100**	**100**	**100**
ironman	12	4.82	3.61	10.8	10.2	*21.7*	12.7	16.9	4.22	13.3	9.04	3.01	4.82	3.61	5.42	14.5	**46.4**
jogging-1	23.1	22.8	66.1	22.8	21.2	23.1	**97.4**	22.8	22.8	22.8	21.5	96.1	50.2	*97.1*	**97.4**	23.8	24.4
jogging-2	0.651	**99.3**	57.7	18.6	84.7	18.2	*95.4*	16	15.6	18.6	16.3	16.6	86	16.3	18.6	18.6	18.6
kitesurf	45.2	40.5	31	42.9	3.57	40.5	46.4	32.1	**100**	46.4	*97.6*	54.8	4.76	51.2	57.1	**100**	**100**
lemming	7.71	16.8	47.4	17.2	80.5	16.7	80.2	16.6	48.8	43.6	51.6	**92.3**	*86.4*	42.2	63.2	30.8	28.5
liquor	20.9	20.6	28.5	20.6	**93.2**	68.2	46.3	33.7	22.4	22.3	22.1	20.7	*72.5*	59.6	39.1	20.2	28.1
matrix	11.0	8.00	7.00	12.0	**35.0**	10.0	16.0	*22.0*	10.0	1.00	6.00	8.00	7.00	16.0	11.0	2.00	1.00
motorRolling	2.44	3.05	7.32	3.66	4.88	5.49	**12.8**	4.27	4.27	4.88	4.27	4.27	4.88	4.27	*7.93*	4.27	*7.93*
mountainBike	25.9	46.5	14	94.7	70.2	92.1	25	99.1	**100**	**100**	35.1	31.6	14.5	**100**	*99.6*	**100**	**100**
panda	22.4	24.1	46.1	2.07	24.1	52.7	83.4	17.8	**86.7**	53.5	22.4	*84.2*	60.2	15.8	52.7	53.1	52.7
shaking	0.822	1.1	9.59	4.11	16.4	31.8	3.29	55.6	17	59.5	**83**	12.6	17.8	*69.9*	53.7	*69.9*	69.6
shirt	81.4	0.631	75.3	93.8	**99.7**	0.526	79.8	78.5	11.7	80.8	11.3	99.3	*99.6*	69	98.2	88.3	89.6
singer1	86.9	**100**	22.8	37.9	21.4	**100**	**100**	**100**	49.3	59	50.7	96.6	**100**	40.2	*98.3*	95.7	**100**
singer2	3.28	63.9	18.6	3.55	19.4	*68*	8.47	3.55	64.5	3.55	59.6	10.9	9.29	**98.1**	3.83	3.55	3.55
skating1	8.5	11.5	18	13.3	27.8	76.5	27	79.3	16.3	98.8	19	20	33	56	51	**100**	*99.5*
skiing	2.47	12.3	3.7	8.64	2.47	*13.6*	12.3	12.3	11.1	9.88	7.41	**21**	6.17	11.1	4.94	*13.6*	*13.6*
soccer	20.9	16.8	8.93	20.7	33.2	10.7	10.2	19.6	17.9	13.5	22.4	24.5	28.1	9.18	21.9	**96.7**	*39.8*
stone	11.8	65.5	50.4	65.5	57.1	**100**	83.2	**100**	**100**	89.9	65.5	88.2	45.4	**100**	*99.2*	**100**	**100**
subway	98.3	21.7	72	25.1	82.9	98.3	60	24.6	**100**	24.6	**100**	25.7	23.4	**100**	*99.4*	24.6	24.6
surfer	3.95	63.2	43.4	*97.4*	86.8	7.89	**100**	5.26	**100**	**100**	3.95	90.8	51.3	63.2	*97.4*	**100**	**100**
tiger1	1.43	9.74	35.2	34.1	16	15.2	23.8	16	32.4	25.5	67	17.5	38.4	7.16	*78.8*	71.3	**81.7**
tiger2	11.2	9.59	13.4	27.1	17.5	12.9	35.3	9.59	26.3	11	**80.3**	36.7	14.5	11.8	*69*	63.3	64.7
trellis	20.9	39.9	37.4	17.6	30.9	*85.8*	44.5	**87.5**	47.6	81	50.6	65.2	22	44.8	73.5	68.9	75.9
walking2	40	41.4	34.8	*97.6*	39.2	40.2	37.6	**100**	40.2	46	40.2	40.8	34.6	39.8	85.6	42.4	41.8
woman	20.6	20.3	18.8	20.4	15.9	93	40.2	94	93.8	25	*95.1*	12.9	18.6	94	**99.7**	25	94
Median	20.8	23.4	38.7	28.9	37.1	42.2	45.4	34.1	49	54.5	41.4	39.5	37.1	55.9	71.3	*79.3*	**81.4**

Table 6: The per-video distance precision (DP) (%) for all the trackers in the state-of-the-art comparison (as in table 3 in the paper). The best two results are achieved using the approaches proposed in this paper.

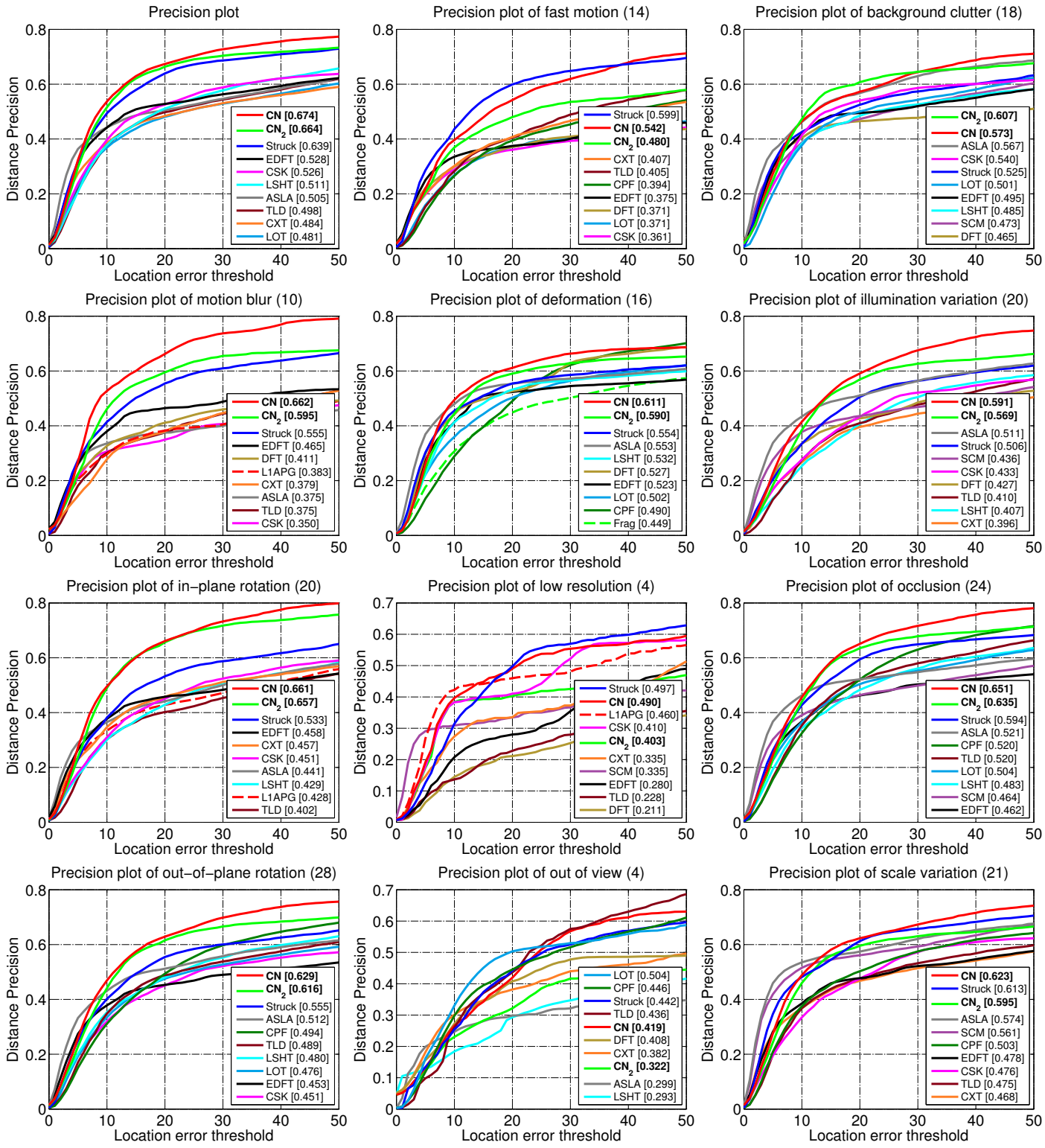

Figure 1: Precision plots of the overall and attribute-based evaluations showing the state-of-the-art comparisons (as in section 4.6 in the paper). The value appearing in the title denotes the number of videos associated with the respective attribute. The ranking score of each tracker is reported in the legend. Overall, the two methods proposed in this paper perform favorably against state-of-the-art algorithms.

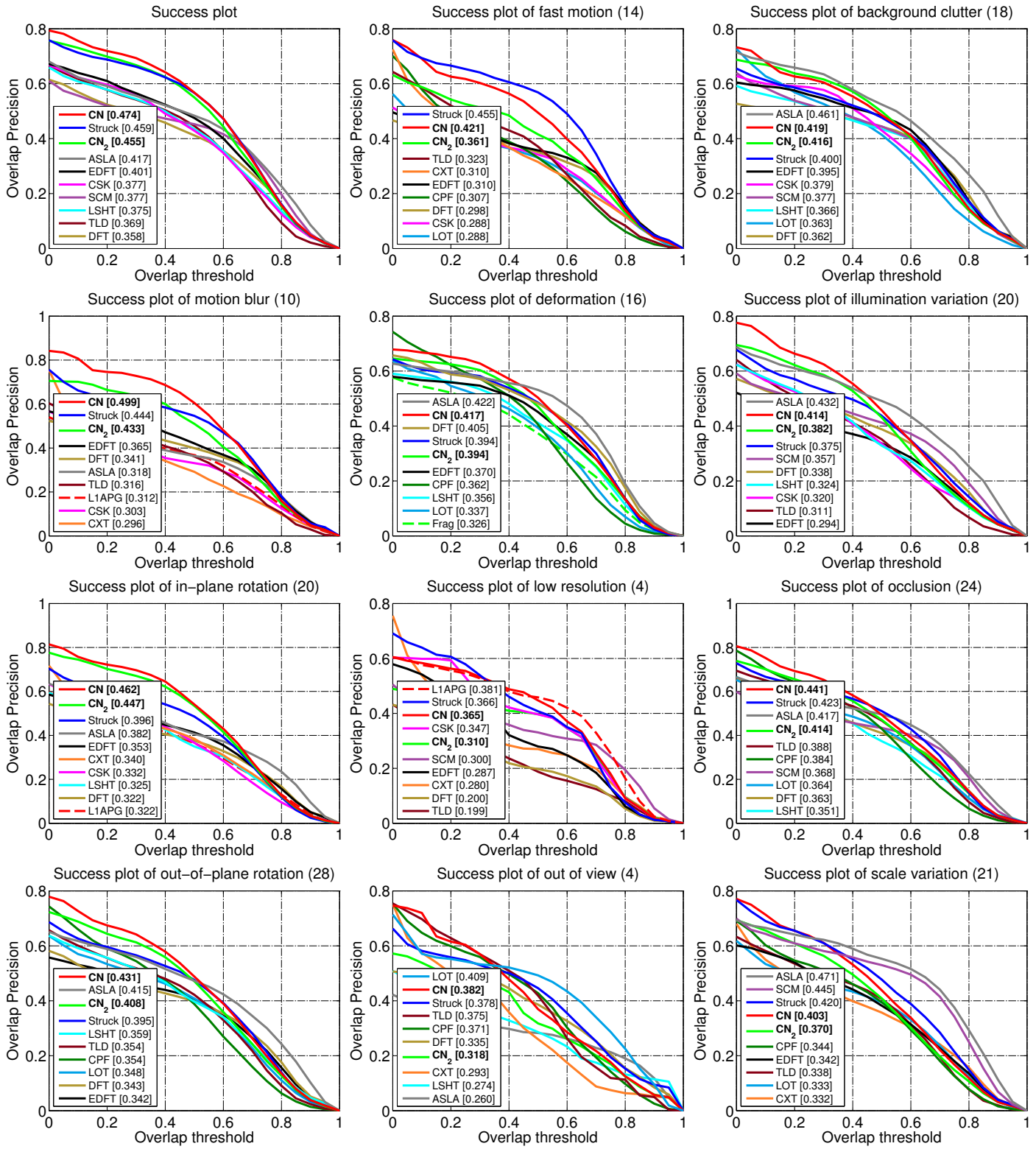

Figure 2: Success plots of the overall and attribute-based evaluations showing the state-of-the-art comparisons (as in section 4.6 in the paper). The value appearing in the title denotes the number of videos associated with the respective attribute. The ranking score of each tracker is reported in the legend. Note that our baseline CSK tracker does not estimate scale variations. Despite this inherent limitation, our two approaches perform favorably against state-of-the-art algorithms.

References

[1] A. Adam, E. Rivlin, and Shimshoni. Robust fragments-based tracking using the integral histogram. In *CVPR*, 2006. 8, 9

[2] B. Babenko, M.-H. Yang, and S. Belongie. Visual tracking with online multiple instance learning. In *CVPR*, 2009. 3

[3] C. Bao, Y. Wu, H. Ling, and H. Ji. Real time robust l1 tracker using accelerated proximal gradient approach. In *CVPR*, 2012. 8, 9

[4] T. B. Dinh, N. Vo, and G. Medioni. Context tracker: Exploring supporters and distracters in unconstrained environments. In *CVPR*, 2011. 8, 9

[5] M. Felsberg. Enhanced distribution field tracking using channel representations. In *ICCV Workshop*, 2013. 8, 9

[6] S. Hare, A. Saffari, and P. Torr. Struck: Structured output tracking with kernels. In *ICCV*, 2011. 8, 9

[7] S. He, Q. Yang, R. Lau, J. Wang, and M.-H. Yang. Visual tracking via locality sensitive histograms. In *CVPR*, 2013. 8, 9

[8] J. Henriques, R. Caseiro, P. Martins, and J. Batista. Exploiting the circulant structure of tracking-by-detection with kernels. In *ECCV*, 2012. 1, 2, 8, 9

[9] X. Jia, H. Lu, and M.-H. Yang. Visual tracking via adaptive structural local sparse appearance model. In *CVPR*, 2012. 8, 9

[10] Z. Kalal, J. Matas, and K. Mikolajczyk. P-n learning: Bootstrapping binary classifiers by structural constraints. In *CVPR*, 2010. 8, 9

[11] S. Oron, A. Bar-Hillel, D. Levi, and S. Avidan. Locally orderless tracking. In *CVPR*, 2012. 8, 9

[12] P. Perez, C. Hue, J. Vermaak, and M. Gangnet. Color-based probabilistic tracking. In *ECCV*, 2002. 8, 9

[13] L. Sevilla-Lara and E. G. Learned-Miller. Distribution fields for tracking. In *CVPR*, 2012. 8, 9

[14] D. Wang, H. Lu, and M.-H. Yang. Least soft-threshold squares tracking. In *CVPR*, 2013. 8, 9

[15] Y. Wu, J. Lim, and M.-H. Yang. Online object tracking: A benchmark. In *CVPR*, 2013. 3

[16] K. Zhang, L. Zhang, and M. Yang. Real-time compressive tracking. In *ECCV*, 2012. 8, 9

[17] W. Zhong, H. Lu, and M.-H. Yang. Robust object tracking via sparsity-based collaborative model. In *CVPR*, 2012. 8, 9