

Supplementary Material

Accurate Scale Estimation for Robust Visual Tracking

1 Results

Here we provide additional experimental results. Table 1, 2 and 3 contain the per-video results of our fast scale search tracker, our exhaustive scale search tracker and the compared state-of-the-art trackers.

Video	Ours Fast	Ours Exhaustive	ASLA [6]	SCM [11]	Struck [3]	TLD [7]	EDFT [2]	LIAPG [1]	DFT [9]	LOT [8]	CSK [5]	LSHT [4]	CT [10]
boy	100	100	43.5	43.9	97.5	82.9	98	89.9	48.3	65.6	84.2	50.7	56.5
car4	100	39.8	100	37.5	39.9	24	27.5	29.6	25.8	1.06	27.6	27.6	18.2
carScale	84.5	84.1	71	66.7	43.3	68.7	44.8	58.3	44.8	46.4	44.8	44.8	44.8
couple	10.7	10.7	22.1	47.1	60.7	22.9	21.4	60	8.57	57.9	8.57	9.29	29.3
crossing	100	100	100	100	95.8	45.8	75	25	64.2	32.5	28.3	40	76.7
david	100	97.5	94.9	28	23.6	61.1	55.4	69.2	23.4	12.5	23.6	28.2	27.2
dog1	100	53.5	89.9	85.3	65.2	75.6	64.4	84.3	52.1	98.6	65.3	54.3	59.8
doll	99.6	99.7	92.2	99.3	68.9	69.3	49.3	32.6	35	85.7	21.9	23	33.2
dudek	98.1	88.6	90.5	80.3	98.1	67	82.3	79.4	80.1	60.9	95.5	89.9	86.8
fleetface	66.5	68.9	64.5	86.6	78.1	44.1	54.7	54.7	55.6	57.6	66.9	65.5	51.9
freeman1	36.8	42.6	32.8	32.2	20.2	23.3	12.6	14.1	17.8	16.9	13.2	18.4	11.7
freeman3	31.3	30	91.7	99.8	17.6	64.6	28.9	33.3	33	10.2	33	15.7	0.217
freeman4	41.7	23	17	57.6	18.7	21.6	17	33.9	18	21.6	17	20.1	0.707
girl	24.2	25.4	86.8	35.6	97	72.6	48.6	97	25.2	55.2	39	14.4	6.6
ironman	13.3	7.83	15.1	9.04	4.82	8.43	4.22	7.23	3.61	9.64	12.7	2.41	6.02
lemming	26.9	26.6	16.7	16.6	63.8	63.4	44.8	17.2	47.4	74.3	42.8	40.7	7.71
liquor	40.9	43.6	68.8	37	40.5	75.5	23.3	20.6	22.9	96.6	28	60.1	20.9
matrix	18	23	7	17	11	7	9	10	6	24	1	2	6
motorRolling	6.71	6.71	9.76	9.15	15.2	17.7	6.1	6.71	6.1	5.49	7.32	9.15	10.4
shaking	100	100	23.3	55.6	52.9	3.29	17.5	3.84	82.5	10.7	61.6	69.9	0.822
singer1	100	36.2	100	100	29.9	98.6	27.6	37.9	27.6	22.5	27.6	27.6	27.6
skating1	54.8	51	73.3	64.8	31.3	9	15.5	13	16.3	26.3	36.8	18.3	7.75
skiing	4.94	7.41	12.3	9.88	4.94	7.41	9.88	8.64	6.17	1.23	7.41	3.7	1.23
soccer	52.8	59.7	10.5	18.1	18.1	11.2	17.6	19.9	21.9	23.7	13.8	9.18	18.9
trellis	96.8	98.8	85.8	87.2	54	40.6	47.6	15.3	51.8	31.6	59.1	44.3	17
walking	99.8	88.8	99.8	95.1	52.4	35.2	55.1	99.8	55.1	96.8	51.9	54.4	0.243
walking2	100	59.4	39.8	100	43	20.8	38.2	97.6	38.2	39	38.6	38.4	37.4
woman	93.3	93.8	88.6	78.2	93.5	32.8	93.3	19.6	93.5	10.1	24.3	83.9	17.6
Median	75.5	52.2	69.9	56.6	43.1	37.9	33.6	31.1	30.3	28.9	28.2	27.9	17.9

Table 1: The per-video overlap precision (OP) (%) for our fast scale tracker, our exhaustive scale tracker and the compared state-of-the-art trackers.

Video	Ours Fast	Ours Exhaustive	ASLA [6]	SCM [11]	Struck [3]	TLD [7]	EDFT [2]	LIAPG [1]	DFT [9]	LOT [8]	CSK [5]	LSHT [4]	CT [10]
boy	100	100	44	44	100	100	100	92.9	48.5	66.6	84.4	56.3	66.6
car4	100	100	100	87.9	100	23.8	29	30.2	26.3	0.759	35.8	58	3.79
carScale	75.8	75	73.4	68.3	64.3	64.3	64.7	64.7	65.1	47.6	65.1	85.3	64.7
couple	10.7	10.7	22.9	50.7	83.6	31.4	21.4	60.7	8.57	63.6	8.57	10.7	30.7
crossing	100	100	100	100	100	57.5	100	25	68.3	65.8	100	55.8	98.3
david	100	100	100	29.1	32.7	65.8	100	80.5	31.4	34.4	49.9	76	79.2
dog1	100	100	99.9	93.1	99.3	98.8	98.2	100	62.4	99.1	100	85.3	77
doll	99.3	99.3	92.3	99.6	90.9	98.5	60.9	97.4	42.7	95.2	58.1	37.4	56.8
dudek	80.8	83.8	83	75	90.6	64.3	73.8	69.5	66.2	53.5	78.1	66.3	38.4
fleetface	57.6	62.9	60.3	60.4	59.7	48.1	35.8	41	35.8	55.3	60.8	57.9	38
freeman1	95.4	81.3	40.2	37.1	86.2	71.5	54.9	38.3	94.2	48.5	55.5	96.3	39.9
freeman3	91.3	91.5	100	100	80.2	83.5	100	74.6	65	49.8	57.2	34.6	25.9
freeman4	67.8	30.4	24	77.7	41.3	37.5	19.8	62.9	30	53.4	19.1	90.1	25.4
girl	92.8	100	100	44.6	100	94	73.2	100	29.6	64	55.4	21.2	50.4
ironman	15.1	10.2	21.7	16.9	5.42	12.7	4.22	10.8	9.04	10.2	13.3	3.61	12
lemming	39.1	27.3	16.7	16.6	63.2	80.2	48.8	17.2	51.6	80.5	43.6	42.2	7.71
liquor	40.4	43.3	68.2	33.7	39.1	46.3	22.4	20.6	22.1	93.2	22.3	59.6	20.9
matrix	25	35	10	22	11	16	10	12	6	35	1	16	11
motorRolling	4.88	4.27	5.49	4.27	7.93	12.8	4.27	3.66	4.27	4.88	4.88	4.27	2.44
shaking	100	100	31.8	55.6	53.7	3.29	17	4.11	83	16.4	59.5	69.9	0.822
singer1	100	100	100	100	98.3	100	49.3	37.9	50.7	21.4	59	40.2	86.9
skating1	97.5	100	76.5	79.3	51	27	16.3	13.3	19	27.8	98.8	56	8.5
skiing	13.6	13.6	13.6	12.3	4.94	12.3	11.1	8.64	7.41	2.47	9.88	11.1	2.47
soccer	84.7	74.5	10.7	19.6	21.9	10.2	17.9	20.7	22.4	33.2	13.5	9.18	20.9
trellis	100	100	85.8	87.5	73.5	44.5	47.6	17.6	50.6	30.9	81	44.8	20.9
walking	100	100	100	100	100	42.2	100	100	100	100	100	100	12.1
walking2	100	64.6	40.2	100	85.6	37.6	40.2	97.6	40.2	39.2	46	39.8	40
woman	93.8	94	93	94	99.7	40.2	93.8	20.4	95.1	15.9	25	94	20.6
Median	93.3	87.6	70.8	64.3	76.8	45.4	48.2	38.1	41.4	48	55.5	55.9	25.7

Table 2: The per-video distance precision (DP) (%) for our fast scale tracker, our exhaustive scale tracker and the compared state-of-the-art trackers.

Video	Ours Fast	Ours Exhaustive	ASLA [6]	SCM [11]	Struck [3]	TLD [7]	EDFT [2]	LIAPG [1]	DFT [9]	LOT [8]	CSK [5]	LSHT [4]	CT [10]
boy	2.09	<i>2.29</i>	84.4	54.9	3.35	4.09	2.34	7.03	106	66	20.1	32.6	32.1
car4	1.86	5.15	<i>2.18</i>	13.1	3.96	87	94.1	77	61.9	172	18.8	16.7	80.5
carScale	19.1	18.9	<i>12.5</i>	17.4	34.2	34.3	76.4	79.8	75.8	102	83	11.2	74
couple	119	105	87.4	<i>28.4</i>	12.7	64.3	89.4	28.4	109	37	144	114	77.8
crossing	1.54	1.84	<i>1.62</i>	1.66	2.63	27	7.61	63.4	22.3	34.1	9.18	29.9	8.08
david	3.79	5.11	<i>4.32</i>	78.1	43.2	34.3	9.2	14	42.9	24	17.7	14.8	14.3
dog1	4.88	4.02	4.87	7.04	6.02	4.08	6.48	3.72	41.2	4.51	<i>3.79</i>	9.63	14.7
doll	3.06	2.47	7.48	<i>2.76</i>	8.68	6.75	96.4	5.89	59.5	6.46	44.7	19.3	25.8
dudek	13.3	<i>11.9</i>	13.5	16.7	10.7	23.8	51.6	23.5	18.7	85.5	14.1	17.7	25.7
fleetface	33.6	29.2	38.2	<i>24.6</i>	20.5	36	67.8	114	68	33.5	25.6	25	60.9
freeman1	7.78	11.8	105	125	13.4	17.3	112	62.4	<i>10.4</i>	86.1	125	10.9	115
freeman3	16.6	15.3	<i>3.55</i>	2.94	11.6	17.3	5.95	33.1	32.6	40.5	53.7	55.8	22.8
freeman4	20	46.4	83.7	11.4	46.1	79.3	83.5	22.1	57.5	38.6	79	<i>11.6</i>	93.8
girl	11.9	8.26	3.52	24.7	2.74	8.32	18	<i>2.8</i>	24	22.9	19.4	41.9	19.2
ironman	79.9	178	162	172	100	93	231	163	240	<i>89.3</i>	185	235	175
lemming	82.6	154	182	189	39.7	16.8	80.8	178	77.8	<i>20.2</i>	114	82.4	134
liquor	98.5	148	42.3	110	133	<i>40.3</i>	215	213	221	8.71	155	41	179
matrix	51.8	59.5	61.4	<i>53.7</i>	216	57.2	415	61.7	106	83.3	114	66.5	69
motorRolling	323	215	206	172	143	87.8	183	207	174	<i>132</i>	393	194	166
shaking	<i>8.17</i>	6.67	23.1	22.3	21.9	68.5	104	110	26.3	75	16.7	15.7	126
singer1	3.51	6.22	<i>3.03</i>	2.94	12.4	10.6	16.6	53.4	18.8	148	15.4	21	15
skating1	8.51	7.47	63.7	49.2	82.3	104	199	159	174	111	<i>7.78</i>	82.3	184
skiing	<i>233</i>	242	251	222	261	264	283	266	276	245	248	260	277
soccer	<i>15</i>	13.6	81	73.8	73.5	69.4	243	101	140	57.1	70.9	82.9	69.1
trellis	2.67	<i>2.83</i>	7.8	6.84	15.3	55.9	59.6	62.2	44.9	47.9	18.9	61.2	51.1
walking	1.7	3.28	<i>1.71</i>	2.37	5.36	110	5.77	3.31	5.87	2.38	7.6	6.59	214
walking2	<i>3.18</i>	15.7	37.3	1.58	12.9	63.1	28.7	5.06	29.1	64.8	18.5	50.6	64.6
woman	9.82	10.3	<i>8.27</i>	8.71	3.44	78.9	16.7	129	8.5	122	208	8.8	116
Median	10.9	<i>11.8</i>	30.2	23.5	14.3	48.1	78.6	62.3	58.5	60.9	35.1	31.2	71.5

Table 3: The per-video center location error (CLE) (in pixels) for our fast scale tracker, our exhaustive scale tracker and the compared state-of-the-art trackers.

References

- [1] C. Bao, Y. Wu, H. Ling, and H. Ji. Real time robust l1 tracker using accelerated proximal gradient approach. In *CVPR*, 2012.
- [2] M. Felsberg. Enhanced distribution field tracking using channel representations. In *ICCV Workshop*, 2013.
- [3] S. Hare, A. Saffari, and P. Torr. Struck: Structured output tracking with kernels. In *ICCV*, 2011.
- [4] S. He, Q. Yang, R. Lau, J. Wang, and M.-H. Yang. Visual tracking via locality sensitive histograms. In *CVPR*, 2013.
- [5] J. Henriques, R. Caseiro, P. Martins, and J. Batista. Exploiting the circulant structure of tracking-by-detection with kernels. In *ECCV*, 2012.
- [6] X. Jia, H. Lu, and M.-H. Yang. Visual tracking via adaptive structural local sparse appearance model. In *CVPR*, 2012.
- [7] Z. Kalal, J. Matas, and K. Mikolajczyk. P-n learning: Bootstrapping binary classifiers by structural constraints. In *CVPR*, 2010.
- [8] S. Oron, A. Bar-Hillel, D. Levi, and S. Avidan. Locally orderless tracking. In *CVPR*, 2012.
- [9] L. Sevilla-Lara and E. G. Learned-Miller. Distribution fields for tracking. In *CVPR*, 2012.
- [10] K. Zhang, L. Zhang, and M. Yang. Real-time compressive tracking. In *ECCV*, 2012.
- [11] W. Zhong, H. Lu, and M.-H. Yang. Robust object tracking via sparsity-based collaborative model. In *CVPR*, 2012.