

Long-term Correlation Tracking

Supplementary Document

Chao Ma^{1,2}, Xiaokang Yang¹, Chongyang Zhang¹, and Ming-Hsuan Yang²

¹Shanghai Jiao Tong University ²University of California at Merced

{chaoma, xkyang, sunny_zhang}@sjtu.edu.cn, mhyang@ucmerced.edu

Quantitative Evaluation. We quantitatively evaluate the proposed long-term correlation tracking (LCT) algorithm on the 50 benchmark sequences with comparisons to the 11 state-of-the-art trackers, CSK [4], STC [10], KCF [5] MIL [1], Struck [3], CT [11], ASLA [6], TLD [7], SCM [12], MEEM [9], and TGPR [2]). We report the distance precision at a threshold of 20 pixels in Table 1 and the overlap success rate at a threshold of 0.5 in Table 2. We report the distance precision plots over eight tracking challenges in our attribute-based evaluation in Figure 1 as mentioned on line 643 in the manuscript.

Additional Qualitative Evaluation. We present the tracking results of our algorithm with comparison to other five state-of-the-art trackers (KCF [5], MEEM [9], STC [10], Struck [3], and TLD [7]) on 20 more benchmark video sequences [8] in Figure 2-5. Note that a case of tracking failure is denoted by $\times$ in the upper right of the image frame in the manuscript and this supplementary material. Due to the limitation of 100 MB size, these sequences are not included in the video of this supplementary material. In addition, we compare the center location errors frame-by-frame on these 20 sequences in Figure 6.

References

- [1] B. Babenko, M.-H. Yang, and S. Belongie. Robust object tracking with online multiple instance learning. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 33(8), 2011.
- [2] J. Gao, H. Ling, W. Hu, and J. Xing. Transfer learning based visual tracking with gaussian processes regression. In *Proceedings of the European Conference on Computer Vision*, 2014.
- [3] S. Hare, A. Saffari, and P. H. S. Torr. Struck: Structured output tracking with kernels. In *Proceedings of IEEE International Conference on Computer Vision*, 2011.
- [4] J. F. Henriques, R. Caseiro, P. Martins, and J. Batista. Exploiting the circulant structure of tracking-by-detection with kernels. In *Proceedings of the European Conference on Computer Vision*, 2012.
- [5] J. F. Henriques, R. Caseiro, P. Martins, and J. Batista. High-speed tracking with kernelized correlation filters. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2015.
- [6] X. Jia, H. Lu, and M.-H. Yang. Visual tracking via adaptive structural local sparse appearance model. In *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, 2012.
- [7] Z. Kalal, K. Mikolajczyk, and J. Matas. Tracking-learning-detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 34(7):1409–1422, 2012.
- [8] Y. Wu, J. Lim, and M.-H. Yang. Online object tracking: A benchmark. In *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, 2013.
- [9] J. Zhang, S. Ma, and S. Sclaroff. MEEM: robust tracking via multiple experts using entropy minimization. In *Proceedings of the European Conference on Computer Vision*, 2014.
- [10] K. Zhang, L. Zhang, Q. Liu, D. Zhang, and M.-H. Yang. Fast visual tracking via dense spatio-temporal context learning. In *Proceedings of the European Conference on Computer Vision*, 2014.
- [11] K. Zhang, L. Zhang, and M.-H. Yang. Real-time compressive tracking. In *Proceedings of the European Conference on Computer Vision*, 2012.
- [12] W. Zhong, H. Lu, and M.-H. Yang. Robust object tracking via sparse collaborative appearance model. *IEEE Transactions on Image Processing*, 23(5):2356–2368, 2014.

Table 1. Comparisons with state-of-the-art trackers on the 50 benchmark sequences in the distance precision at a threshold of 20 pixels. Our LCT algorithm performs favorably against the state-of-the-art algorithms.

	LCT	CSK	STC	KCF	MIL	Struck	CT	ASLA	TLD	SCM	MEEM	TGPR
<i>basketball</i>	1.000	1.000	0.560	0.923	0.284	0.120	0.299	0.599	0.028	0.661	0.250	0.993
<i>bolt</i>	1.000	0.034	0.046	0.989	0.014	0.020	0.011	0.017	0.306	0.031	0.054	0.026
<i>boy</i>	1.000	0.844	0.761	1.000	0.846	1.000	0.930	0.440	1.000	0.440	1.000	1.000
<i>car4</i>	0.989	0.355	0.967	0.953	0.354	0.992	0.281	1.000	0.874	0.974	0.686	1.000
<i>carDark</i>	1.000	1.000	1.000	1.000	0.379	1.000	0.005	1.000	0.639	1.000	1.000	1.000
<i>carScale</i>	0.730	0.651	0.647	0.806	0.627	0.647	0.718	0.742	0.853	0.647	0.813	0.806
<i>coke</i>	0.914	0.873	0.155	0.838	0.151	0.948	0.113	0.165	0.684	0.430	0.945	0.942
<i>couple</i>	0.571	0.086	0.086	0.257	0.679	0.736	0.693	0.086	1.000	0.114	1.000	0.107
<i>crossing</i>	1.000	1.000	0.533	1.000	1.000	1.000	1.000	1.000	0.617	1.000	1.000	0.950
<i>david</i>	1.000	0.499	0.837	1.000	0.699	0.329	0.815	1.000	1.000	1.000	0.979	0.987
<i>david2</i>	1.000	1.000	1.000	1.000	0.978	1.000	0.004	1.000	1.000	1.000	1.000	1.000
<i>david3</i>	1.000	0.659	0.925	1.000	0.738	0.337	0.413	0.548	0.111	0.496	0.988	1.000
<i>deer</i>	0.817	1.000	0.042	0.817	0.127	1.000	0.042	0.028	0.732	0.028	1.000	1.000
<i>dog1</i>	1.000	1.000	0.700	1.000	0.919	0.996	0.950	0.997	1.000	0.976	0.979	1.000
<i>doll</i>	0.981	0.579	0.763	0.968	0.732	0.919	0.684	0.923	0.983	0.978	0.978	0.971
<i>dudek</i>	0.907	0.807	0.554	0.877	0.688	0.897	0.418	0.755	0.597	0.883	0.788	0.681
<i>faceocc1</i>	0.915	0.947	0.250	0.730	0.221	0.575	0.330	0.180	0.203	0.933	0.665	0.831
<i>faceocc2</i>	1.000	1.000	0.974	0.972	0.740	1.000	0.681	0.792	0.856	0.860	0.975	0.979
<i>fish</i>	1.000	0.042	1.000	1.000	0.387	1.000	0.882	1.000	1.000	0.863	1.000	1.000
<i>fleeface</i>	0.546	0.567	0.481	0.460	0.358	0.639	0.438	0.301	0.506	0.529	0.574	0.393
<i>football</i>	1.000	0.798	0.801	0.796	0.790	0.751	0.798	0.735	0.804	0.765	0.992	1.000
<i>football1</i>	0.973	0.757	0.514	0.959	1.000	1.000	0.351	0.797	0.554	0.568	1.000	0.986
<i>freeman1</i>	0.975	0.555	0.371	0.402	0.939	0.801	0.396	0.390	0.540	0.982	0.923	0.985
<i>freeman3</i>	0.793	0.572	0.596	0.911	0.048	0.789	0.209	1.000	0.767	1.000	0.965	0.122
<i>freeman4</i>	0.951	0.187	0.233	0.534	0.201	0.375	0.064	0.219	0.410	0.509	0.565	0.519
<i>girl</i>	1.000	0.554	0.594	0.864	0.714	1.000	0.608	1.000	0.918	1.000	1.000	0.904
<i>ironman</i>	0.145	0.133	0.151	0.217	0.108	0.114	0.096	0.133	0.120	0.157	0.036	0.096
<i>jogging-1</i>	0.971	0.228	0.228	0.235	0.231	0.241	0.231	0.231	0.974	0.228	0.456	0.225
<i>jogging-2</i>	0.977	0.186	0.186	0.163	0.186	0.254	0.166	0.182	0.857	1.000	0.961	0.997
<i>jumping</i>	0.968	0.051	0.054	0.342	0.997	1.000	0.096	0.450	1.000	0.153	1.000	0.109
<i>lemming</i>	0.778	0.436	0.312	0.487	0.823	0.628	0.677	0.168	0.859	0.166	0.731	0.275
<i>liquor</i>	0.824	0.223	0.403	0.976	0.199	0.390	0.209	0.226	0.588	0.276	0.565	0.657
<i>matrix</i>	0.360	0.010	0.100	0.170	0.180	0.120	0.020	0.050	0.160	0.350	0.090	0.110
<i>mhyang</i>	1.000	1.000	1.000	1.000	0.460	1.000	0.819	1.000	0.978	1.000	1.000	1.000
<i>motorRolling</i>	0.043	0.043	0.073	0.049	0.043	0.085	0.037	0.061	0.116	0.037	0.043	0.061
<i>mountainBike</i>	0.996	1.000	1.000	1.000	0.667	0.921	0.175	0.904	0.259	0.969	0.934	1.000
<i>shaking</i>	0.981	0.564	0.981	0.019	0.282	0.192	0.047	0.485	0.405	0.814	0.970	0.641
<i>singer1</i>	1.000	0.670	1.000	0.843	0.501	0.641	0.840	1.000	1.000	1.000	0.419	0.219
<i>singer2</i>	0.975	0.036	0.571	0.948	0.404	0.036	0.005	0.036	0.071	0.112	0.033	0.954
<i>skating1</i>	1.000	0.988	0.690	1.000	0.130	0.465	0.090	0.765	0.318	0.768	0.445	0.700
<i>skiing</i>	0.136	0.099	0.136	0.074	0.074	0.037	0.086	0.136	0.123	0.136	0.136	0.111
<i>soccer</i>	0.885	0.135	0.135	0.791	0.191	0.253	0.219	0.122	0.115	0.268	0.349	0.143
<i>subway</i>	1.000	0.240	0.246	1.000	0.994	0.983	0.989	0.229	0.251	1.000	1.000	1.000
<i>suv</i>	0.980	0.568	0.805	0.979	0.123	0.572	0.250	0.575	0.909	0.978	0.958	0.531
<i>sylvester</i>	0.993	0.910	0.897	0.843	0.651	0.995	0.901	0.821	0.949	0.946	0.967	0.946
<i>tiger1</i>	0.868	0.255	0.261	0.851	0.095	0.175	0.215	0.226	0.456	0.126	0.831	0.269
<i>tiger2</i>	0.696	0.110	0.145	0.356	0.414	0.630	0.364	0.142	0.386	0.112	0.490	0.792
<i>trellis</i>	1.000	0.810	0.738	1.000	0.230	0.877	0.387	0.861	0.529	0.873	0.956	0.981
<i>walking</i>	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.964	1.000	0.998	1.000
<i>walking2</i>	0.404	0.468	0.794	0.434	0.406	0.982	0.432	0.404	0.426	1.000	0.432	0.996
<i>woman</i>	0.940	0.250	0.615	0.938	0.206	1.000	0.204	0.203	0.191	0.940	1.000	0.940
<i>Average</i>	0.854	0.545	0.547	0.741	0.475	0.656	0.406	0.532	0.608	0.649	0.744	0.705

Table 2. Comparisons with state-of-the-art trackers on the 50 benchmark sequences in the overlap success rate at a threshold of 0.5. Our LCT algorithm performs well compared to the state-of-the-art algorithms.

	LCT	CSK	STC	KCF	MIL	Struck	CT	ASLA	TLD	SCM	MEEM	TGPR
<i>basketball</i>	0.906	0.874	0.236	0.898	0.274	0.102	0.259	0.556	0.025	0.611	0.236	0.850
<i>bolt</i>	0.720	0.017	0.043	0.934	0.011	0.017	0.006	0.014	0.146	0.014	0.037	0.014
<i>boy</i>	0.998	0.842	0.663	0.992	0.385	0.975	0.688	0.439	0.935	0.439	0.987	0.990
<i>car4</i>	1.000	0.276	0.225	0.367	0.276	0.398	0.275	1.000	0.792	0.973	0.349	0.398
<i>carDark</i>	0.885	0.992	0.997	0.723	0.178	1.000	0.003	1.000	0.529	0.997	1.000	1.000
<i>carScale</i>	0.702	0.448	0.528	0.444	0.448	0.433	0.448	0.690	0.437	0.651	0.373	0.421
<i>coke</i>	0.914	0.739	0.089	0.722	0.117	0.942	0.093	0.148	0.289	0.337	0.942	0.914
<i>couple</i>	0.507	0.086	0.086	0.243	0.671	0.543	0.686	0.086	1.000	0.107	0.750	0.107
<i>crossing</i>	1.000	0.317	0.175	0.925	0.983	0.942	0.983	1.000	0.517	1.000	0.983	0.808
<i>david</i>	0.907	0.236	0.584	0.622	0.229	0.236	0.427	0.958	0.970	0.913	0.588	0.771
<i>david2</i>	0.972	1.000	0.752	1.000	0.324	1.000	0.002	0.946	0.952	0.911	1.000	1.000
<i>david3</i>	0.790	0.627	0.333	0.992	0.683	0.337	0.349	0.508	0.103	0.484	0.853	0.988
<i>deer</i>	0.817	1.000	0.042	0.817	0.127	1.000	0.042	0.028	0.732	0.028	1.000	1.000
<i>dog1</i>	0.929	0.653	0.573	0.653	0.650	0.653	0.652	0.921	0.673	0.847	0.636	0.713
<i>doll</i>	0.968	0.218	0.100	0.552	0.433	0.688	0.531	0.922	0.624	0.987	0.726	0.714
<i>dudek</i>	0.987	0.947	0.724	0.976	0.857	0.980	0.852	0.897	0.842	0.976	0.985	0.879
<i>faceocc1</i>	0.901	1.000	0.243	1.000	0.765	1.000	0.854	0.311	0.834	1.000	1.000	0.982
<i>faceocc2</i>	0.649	1.000	0.980	0.996	0.936	1.000	0.744	0.809	0.829	0.874	0.998	0.993
<i>fish</i>	1.000	0.042	0.372	1.000	0.387	1.000	0.889	1.000	0.964	0.863	1.000	1.000
<i>fleeftace</i>	0.687	0.676	0.463	0.669	0.537	0.666	0.638	0.608	0.567	0.706	0.795	0.610
<i>football</i>	0.989	0.657	0.619	0.682	0.738	0.660	0.785	0.652	0.412	0.586	0.959	0.970
<i>football1</i>	0.770	0.392	0.351	0.959	0.784	0.878	0.081	0.419	0.392	0.405	0.919	0.811
<i>freeman1</i>	0.623	0.144	0.169	0.160	0.153	0.215	0.101	0.307	0.212	0.807	0.224	0.218
<i>freeman3</i>	0.578	0.330	0.207	0.274	0.009	0.200	0.002	0.937	0.580	0.930	0.309	0.067
<i>freeman4</i>	0.693	0.170	0.170	0.184	0.021	0.155	0.004	0.166	0.269	0.244	0.336	0.332
<i>girl</i>	0.912	0.398	0.302	0.756	0.294	0.980	0.178	0.906	0.764	0.882	0.938	0.882
<i>ironman</i>	0.114	0.127	0.042	0.157	0.048	0.048	0.090	0.127	0.066	0.133	0.018	0.060
<i>jogging-1</i>	0.967	0.225	0.208	0.225	0.225	0.225	0.225	0.225	0.967	0.212	0.397	0.225
<i>jogging-2</i>	0.941	0.182	0.173	0.160	0.163	0.248	0.140	0.182	0.831	0.990	0.948	0.990
<i>jumping</i>	0.936	0.048	0.048	0.281	0.476	0.799	0.006	0.166	0.847	0.121	0.987	0.096
<i>lemming</i>	0.778	0.429	0.153	0.431	0.811	0.640	0.680	0.168	0.594	0.166	0.701	0.268
<i>liquor</i>	0.841	0.278	0.251	0.982	0.201	0.406	0.209	0.239	0.582	0.321	0.609	0.681
<i>matrix</i>	0.250	0.010	0.100	0.130	0.110	0.120	0.020	0.020	0.070	0.300	0.090	0.040
<i>mhyang</i>	0.926	1.000	0.860	1.000	0.389	1.000	0.730	1.000	0.893	0.997	1.000	0.982
<i>motorRolling</i>	0.024	0.073	0.140	0.079	0.073	0.159	0.055	0.098	0.171	0.073	0.079	0.110
<i>mountainBike</i>	0.965	1.000	0.873	0.991	0.575	0.855	0.171	0.899	0.259	0.961	0.816	1.000
<i>shaking</i>	0.951	0.581	0.852	0.014	0.227	0.167	0.041	0.378	0.400	0.896	0.932	0.433
<i>singer1</i>	0.681	0.296	0.507	0.276	0.276	0.299	0.248	1.000	0.991	1.000	0.274	0.228
<i>singer2</i>	0.527	0.036	0.459	0.970	0.475	0.036	0.011	0.036	0.101	0.164	0.036	1.000
<i>skating1</i>	0.890	0.368	0.230	0.363	0.103	0.370	0.100	0.690	0.228	0.423	0.315	0.535
<i>skiing</i>	0.099	0.074	0.111	0.062	0.074	0.037	0.074	0.111	0.074	0.086	0.111	0.086
<i>soccer</i>	0.648	0.138	0.115	0.390	0.156	0.156	0.202	0.125	0.122	0.237	0.319	0.130
<i>subway</i>	1.000	0.223	0.223	0.994	0.794	0.909	0.766	0.217	0.229	0.994	0.674	0.994
<i>suv</i>	0.980	0.575	0.513	0.985	0.130	0.575	0.231	0.580	0.839	0.984	0.960	0.535
<i>sylvester</i>	0.855	0.717	0.610	0.819	0.546	0.929	0.828	0.749	0.928	0.886	0.926	0.923
<i>tiger1</i>	0.917	0.264	0.052	0.857	0.097	0.183	0.246	0.244	0.456	0.129	0.865	0.272
<i>tiger2</i>	0.619	0.107	0.090	0.364	0.447	0.649	0.370	0.142	0.173	0.112	0.488	0.816
<i>trellis</i>	0.937	0.591	0.580	0.840	0.244	0.784	0.350	0.858	0.473	0.854	0.819	0.793
<i>walking</i>	0.981	0.519	0.721	0.515	0.541	0.566	0.502	0.998	0.383	0.959	0.519	0.738
<i>walking2</i>	0.406	0.388	0.442	0.378	0.380	0.434	0.384	0.398	0.340	1.000	0.350	0.706
<i>woman</i>	0.687	0.245	0.258	0.936	0.188	0.935	0.159	0.194	0.166	0.858	0.930	0.935
<i>Average</i>	0.769	0.443	0.365	0.622	0.373	0.559	0.341	0.511	0.521	0.616	0.649	0.628

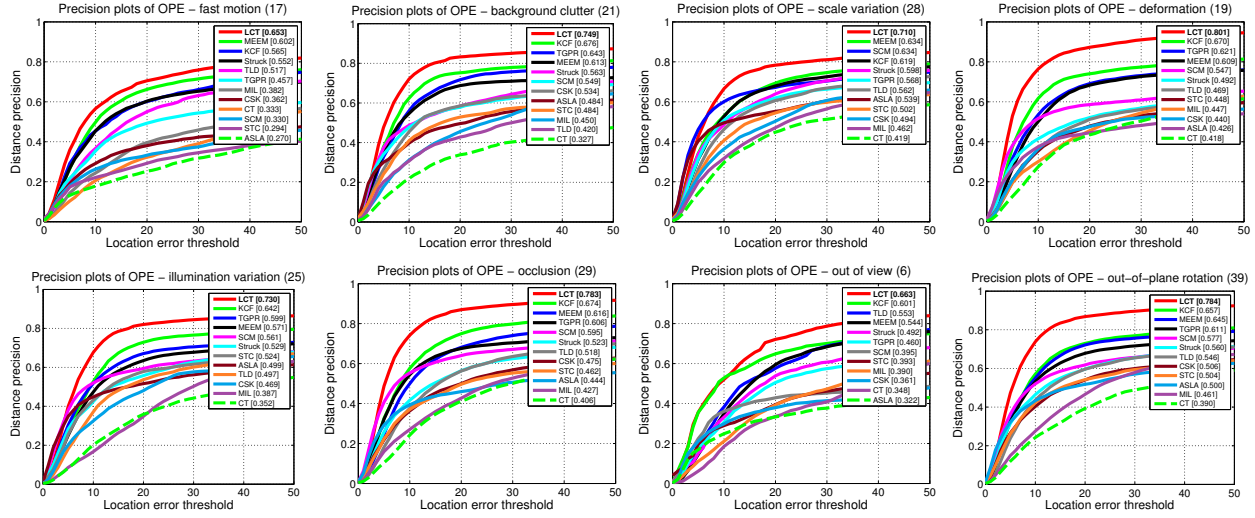


Figure 1. Distance precision plots over eight tracking challenges of fast motion, background clutter, scale variation, deformation, illumination variation, occlusion, out-of-view, and out-of-plane rotation. The legend contains the AUC score for each tracker. The proposed LCT method performs well when evaluating with eight challenging factors.



Figure 2. Tracking results of our proposed algorithm, KCF, STC, and TLD methods on six challenging sequences (from top to down are *Basketball*, *Bolt*, *CarDark*, *Crossing*, and *David2*, respectively). A case of tracking failure is denoted by $\times$.

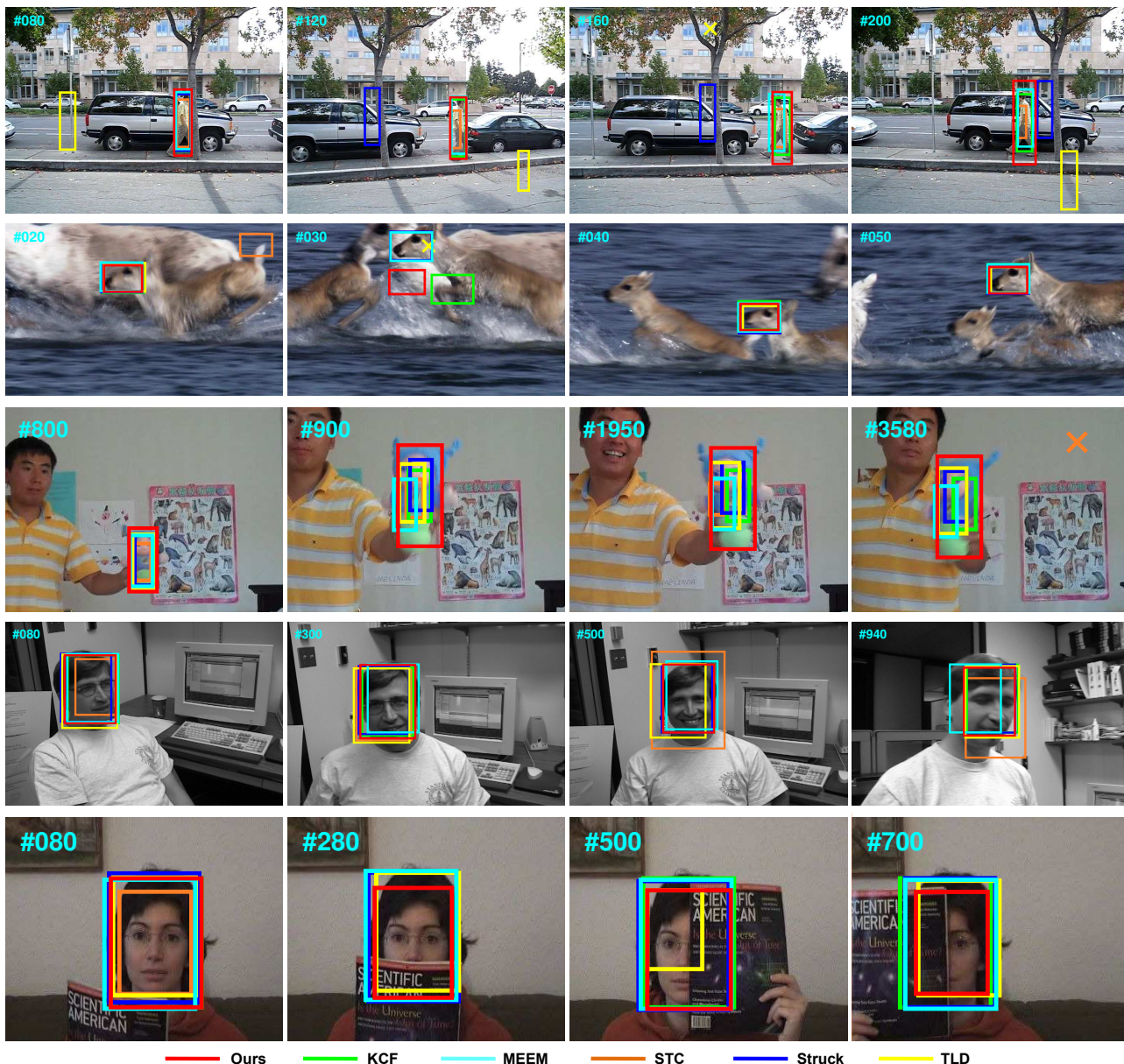


Figure 3. Tracking results of our proposed algorithm, KCF, STC, and TLD methods on six challenging sequences (from top to down are *David3*, *Deer*, *Doll*, *Dudek*, and *FaceOcc1*, respectively). A case of tracking failure is denoted by $\times$.

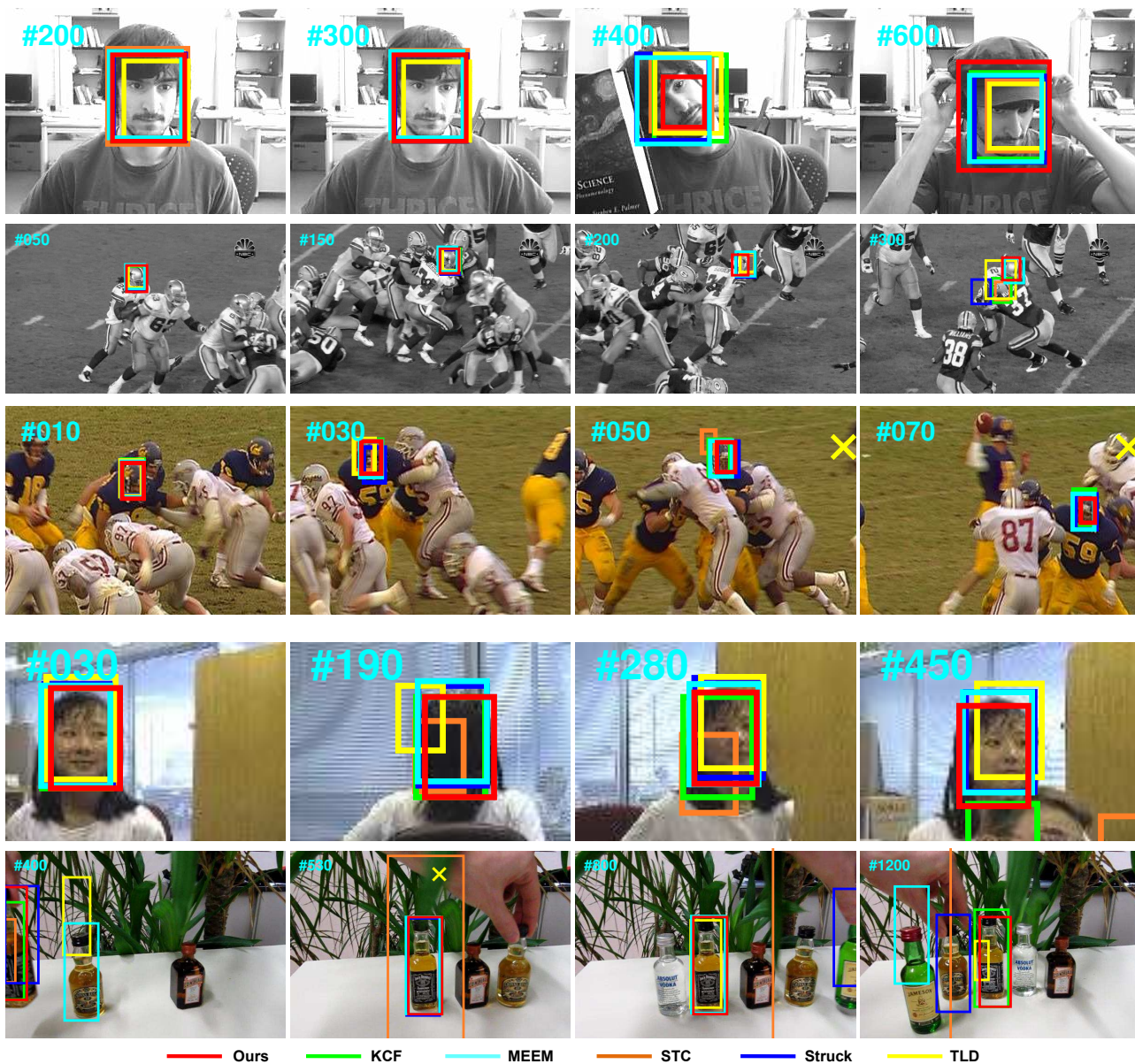


Figure 4. Tracking results of our proposed algorithm, KCF, STC, and TLD methods on six challenging sequences (from top to down are *FaceOcc2*, *Football*, *Football1*, *Girl*, and *Liquor*, respectively). A case of tracking failure is denoted by $\times$.

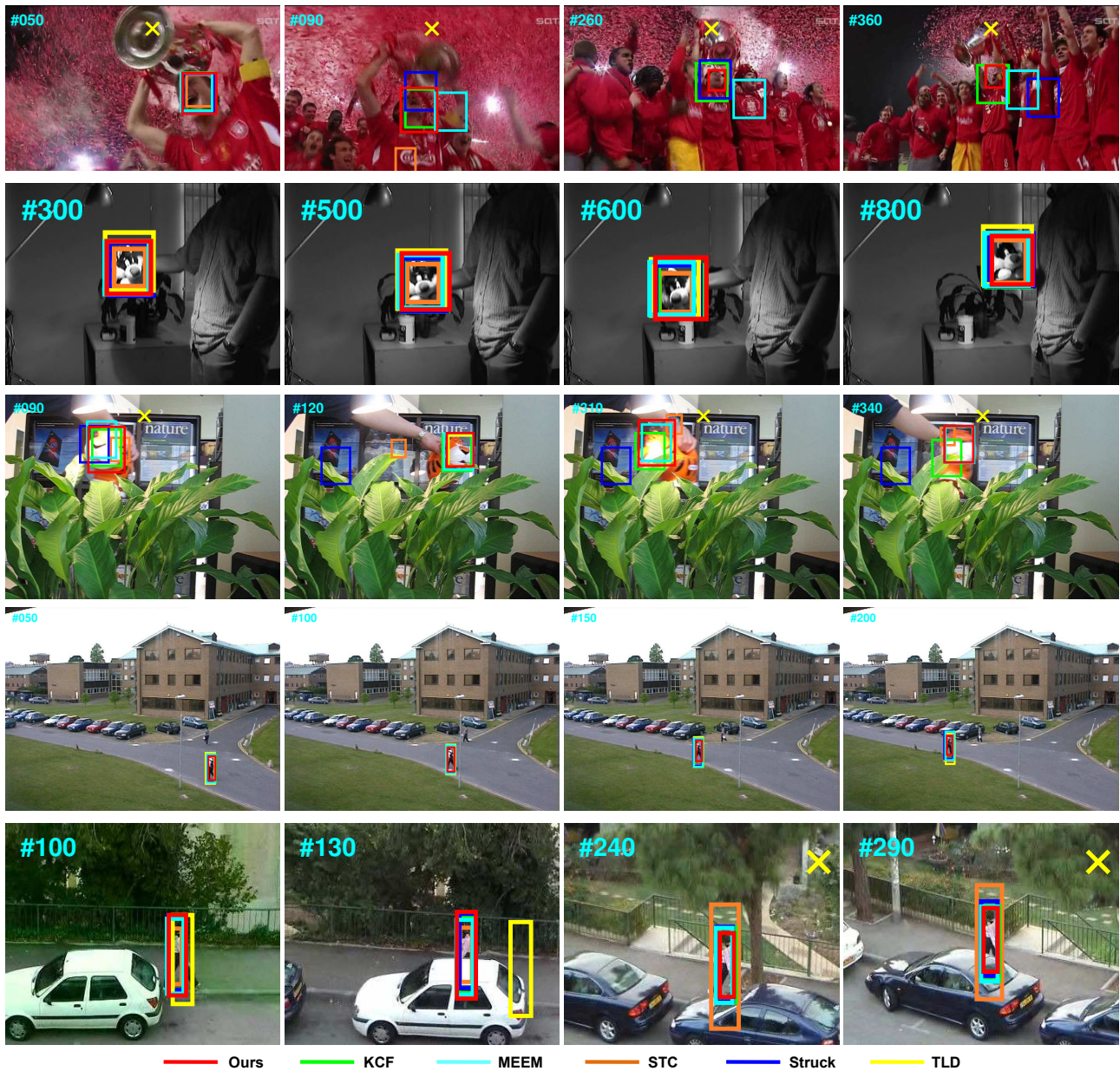


Figure 5. Tracking results of our proposed algorithm, KCF, STC, and TLD methods on six challenging sequences (from top to down are *Soccer*, *Sylvester*, *Tiger1*, *Walking*, and *Woman*, respectively). A case of tracking failure is denoted by $\times$.

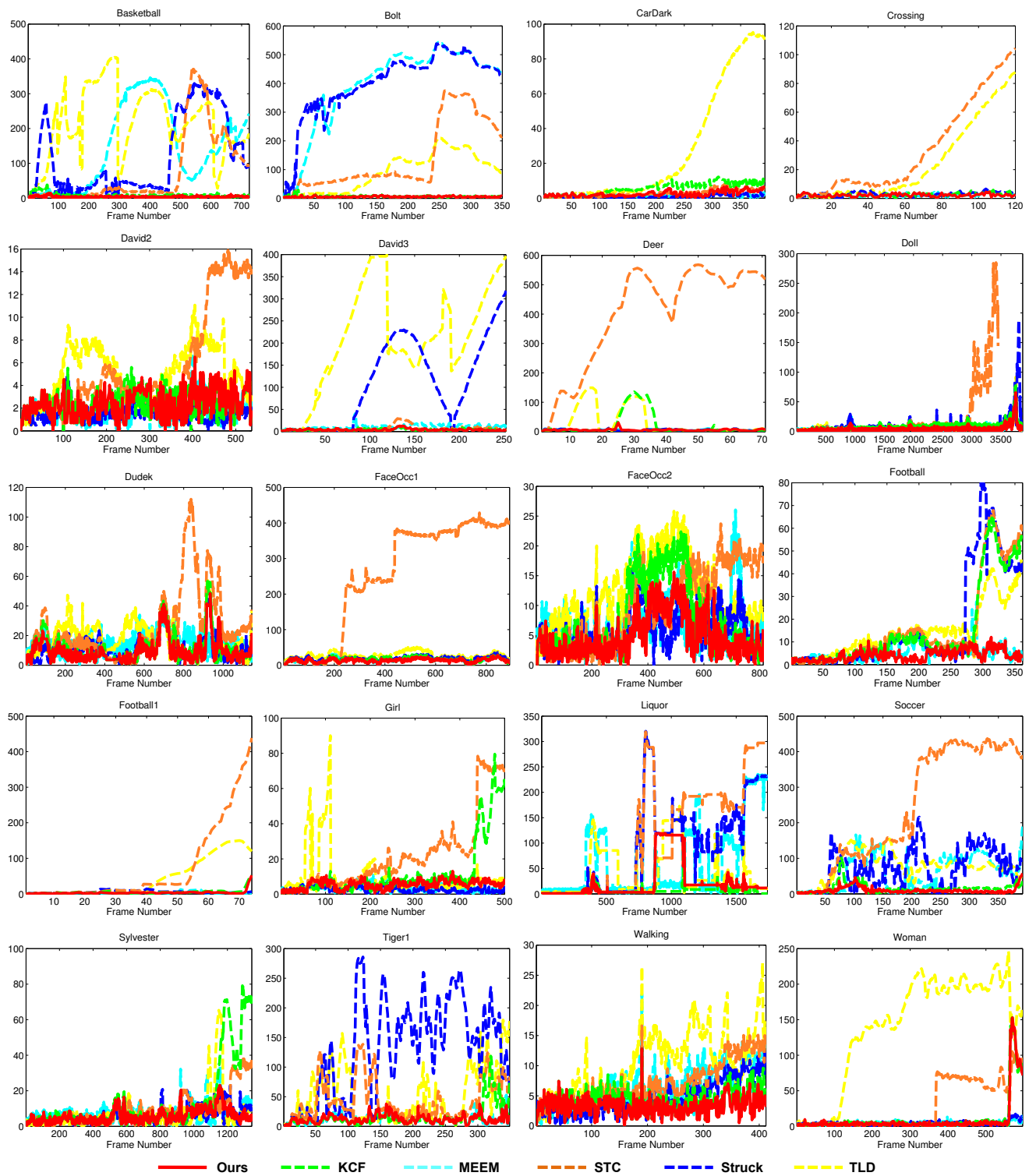


Figure 6. Frame-by-frame comparisons of the center location errors (in pixels) on twenty challenging sequences in Figure 2-5. Our method performs favorably against the state-of-the-art trackers.