

Zhao-yun Chen, Lei Luo, Da-fei Huang, Mei Wen, Chun-yuan Zhang, 2017.
Exploiting a depth context model in visual tracking with correlation filter. *Frontiers of Information Technology & Electronic Engineering*, **18**(5): 667-679.
<http://dx.doi.org/10.1631/FITEE.1500389>

Exploiting a depth context model in visual tracking with correlation filter

Key words: Visual tracking; Depth context model; Correlation filter;
Region growing

Corresponding author: Lei Luo

E-mail: l.luo@nudt.edu.cn

 ORCID: <http://orcid.org/0000-0002-9329-1411>

Motivation

- Visual object tracking is a fundamental task for a wide range of computer vision applications.
- Correlation filter based tracker have been proven to be competitive against other in accuracy and with much higher computational efficiency.
- Although scale-adaptive variants have been proposed, there is no correlation filter variant of flexible adaptability to target's scale and aspect ratio changes.
- The correlation filter based trackers are prone to drifting if the occluders have a similar appearance to the target.
- A reliable depth map can provide some valuable information which can significantly improve the tracking performance against occlusion and scale variation.

Main idea

- A depth context model is constructed for estimation of the target location which is resistant to heavy occlusion, fast motion and large deformation.
- A region growing method is adopted to enable the adaptability to the target's scale and aspect ratio changes.
- A scheme for occlusion detection and learning rate suppression is proposed to improve the performance when occlusion exists.

Workflow

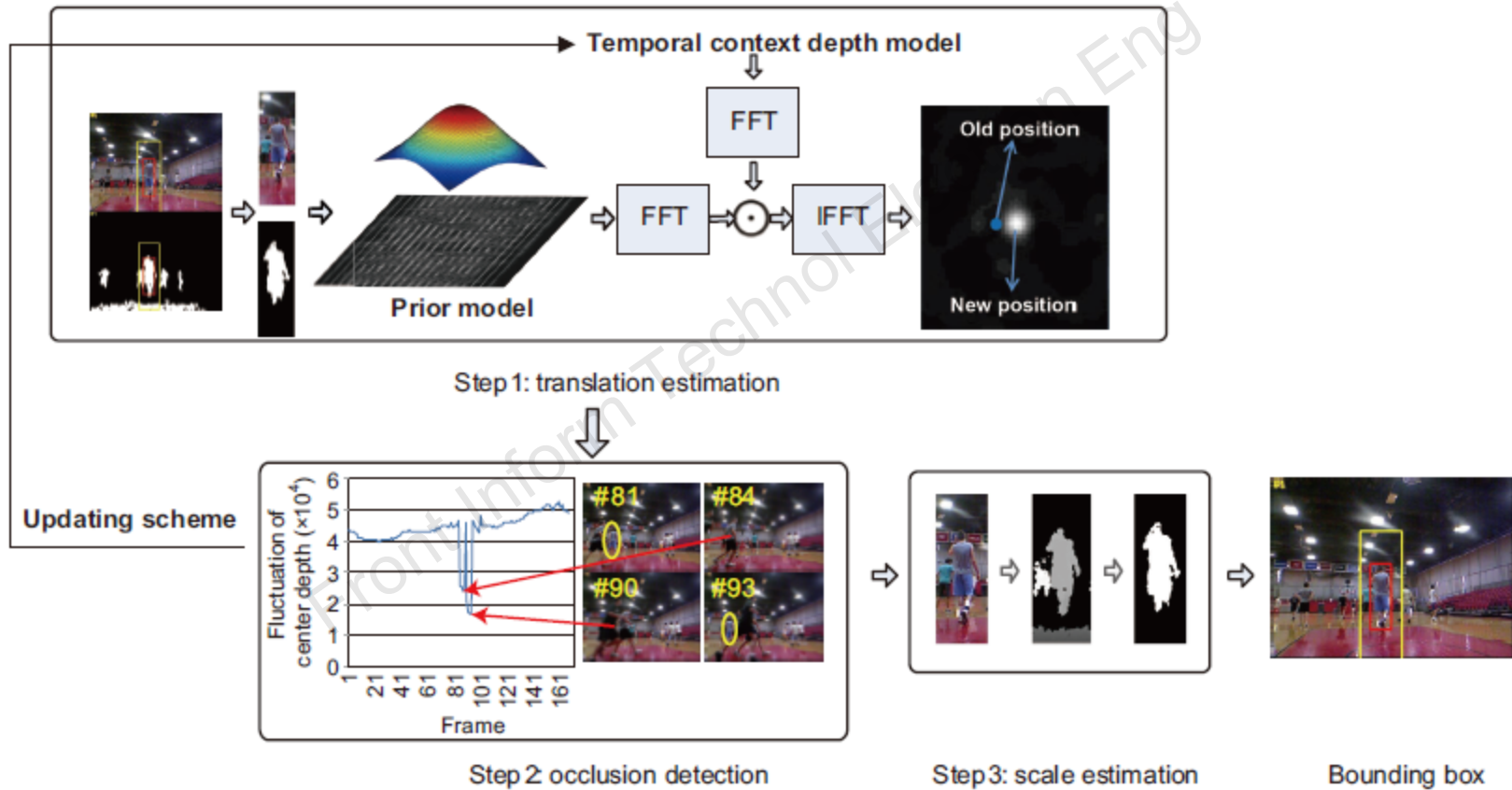


Fig. 1 Workflow of the proposed algorithm

Method

1. We construct the depth context model for robust estimation of the target location, which is based on STC (Zhang *et al.*, 2014).
2. A region-growing method (Adams and Bischof, 1994) based on the depth map is adopted to provide an accurate scale estimation for the target.
3. The occlusion detection is according to the depth of the target location and a corresponding updating scheme is adopted.

Major results

- Higher value indicates the better performance. **Red fonts** indicate the best performances. **Blue fonts** indicate the second best ones.

Table 2 Overall performances

Sequence number	Success rate						Distance precision					
	KCF	DSST	TLD	Struck	STC	Ours	KCF	DSST	TLD	Struck	STC	Ours
1	0.389	0.493	0.115	0.885	0.089	0.751	0.343	0.441	0.231	0.745	0.157	0.732
2	0.261	0.205	0.205	0.205	0.215	0.574	0.185	0.200	0.207	0.265	0.309	0.590
3	0.727	0.773	0.435	0.687	0.354	0.788	0.636	0.803	0.469	0.612	0.297	0.818
4	1.000	0.997	0.116	1.000	0.305	1.000	0.360	0.945	0.135	1.000	0.379	1.000
5	0.827	1.000	1.000	0.827	0.835	0.976	0.770	1.000	1.000	0.879	0.921	1.000
6	0.526	0.617	0.658	0.296	0.133	0.704	0.643	0.658	0.631	0.350	0.154	0.663
7	0.945	0.500	0.939	0.506	0.494	0.967	0.990	0.525	0.954	0.537	0.513	0.994
8	0.374	0.186	0.014	0.060	0.110	0.900	0.207	0.019	0.125	0.217	0.097	0.705
9	0.362	0.991	0.216	0.310	0.647	0.707	0.094	0.991	0.267	0.423	0.687	0.664
10	0.712	1.000	1.000	0.781	1.000	0.928	0.329	1.000	1.000	0.964	1.000	1.000
11	0.766	0.846	0.940	0.169	0.453	0.960	0.861	0.896	0.867	0.158	0.516	0.886
12	0.699	0.456	0.641	0.544	0.485	0.505	0.553	0.476	0.657	0.578	0.579	0.646
13	0.617	0.877	0.773	0.565	0.857	0.571	0.643	0.935	0.795	0.421	0.864	0.805
14	0.930	0.698	0.532	0.804	0.544	0.893	0.791	0.443	0.601	0.503	0.531	0.821
15	0.981	1.000	0.474	0.805	1.000	1.000	1.000	1.000	0.506	0.937	0.989	1.000
16	0.993	1.000	0.437	0.927	0.530	0.995	0.903	0.967	0.439	0.907	0.536	1.000
17	0.325	0.282	0.479	0.620	0.081	0.641	0.338	0.282	0.493	0.505	0.154	0.526
18	0.314	0.771	0.610	0.559	0.678	0.797	0.119	0.975	0.567	0.607	0.796	0.847
19	0.055	0.894	0.070	0.191	0.166	0.754	0.0603	0.879	0.094	0.179	0.264	0.950
20	0.977	0.953	0.957	0.875	0.785	0.989	1.000	1.000	0.964	0.896	0.851	1.000
Average	0.638	0.696	0.537	0.562	0.471	0.846	0.543	0.677	0.560	0.580	0.510	0.843
FPS	140.811	12.159	8.610	6.150	28.700	14.600						

Higher value indicates better performance. **Red fonts** indicate the best performances. **Blue fonts** indicate the second best ones.
 FPS: number of frames per second. References to color refer to the online version of this table

Major results (Cont'd)

- The evaluations demonstrate that the proposed tracker performs favorably against state-of-the-art algorithms.

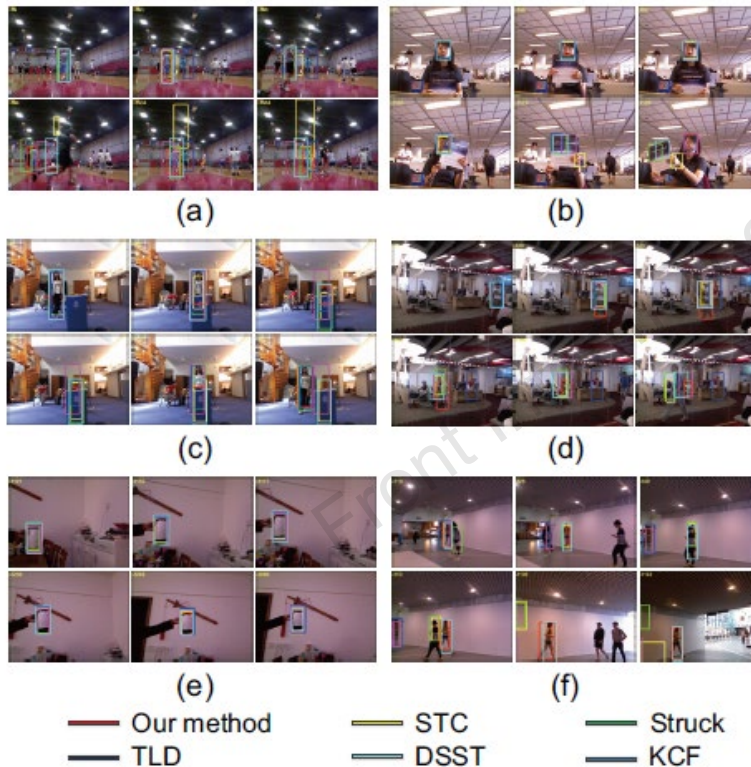


Fig. 6 Screenshots of tracking results:

- (a) Basketball1;
- (b) Face_occ5;
- (c) Tracking7.1;
- (d) Library2.2_occ;
- (e) Cup_move_1;
- (f) Walking_occ_long

Conclusions

- Our method uses a depth context model for estimating the translation of the target.
- A modified region growing method is also adopted for scale estimation.
- We additionally propose an accurate model updating scheme based on the occlusion detection using a depth map.
- Extensive experimental results show that the proposed algorithm performs favorably against outperforms the state-of-the-art trackers in terms of efficiency, accuracy and robustness.