Real-time RGB-D Tracking with Depth Scaling Kernelised Correlation Filters and Occlusion Handling

Massimo Camplani
massimo.camplani@bristol.ac.uk
Sion Hannuna
sh1670@bristol.ac.uk
Majid Mirmehdi
majid@cs.bris.ac.uk
Dima Damen
dima.damen@bristol.ac.uk
Adeline Paiement
csatmp@bristol.ac.uk
Lili Tao
lili.tao@bristol.ac.uk
Tilo Burghardt
tb2935@bristol.ac.uk

Visual Information Laboratory
Faculty of Engineering
University of Bristol
Bristol, UK

The recent surge in popularity of real-time RGB-D sensors has encouraged research into combining colour and depth data for tracking. The results from a few, recent works in RGB-D tracking have demonstrated that state-of-the-art RGB tracking algorithms can be outperformed by approaches that fuse colour and depth, for example [1, 3, 4, 5].

In this paper, we propose a real-time RGB-D tracker which we refer to as the Depth Scaling Kernalised Correlations Filters (DS-KCF). It is based on, and improves upon, the RGB Kernelised Correlation Filters tracker (KCF) from [2]. KCF is based on the use of the 'kernel trick' to extend correlation filters for very fast RGB tracking. The KCF tracker has important characteristics, in particular its ability to combine high accuracy and processing speed as demonstrated in [2, 6]. It is based on a simple processing chain that comprises training, detection, retraining and model update obtained by linear interpolation. The key to KCF is that it exploits the properties of circulant matrices to achieve efficient learning by implicitly encoding convolution and by allowing to operate in the Fourier domain using mainly element wise operations.

The proposed DS-KCF tracker[1] extends the RGB KCF tracker in three ways: (i) we employ an the efficient combination of colour and depth features (ii) we propose an efficient a novel management of scale changes and (iii) occlusions handling. The improvements we implement provide higher rates of accuracy while still operating at better than real-time frame rates (35fps on average). In particular, depth data in the target region is segmented with a fast K-means approach to extract relevant features for the target's depth distribution. Modelled as a Gaussian distribution, this data allows to identify scale changes and efficiently model them in the Fourier domain. The advantage of the proposed approach is that only a single target model is kept and updated. Furthermore, region depth distribution enables the detection of possible occlusions identified as sudden changes in the target region's depth histogram, and recovering lost tracks by searching for the unoccluded object in specifically identified key areas. During an occlusion, the model is not updated and the occluding object is tracked to guide the target's search space.

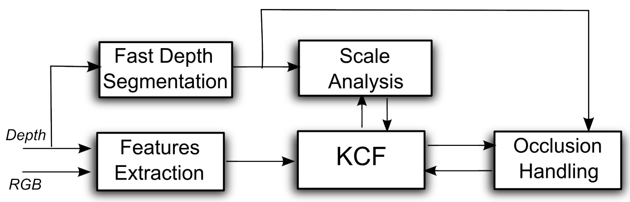

Figure 1: Block diagram of the proposed DS-KCF tracker.

Comparison on the Princeton Dataset [4] - We compare tracking performance by reporting the precision value for an error threshold equal to 20 pixels (P20), the area under the curve (AUC) of success plot measure, and the number of processed frames per second (fps). Table 1 shows that the proposed DS-KCF tracker outperforms the baseline KCF leading to better results both in terms of AUC and P20 measures. DS-KCF also outperforms the other two RGB-D trackers tested, Prin-Track [4] and OAPF [3]. Furthermore, the average processing rate in the Prin-Track (RGB-D) is 0.14fps and 0.9fps for the OAPF tracker in striking contrast to 40fps for DS-KCF. Example results of the trackers are shown in Figure 2.

Table 1: Trackers' performance on Princeton Validation Dataset [4].

	Average		
	AUC	**P20**	**fps**
KCF [2]	56.6	73.3	**103**
DS-KCF	**79.5**	94.2	40
Prin-Track(RGB-D) [4]	74.1	**96.0**	0.14
OAPF [3]	76.5	-	0.90

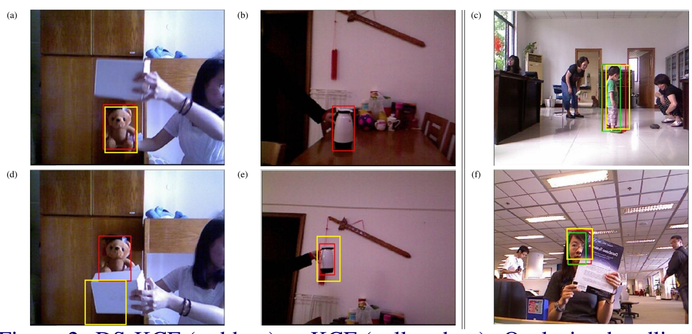

Figure 2: DS-KCF (red box) v. KCF (yellow box). Occlusion handling: (a) before occlusion, and (d) after occlusion. Change of scale: (b) initial target, and (e) target after change of scale. DS-KCF (red) v. Prin-Track(RGBD) (yellow) v. groundtruth (green) on (c) 'child_no1', and (f) 'face_occ5'.

[1] G. García, D. Klein, J. Stückler, S. Frintrop, and A. Cremers. Adaptive multi-cue 3D tracking of arbitrary objects. In Pattern Recognition, pages 357–366. 2012.

[2] J. F. Henriques, R. Caseiro, P. Martins, and J. Batista. High-speed tracking with kernelized correlation filters. Pattern Analysis and Machine Intelligence, IEEE Transactions on, 2015.

[3] T. Meshgi, S. Maeda, S. Oba, H. Skibbe, Y. Li, and S. Ishii. Occlusion aware particle filter tracker to handle complex and persistent occlusions. Computer Vision and Image Understanding, 2015 to appear.

[4] S. Song and J. Xiao. Tracking revisited using RGBD camera: Unified benchmark and baselines. In ICCV, pages 233–240, 2013.

[5] Q. Wang, J. Fang, and Y. Yuan. Multi-cue based tracking. Neurocomputing, 131(0):227 – 236, 2014.

[6] Y. Wu, J. Lim, and M. Yang. Online Object Tracking: A Benchmark. In CVPR, pages 2411–2418, 2013.

[1] DS-KCF code is available at `http://www.irc-sphere.ac.uk/work-package-2/DS-KCF`