Multi-Template Scale-Adaptive Kernelized Correlation Filters

Adel Bibi, Bernard Ghanem

King Abdullah University of Science and Technology (KAUST)

Abstract

- We identify the major drawbacks of a very computationally efficient and state-of-art-tracker known as the Kernelized Correlation Filter (KCF), which uses a fixed scale and a heuristic update strategy of the filter taps.
- We show that we can use multiple templates for the filter tap update along with non-linear kernel and multi-dimensional features in a joint fixed-point optimization by exploiting the underlying circulant structure.
- Our proposed method addresses the fixed-scale drawback by computing the MAP scale estimate over multiple scales rather simply using ML scale estimate.

Motivation

Motivated by the superior performance of KCF using multiple templates, which is sometimes even more prominent than using non-linear kernels, we re-formulate the filter update rule to incorporate previous templates and multi dimensional features in non-linear kernel spaces in one framework that uses fixed-point optimization. These two ideas were claimed to be infeasible when they are applied jointly.

Results on VOT 2015

Tracker	Acc. Rank	Rob. Rank	Rank
KCF_Gauss_HOG	**3.23**	2.83	**3.03**
KCF_Lin_HOG	3.35	**2.79**	3.07
KCF_Lin_Gray_Multi	3.41	3.20	3.30
KCF_Gauss_Gray	3.79	3.00	3.39
KCF_Lin_Gray	3.99	4.04	4.02

Problem Formulation

Multiple Templates: We aim to solve the filter update problem by incorporating previous templates. The model below is for two templates but can easily be extended to any number of templates.

Formulation

$$\min_{\mathbf{w}_1,\mathbf{w}_2} \left\|\begin{pmatrix}\Phi_1\mathbf{w}_1-\mathbf{y}\\ \Phi_2\mathbf{w}_2-\mathbf{y}\end{pmatrix}\right\|_2^2+\lambda\left\|\begin{pmatrix}\mathbf{w}_1\\ \mathbf{w}_2\end{pmatrix}\right\|_2^2$$

$$\text{subject to:}\quad \mathbf{w}_1=\mathbf{w}_2.$$

$$\min_{\mathbf{w}_1} ||\Phi_1\mathbf{w}_1-\mathbf{y}||_2^2+\lambda||\mathbf{w}_1||_2^2+\mu||\mathbf{w}_1-\mathbf{w}_2^j||_2^2$$

$$\min_{\mathbf{w}_2} ||\Phi_2\mathbf{w}_2-\mathbf{y}||_2^2+\lambda||\mathbf{w}_2||_2^2+\mu||\mathbf{w}_2-\mathbf{w}_1^{j+1}||_2^2$$

Solution

$$\hat{\mathbf{a}}_1=\frac{\hat{\psi}}{\hat{\mathbf{k}}^{x_1x_1}+(\lambda+\mu)}$$

$$\hat{\psi}=\mathbf{F}(-(k\mathbf{I}+(\lambda k+\mu(k-1))(\Phi_1\Phi_1^T)^{-1})\tilde{\mathbf{b}}+\mathbf{y})$$

$$\tilde{\mathbf{b}}=\mathbf{F}(\hat{k}^{x_2x_1}\odot\hat{\mathbf{a}}_2^*)$$

$$(\Phi_1\Phi_1^T)^{-1}\tilde{\mathbf{b}}=\mathbf{F}((\hat{k}^{x_1x_1})^{-1}\odot\hat{k}^{x_2x_1}\odot\hat{\mathbf{a}}_2^*)$$

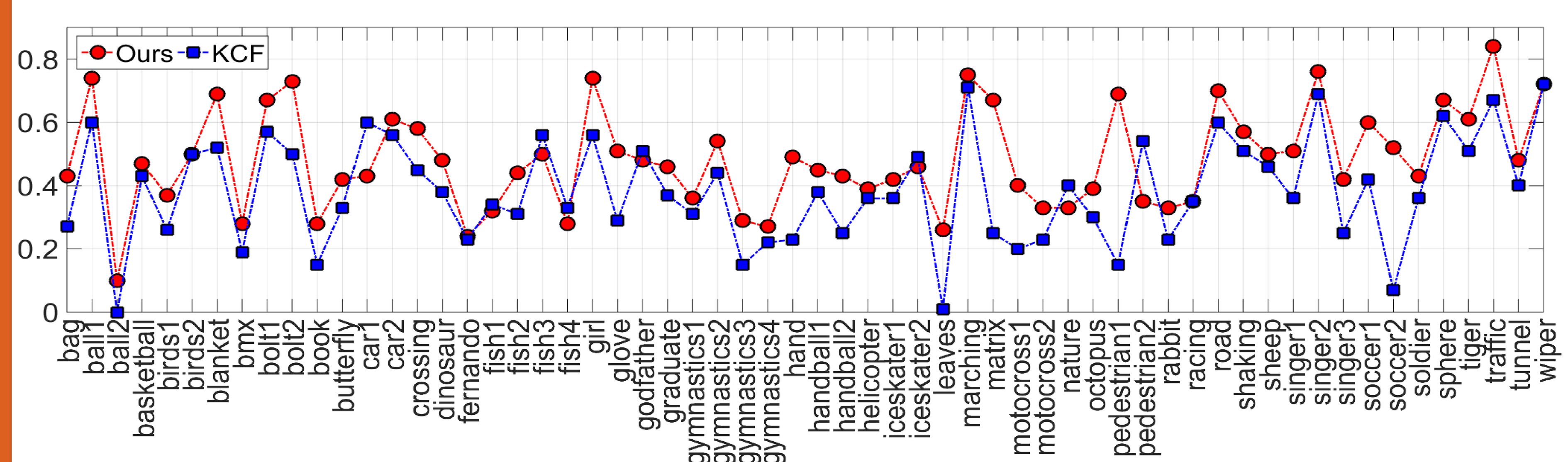
Scale Integration: We perform max-pooling on the aposterior distribution which allows a smoother scale change transition across consecutive frames.

$$\max_i \; P(s_i|\mathbf{y})=P(\mathbf{y}|s_i)P(s_i)$$

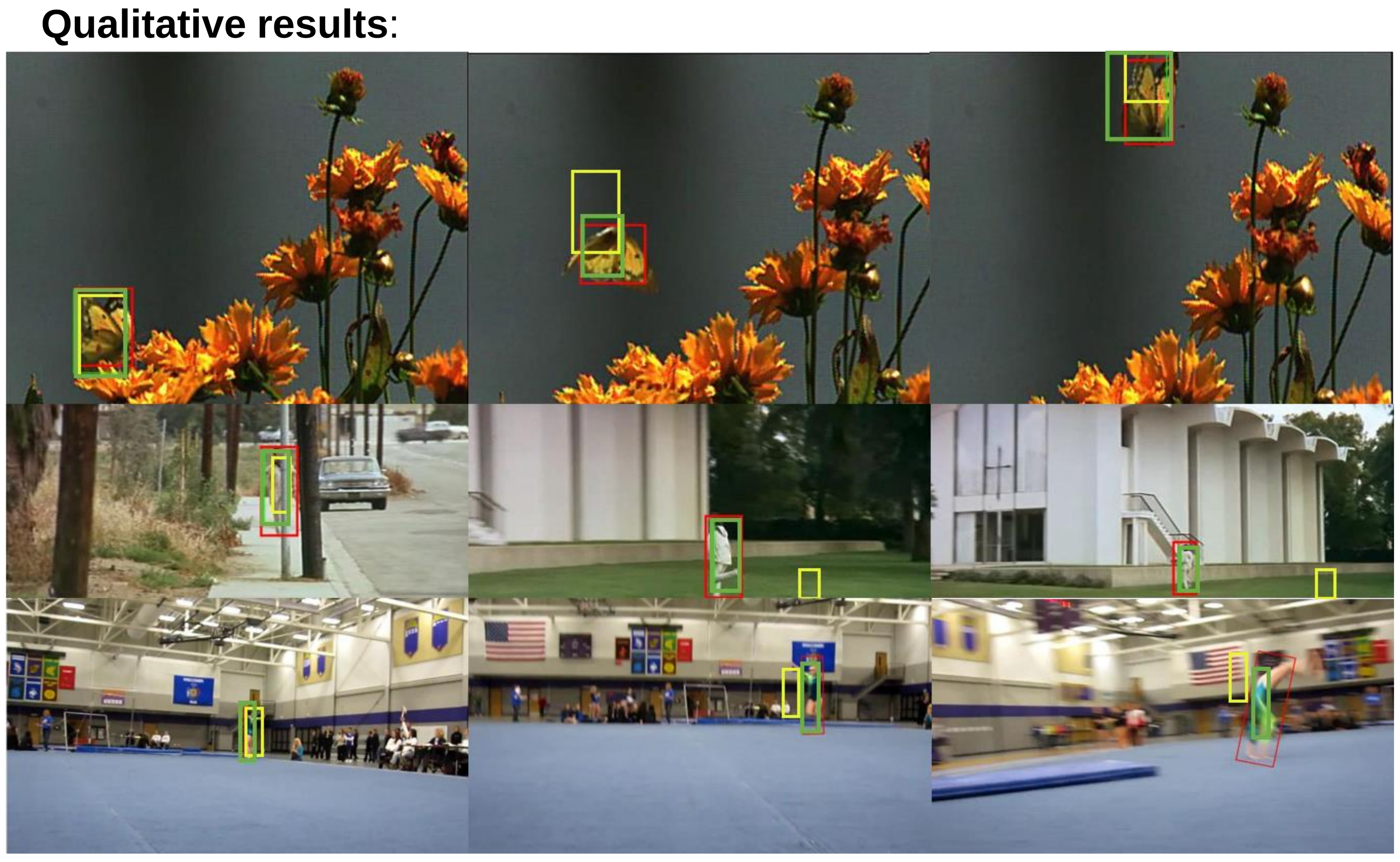
Experiments

Quantitative Comparison on VOT2014/2015 Datasets: We compare our method against state-of-art trackers on VOT2014 and different versions of our tracker on VOT2015.

Tracker	Acc. Rank	Rob. Rank	Rank	FPS	
KCF+MT+Sc (ours)	**2.68**	**2.96**	**2.82**	25.1	VOT 2014
KCF_Scale	3.00	3.48	3.24	47.42	VOT 2014
KCF	3.46	3.13	3.29	**66.5**	VOT 2014
Struck	3.68	3.85	3.75	19.77	VOT 2014
MIL	5.02	3.88	4.45	1.94	VOT 2014
IVT	5.00	4.65	4.81	27.51	VOT 2014
NCC	5.21	6.08	5.65	27.9	VOT 2014
KCF+MT+Sc (ours)	**1.80**	**1.95**	**1.88**	24.3	VOT 2015
KCF+MT	2.04	2.04	2.04	31.54	VOT 2015
KCF	2.16	2.01	2.08	**44.76**	VOT 2015



Qualitative results:



Ground Truth   Ours   KCF

References:

J. F. Henriques, R. Caseiro, P. Martins, and J. Batista. Highspeed tracking with kernelized correlation filters. PAMI 2015.

M. Danelljan, G. H¨ager, F. Khan, and M. Felsberg. Accurate scale estimation for robust visual tracking. BMVC 2014.

Acknowledgements. Research reported was supported by competitive funding from KAUST.