Context-Aware Correlation Filter Tracking

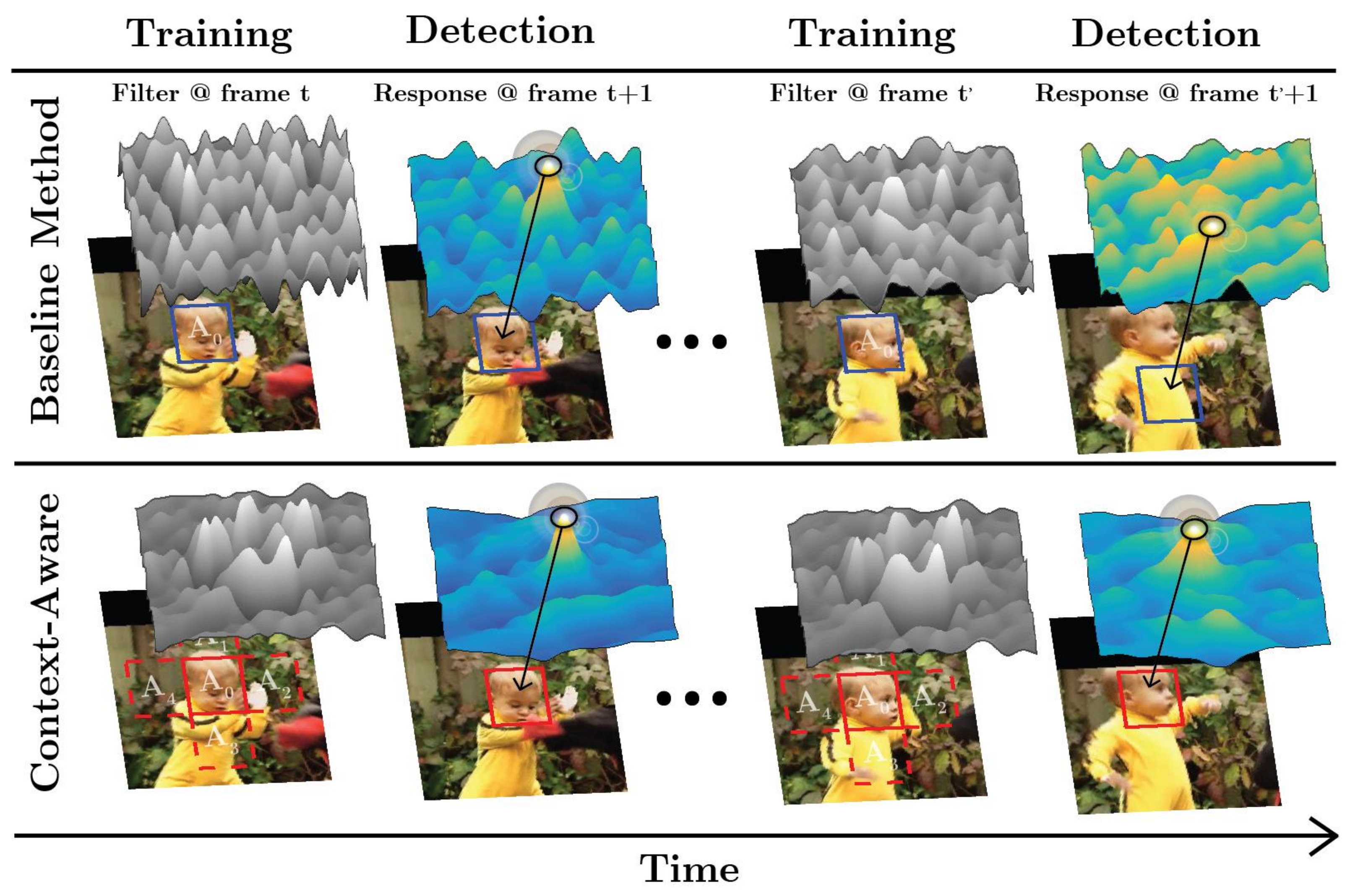
Matthias Mueller, Neil Smith, Bernard Ghanem
King Abdullah University of Science and Technology (KAUST)

Code Repository: https://goo.gl/gJ2vTs
Project Website: https://goo.gl/6zOpKO

Contributions

- Generic and efficient framework to incorporate global context within CF trackers
- Reformulation of the ridge regression problem with closed form solution for single and multi-dimensional features in the primal and dual domain
- Significant performance improvement of many CF trackers with only a modest impact on their frame rate

Methodology



- Learn a filter with high response for the object patch $\mathbf{a}_0$ and close to zero response for k context patches $\mathbf{a}_i$
- Approximate dense sampling by circulantly shifting $\mathbf{a}_0$ and $\mathbf{a}_i$ with efficient solution in the Fourier domain

$$\min_{\mathbf{w}} \; \underbrace{\|\mathbf{A}_0\mathbf{w} - \mathbf{y}\|_2^2 + \lambda_1\|\mathbf{w}\|_2^2}_{\text{Standard Formulation}} + \underbrace{\lambda_2 \sum_{i=1}^{k} \|\mathbf{A}_i\mathbf{w}\|_2^2}_{\text{Context Term}}$$

Context-Aware CF Solution

- Primal Domain: Single-Channel Features
  - Training:
$$\hat{\mathbf{w}} = \frac{\hat{\mathbf{a}}_0^* \odot \hat{\mathbf{y}}}{\hat{\mathbf{a}}_0^* \odot \hat{\mathbf{a}}_0 + \lambda_1 + \lambda_2 \sum_{i=1}^{k} \hat{\mathbf{a}}_i^* \odot \hat{\mathbf{a}}_i}$$
  - Detection: $\hat{\mathbf{r}}_p = \hat{\mathbf{z}} \odot \hat{\mathbf{w}}$

- $\hat{\mathbf{a}}_0$ and $\hat{\mathbf{a}}_i$ are the DFTs of $\mathbf{a}_0$ and $\mathbf{a}_i$
- $\mathbf{a}_0^*$ and $\mathbf{a}_i^*$ are the conjugates of $\hat{\mathbf{a}}_0$ and $\hat{\mathbf{a}}_i$
- $\hat{\mathbf{y}}$ is the DFT of regression target $\mathbf{y}$
- λ_1 and λ_2 are regularization parameters
- $\hat{\mathbf{w}}$ is the DFT of the learned filter $\mathbf{w}$
- $\hat{\mathbf{z}}$ is the DFT of a new image patch $\mathbf{z}$
- $\hat{\mathbf{r}}_p$ is the DFT of the primal response

- Dual Domain: Single-Channel Features
  - Training:
$$\hat{\boldsymbol{\alpha}} = \begin{bmatrix} \mathrm{diag}(\mathbf{d}_{00}) & \dots & \mathrm{diag}(\mathbf{d}_{0k}) \\ \vdots & \ddots & \vdots \\ \mathrm{diag}(\mathbf{d}_{k0}) & \dots & \mathrm{diag}(\mathbf{d}_{kk}) \end{bmatrix}^{-1} \begin{bmatrix} \hat{\mathbf{y}} \\ \vdots \\ \mathbf{0} \end{bmatrix}$$
  - Detection:
$$\hat{\mathbf{r}}_d = \hat{\mathbf{z}} \odot \hat{\mathbf{a}}_0^* \odot \hat{\boldsymbol{\alpha}}_0 + \sqrt{\lambda_2} \sum_{i=1}^{k} \hat{\mathbf{z}} \odot \hat{\mathbf{a}}_i^* \odot \hat{\boldsymbol{\alpha}}_i$$

- Vectors $\mathbf{d}_{jl}$ with $j,l \in \{1,...,k\}$ are given by:
$$\begin{cases} \mathbf{d}_{00} &= \hat{\mathbf{a}}_0 \odot \hat{\mathbf{a}}_0^* + \lambda_1 \\ \mathbf{d}_{jj} &= \lambda_2 \left(\hat{\mathbf{a}}_j \odot \hat{\mathbf{a}}_j^*\right) + \lambda_1, \; j \neq 0 \\ \mathbf{d}_{jl} &= \sqrt{\lambda_2} \left(\hat{\mathbf{a}}_j \odot \hat{\mathbf{a}}_l^*\right), \; j \neq l \end{cases}$$
- $\hat{\boldsymbol{\alpha}}$ is the DFT of the dual variable $\boldsymbol{\alpha}$
- $\hat{\mathbf{r}}_d$ is the DFT of the dual response

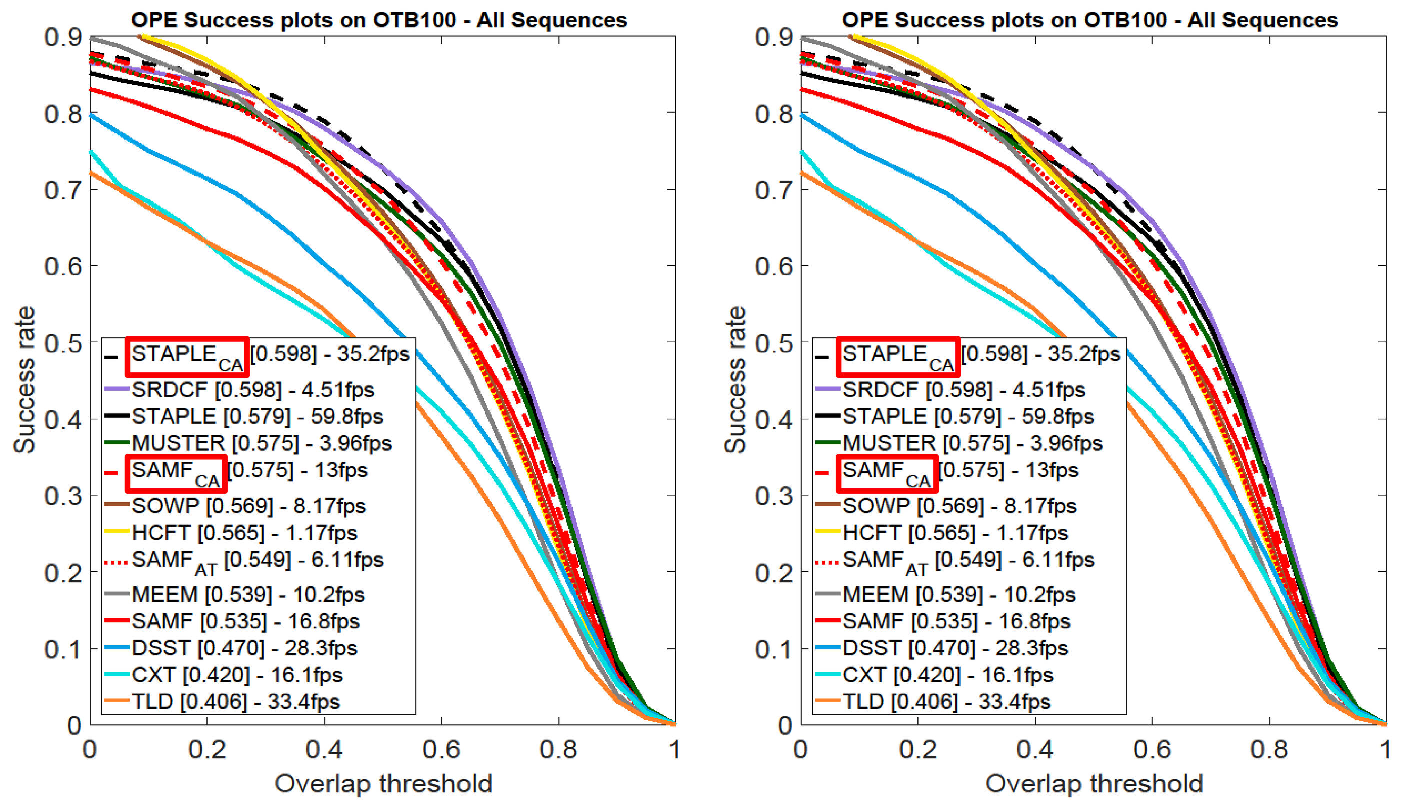
- Dual Domain: Multi-Channel Features
  - Training:
$$\hat{\mathbf{w}} = \begin{bmatrix} \bar{\mathbf{C}}_{11} & \dots & \bar{\mathbf{C}}_{1m} \\ \vdots & \ddots & \vdots \\ \bar{\mathbf{C}}_{m1} & \dots & \bar{\mathbf{C}}_{mm} \end{bmatrix}^{-1} \begin{bmatrix} \mathrm{diag}(\hat{\mathbf{a}}_{01}^* \odot \hat{\mathbf{y}}) \\ \vdots \\ \mathrm{diag}(\hat{\mathbf{a}}_{0m}^* \odot \hat{\mathbf{y}}) \end{bmatrix}$$
  - Detection: $\hat{\mathbf{r}}_p = \sum_{j=1}^{m} \hat{\mathbf{z}}_j \odot \hat{\mathbf{w}}_j$

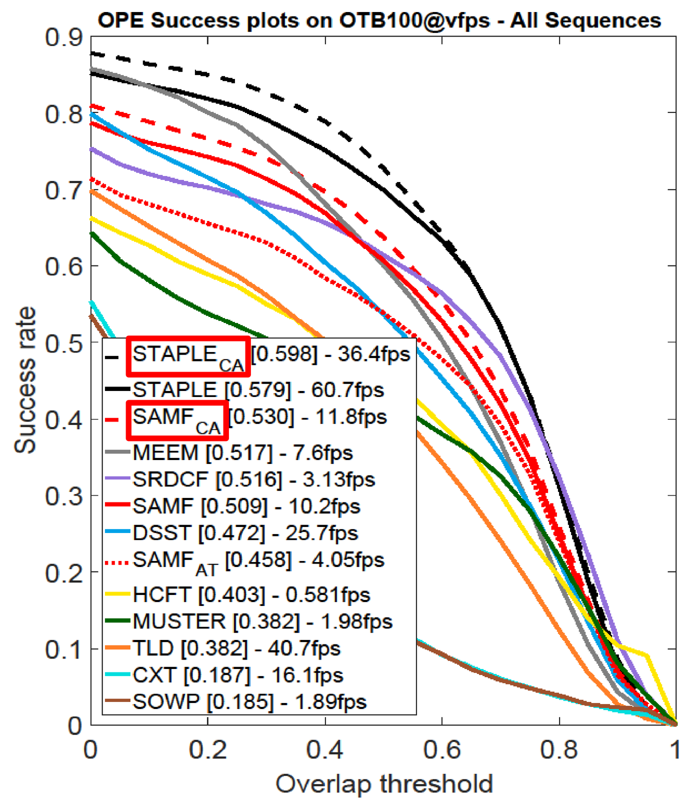
- Blocks $\mathbf{C}_{jl}$ with $j,l \in \{1,...,m\}$ are given by:
$$\begin{cases} \bar{\mathbf{C}}_{jj} &= \mathrm{diag}\left(\hat{\mathbf{a}}_{0j}^* \odot \hat{\mathbf{a}}_{0j} + \lambda_2 \sum_{i=1}^{k} \hat{\mathbf{a}}_{ij}^* \odot \hat{\mathbf{a}}_{ij}\right) + \lambda_1 \mathbf{I} \\ \bar{\mathbf{C}}_{jl} &= \mathrm{diag}\left(\hat{\mathbf{a}}_{0j}^* \odot \hat{\mathbf{a}}_{0l} + \lambda_2 \sum_{i=1}^{k} \hat{\mathbf{a}}_{ij}^* \odot \hat{\mathbf{a}}_{il}\right), \; j \neq l \end{cases}$$
- $\hat{\mathbf{a}}_{0j}$ and $\hat{\mathbf{a}}_{ij}$ are the DFTs of $\mathbf{a}_{0j}$ and $\mathbf{a}_{ij}$
- $\mathbf{a}_{0j}^*$ and $\mathbf{a}_{ij}^*$ are the conjugates of $\hat{\mathbf{a}}_{0j}$ and $\hat{\mathbf{a}}_{ij}$
- $\hat{\mathbf{z}}_j$ is the DFT of a new image patch $\mathbf{z}_j$

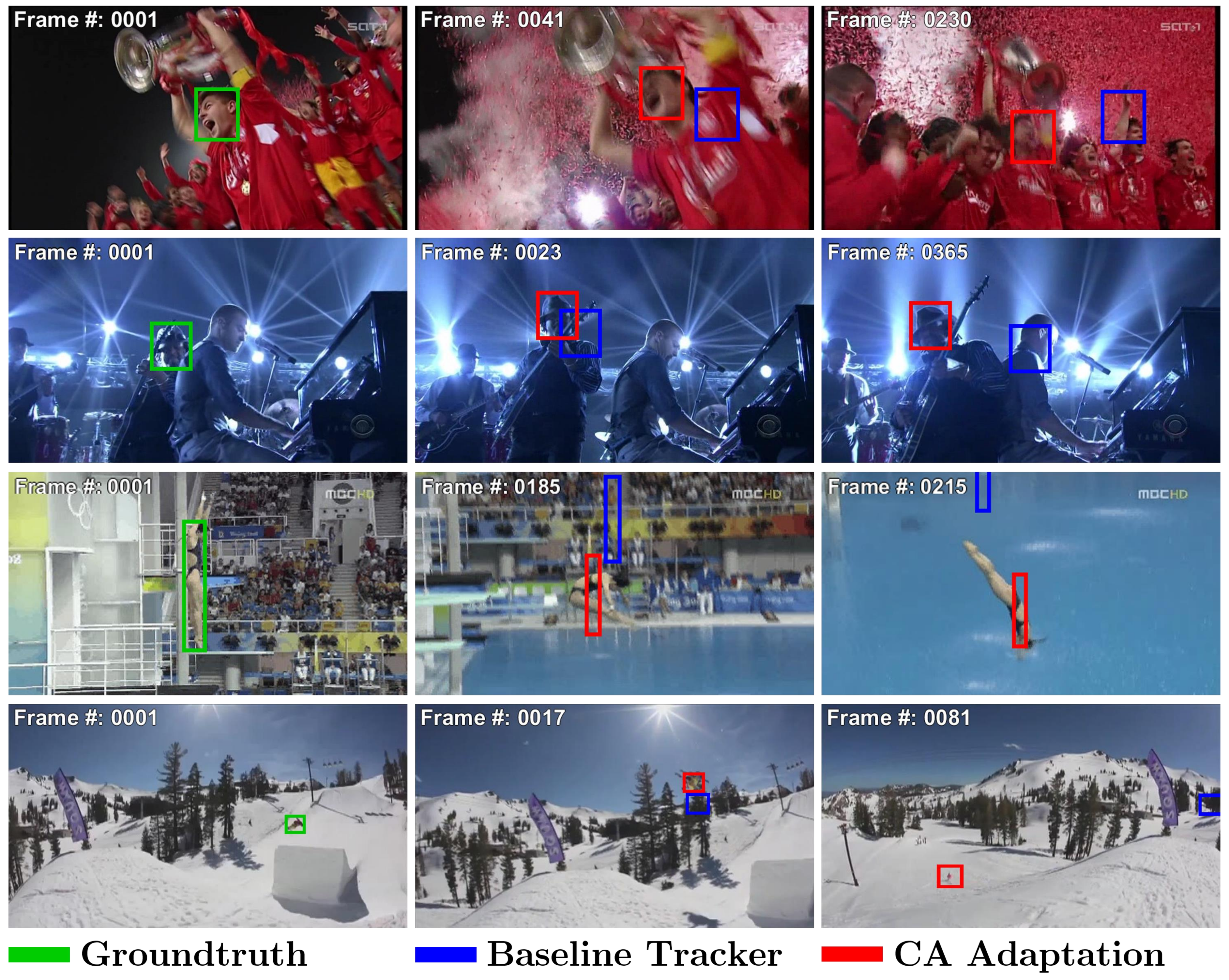
- Dual Domain: Multi-Channel Features
  - Training:
$$\hat{\bar{\boldsymbol{\alpha}}} = \begin{bmatrix} \mathrm{diag}(\bar{\mathbf{d}}_{00}) & \dots & \mathrm{diag}(\bar{\mathbf{d}}_{0k}) \\ \vdots & \ddots & \vdots \\ \mathrm{diag}(\bar{\mathbf{d}}_{k0}) & \dots & \mathrm{diag}(\bar{\mathbf{d}}_{kk}) \end{bmatrix}^{-1} \begin{bmatrix} \hat{\mathbf{y}} \\ \vdots \\ \mathbf{0} \end{bmatrix}$$
  - Detection:
$$\hat{\mathbf{r}}_d = \sum_{i=1}^{m} \hat{\mathbf{z}}_i \odot \hat{\mathbf{a}}_{0i}^* \odot \hat{\bar{\boldsymbol{\alpha}}}_0 + \sqrt{\lambda_2} \sum_{j=1}^{k} \sum_{i=1}^{m} \hat{\mathbf{z}}_i \odot \hat{\mathbf{a}}_{ji}^* \odot \hat{\bar{\boldsymbol{\alpha}}}_j$$

- Vectors $\bar{\mathbf{d}}_{jl}$ with $j,l \in \{1,...,k\}$ are given by:
$$\begin{cases} \bar{\mathbf{d}}_{00} &= \sum_{i=1}^{m} \left(\hat{\mathbf{a}}_{0i} \odot \hat{\mathbf{a}}_{0i}^*\right) + \lambda_1 \\ \bar{\mathbf{d}}_{jj} &= \lambda_2 \sum_{i=1}^{m} \left(\hat{\mathbf{a}}_{ji} \odot \hat{\mathbf{a}}_{ji}^*\right) + \lambda_1, \; j \neq 0 \\ \bar{\mathbf{d}}_{jl} &= \sqrt{\lambda_2} \sum_{i=1}^{m} \left(\hat{\mathbf{a}}_{ji} \odot \hat{\mathbf{a}}_{li}^*\right), \; j \neq l \end{cases}$$
- $\hat{\mathbf{a}}_{0i}$ and $\hat{\mathbf{a}}_{ji}$ are the DFTs of $\mathbf{a}_{0i}$ and $\mathbf{a}_{ji}$
- $\mathbf{a}_{0i}^*$ and $\mathbf{a}_{ji}^*$ are the conjugates of $\hat{\mathbf{a}}_{0i}$ and $\hat{\mathbf{a}}_{ji}$
- $\hat{\mathbf{z}}_i$ is the DFT of a new image patch $\mathbf{z}_i$

Experimental Results







Acknowledgements. Research reported was supported by competitive funding from King Abdullah University of Science and Technology.