Supplementary Material

Adaptive Decontamination of the Training Set: A Unified Formulation for Discriminative Visual Tracking

Martin Danelljan, Gustav Häger, Fahad Shahbaz Khan, Michael Felsberg
Computer Vision Laboratory, Department of Electrical Engineering, Linköping University, Sweden
{martin.danelljan, gustav.hager, fahad.khan, michael.felsberg}@liu.se

In this supplementary material of [3], we first prove the result $\alpha_k \to \rho_k$ in the case when $\mu \to 0$ (stated on page 4 in [3]). The derivation is performed in section 1. In section 2 we present the per-video and all attribute results on the OTB-2015 dataset [19]. Finally, section 3 contains per-video results on the Temple-Color dataset [14].

1. Derivation of $\alpha_k \to \rho_k$ When $\mu \to 0$

Here, we derive that the computed sample weights $\alpha_k$ converge to the prior weights $\rho_k$ when the flexibility parameter $\mu$ is reduced in our joint formulation. That is, we derive that $\alpha_k \to \rho_k$ when $\mu \to 0$ for fixed model parameters $\theta \in \Omega$. Our joint optimization problem is given by (corresponds to eq. (3) in the paper),

$$\begin{align*}
\text{minimize} & \quad J(\theta, \alpha) = \sum_{k=1}^{t} \alpha_k \sum_{j=1}^{n_k} L(\theta; x_{jk}, y_{jk}) + \frac{1}{\mu} \sum_{k=1}^{t} \alpha_k^2 \rho_k + \lambda R(\theta) \\
\text{subject to} & \quad \alpha_k \geq 0, \quad k = 1, \ldots, t \\
& \quad \sum_{k=1}^{t} \alpha_k = 1.
\end{align*}$$

(1a)

(1b)

(1c)

Here, the prior weights are positive $\rho_k > 0$ and sum up to one,

$$\sum_{k=1}^{t} \rho_k = 1$$

(2)

We let the model parameters $\theta$ be fixed and define the total loss in frame $k$ by,

$$L_k = \sum_{j=1}^{n_k} L(\theta; x_{jk}, y_{jk}).$$

(3)

Minimizing the joint formulation (1) with respect to the weights $\alpha_k$ is then equivalent to solving the following quadratic programming problem,

$$\begin{align*}
\text{minimize} & \quad J_2(\alpha) = \sum_{k=1}^{t} L_k \alpha_k + \frac{1}{\mu} \sum_{k=1}^{t} \alpha_k^2 \rho_k \\
\text{subject to} & \quad \alpha_k \geq 0, \quad k = 1, \ldots, t \\
& \quad \sum_{k=1}^{t} \alpha_k = 1.
\end{align*}$$

(4a)

(4b)

(4c)
We temporarily ignore the inequality constraint (4b) and introduce Lagrange multipliers for the constraint (4c),

\[ L(\alpha, \eta) = \sum_{k=1}^{t} L_k \alpha_k + \frac{1}{\mu} \sum_{k=1}^{t} \frac{\alpha_k^2}{\rho_k} - \eta \cdot \left( \sum_{k=1}^{t} \alpha_k - 1 \right). \] (5)

Here, \( \eta \) denotes the Lagrange multiplier. Differentiation w.r.t. \( \alpha_k \) gives,

\[ \frac{\partial L}{\partial \alpha_k} = L_k + \frac{2}{\mu} \frac{\alpha_k}{\rho_k} - \eta, \quad k = 1, \ldots, t. \] (6)

The stationary point is computed by setting the partial derivatives to zero,

\[ \frac{\partial L}{\partial \alpha_k} = 0 \iff \alpha_k = \frac{\mu \eta}{2} \rho_k - \frac{\mu}{2} L_k \rho_k, \quad k = 1, \ldots, t \] (7)

The Lagrange multiplier \( \eta \) is computed by summing both sides of (7) over \( k \) and using (4c) and (2),

\[ \sum_{k=1}^{t} \alpha_k = \sum_{k=1}^{t} \left( \frac{\mu \eta}{2} \rho_k - \frac{\mu}{2} L_k \rho_k \right) \iff \]
\[ 1 = \frac{\mu \eta}{2} - \frac{\mu}{2} \sum_{k=1}^{t} L_k \rho_k \iff \]
\[ \eta = \frac{2}{\mu} \mu + \sum_{k=1}^{t} L_k \rho_k. \] (8)

Using the result (8) in (7) gives,

\[ \alpha_k = \rho_k + \frac{\mu}{2} \left( \rho_k \sum_{l=1}^{t} L_l \rho_l - L_k \rho_k \right) \] (9)

From (9) it follows that \( \alpha_k \to \rho_k \) when \( \mu \to 0 \). To show that the inequality constraint (4b) also holds in the limit \( \mu \to 0 \), we define the constant

\[ \delta = \min_k \rho_k \cdot \left| \rho_k \sum_{l=1}^{t} L_l \rho_l - L_k \rho_k \right|^{-1}. \] (10)

This choice ensures that \( \alpha_k > 0, \forall k \) for \( 0 < \mu < \delta \). The inequality constraint (4b) is thus satisfied for \( 0 < \mu < \delta \). This proves that the limit \( \mu \to 0 \) of (9) is also the limit of the solution \( \alpha_k \) of (4). Hence, \( \alpha_k \to \rho_k \) in (4) when \( \mu \to 0 \).

2. Detailed Results on OTB-2015

We provide detailed results on OTB-2015 [19] with 100 videos. The videos and ground truth are available at https://sites.google.com/site/benchmarkpami/. Figure 1 contains the success plots for all 11 attributes. Table 2 shows the per-video overlap precision for all trackers.

3. Detailed Results on Temple-Color

We also report detailed results on the Temple-Color dataset [14] with 128 videos. The videos and ground truth are available at http://www.dabi.temple.edu/~hbling/data/TColor-128/TColor-128.html. The per-video overlap precision for all trackers in our comparison are reported in table 2.
Figure 1. Success plots on OTB-2015 [19]. We show the total success plot (top-left) and the success plots for all 11 attributes. The title of each attribute plot contains the name of the attribute and the number of videos associated with it. The area-under-the-curve score is shown in the legend. For clarity, only the top 10 trackers in each plot are displayed in the legend. Our approach obtains the best results on all 11 attributes.
Table 1. Per-video results on the OTB-2015 dataset [19] with 100 videos. The results are shown in terms of overlap precision (in percent), which corresponds to the PASCAL criterion. The two best results for each sequence are shown in red and blue respectively. Our approach achieves a significant gain of 3.8% in average overlap precision, compared to the best existing tracker (SRDCF).
Table 2. Per-video results on the Temple-Color dataset [14] with 128 videos. The results are shown in terms of overlap precision (in percent), which corresponds to the PASCAL criterion. The two best results for each sequence are shown in red and blue respectively. Our approach achieves a significant gain of 3.6% in average overlap precision, compared to the best existing trackers (MEEM and SRDCF).
References