Supplementary Material: Discriminant Tracking Using Tensor Representation with Semi-supervised Improvement

Jin Gao[1], Junliang Xing[1], Weiming Hu[1], and Steve Maybank[2]

[1]National Laboratory of Pattern Recognition, Institute of Automation, CAS, Beijing
[2]Department of Computer Science, Birkbeck College, London
{jgao10, jlxing, wmhu}@nlpr.ia.ac.cn, sjmaybank@dcs.bbk.ac.uk

Abstract

In this supplementary document, we provide proofs of two propositions in Section 1 and Section 2 to support the claims in the main paper; describe how the video in the zip file showcase the experimental results in Section 3.

1. Proof of Proposition 1

Proposition 1. *Solving the objective function Eq. (1) of the main paper in the* 2^{nd}*-order tensor case is equivalent to solving the following constrained optimization problem:*

$$\begin{aligned} \underset{\mathbf{U},\mathbf{V}}{\text{minimize}} \quad & \sum_{i,j} \left\| \mathbf{U}^T \mathcal{X}_i \mathbf{V} - \mathbf{U}^T \mathcal{X}_j \mathbf{V} \right\|_F^2 W_{ij}, \\ \text{subject to} \quad & \sum_{i,j} \left\| \mathbf{U}^T \mathcal{X}_i \mathbf{V} - \mathbf{U}^T \mathcal{X}_j \mathbf{V} \right\|_F^2 W_{ij}^p = d \end{aligned} \tag{1}$$

where $\mathbf{U}^T = \mathbf{M}^1$, $\mathbf{V}^T = \mathbf{M}^2$*, and* d *is a constant.*

Proof. Let $\mathbf{M}^1 = \mathbf{U}^T \in \mathbb{R}^{l_1 \times m_1}$, $\mathbf{M}^2 = \mathbf{V}^T \in \mathbb{R}^{l_2 \times m_2}$, where $l_1 < m_1$, and $l_2 < m_2$. Consider $\mathcal{Y}_i = \mathcal{X}_i \times_1 \mathbf{M}^1 \times_2 \mathbf{M}^2$. In the 2^{nd}-order tensor case, the mode-1 *flattening* of the tensor $\mathcal{A}$ is $\mathbf{A}_{(1)} = \mathcal{A}$ and the mode-2 *flattening* is $\mathbf{A}_{(2)} = \mathcal{A}^T$. Denote $\mathcal{X}_i \times_1 \mathbf{M}^1$ as the tensor $\mathcal{T}$. Then, $\mathcal{T}$ can be computed by matrix multiplication $\mathbf{T}_{(1)} = \mathbf{M}^1 \mathbf{X}_{i(1)} = \mathbf{U}^T \mathcal{X}_i$, followed by retensorization for mode-1 *folding*. That is to say, $\mathcal{T} = \mathbf{U}^T \mathcal{X}_i$. Likewise, $\mathcal{Y}_i$ can be computed by matrix multiplication $\mathbf{Y}_{i(2)} = \mathbf{M}^2 \mathbf{T}_{(2)} = \mathbf{V}^T \mathcal{T}^T$, followed by retensorization for mode-2 *folding*. Then, $\mathcal{Y}_i$ can be rederived as: $\mathcal{Y}_i = \left(\mathbf{Y}_{i(2)}\right)^T = \left(\mathbf{V}^T \mathcal{T}^T\right)^T = \mathcal{T}\mathbf{V} = \mathbf{U}^T \mathcal{X}_i \mathbf{V}$. □

2. Proof of Proposition 2

Proposition 2. *Let* $\mathbf{D}$ *and* $\mathbf{D}^p$ *be diagonal matrices, where* $D_{ii} = \sum_{j \neq i} W_{ij}$ *and* $D_{ii}^p = \sum_{j \neq i} W_{ij}^p$*. The optimization problem Eq. (1) in this supplementary material can be reformulated as either of the following two optimization problems:* $\min_{\mathbf{U},\mathbf{V}} \mathrm{tr}\left(\frac{\mathbf{U}^T(\mathbf{D}_V - \mathbf{W}_V)\mathbf{U}}{\mathbf{U}^T(\mathbf{D}_V^p - \mathbf{W}_V^p)\mathbf{U}} \right)$ *or* $\min_{\mathbf{U},\mathbf{V}} \mathrm{tr}\left(\frac{\mathbf{V}^T(\mathbf{D}_U - \mathbf{W}_U)\mathbf{V}}{\mathbf{V}^T(\mathbf{D}_U^p - \mathbf{W}_U^p)\mathbf{V}} \right)$*, where* $\mathbf{D}_V = \sum_i D_{ii} \mathcal{X}_i \mathbf{V}\mathbf{V}^T \mathcal{X}_i^T$*,* $\mathbf{W}_V = \sum_{i,j} W_{ij} \mathcal{X}_i \mathbf{V}\mathbf{V}^T \mathcal{X}_j^T$*,* $\mathbf{D}_V^p = \sum_i D_{ii}^p \mathcal{X}_i \mathbf{V}\mathbf{V}^T \mathcal{X}_i^T$*,* $\mathbf{W}_V^p = \sum_{i,j} W_{ij}^p \mathcal{X}_i \mathbf{V}\mathbf{V}^T \mathcal{X}_j^T$*,* $\mathbf{D}_U = \sum_i D_{ii} \mathcal{X}_i^T \mathbf{U}\mathbf{U}^T \mathcal{X}_i$*,* $\mathbf{W}_U = \sum_{i,j} W_{ij} \mathcal{X}_i^T \mathbf{U}\mathbf{U}^T \mathcal{X}_j$*,* $\mathbf{D}_U^p = \sum_i D_{ii}^p \mathcal{X}_i^T \mathbf{U}\mathbf{U}^T \mathcal{X}_i$*,* $\mathbf{W}_U^p = \sum_{i,j} W_{ij}^p \mathcal{X}_i^T \mathbf{U}\mathbf{U}^T \mathcal{X}_j$*.*

Proof. In the 2^{nd}-order tensor case, $\mathcal{Y}_i = \mathbf{U}^T \mathcal{X}_i \mathbf{V}$. Since $\|\mathcal{A}\|_F^2 = \operatorname{tr}\left(\mathcal{A}\mathcal{A}^T\right)$ for a 2^{nd}-order tensor $\mathcal{A}$, we see that:

$$
\begin{aligned}
& \frac{1}{2}\sum_{i,j} \left\|\mathbf{U}^T \mathcal{X}_i \mathbf{V} - \mathbf{U}^T \mathcal{X}_j \mathbf{V}\right\|_F^2 W_{ij} \\
= & \frac{1}{2}\sum_{i,j} \left\|\mathcal{Y}_i - \mathcal{Y}_j\right\|_F^2 W_{ij} \\
= & \frac{1}{2}\sum_{i,j} \operatorname{tr}\left(\left(\mathcal{Y}_i - \mathcal{Y}_j\right)\left(\mathcal{Y}_i - \mathcal{Y}_j\right)^T\right) W_{ij} \\
= & \frac{1}{2}\sum_{i,j} \operatorname{tr}\left(\mathcal{Y}_i\mathcal{Y}_i^T + \mathcal{Y}_j\mathcal{Y}_j^T - \mathcal{Y}_i\mathcal{Y}_j^T - \mathcal{Y}_j\mathcal{Y}_i^T\right) W_{ij} \\
= & \operatorname{tr}\left(\sum_i D_{ii}\mathcal{Y}_i\mathcal{Y}_i^T - \sum_{i,j} W_{ij}\mathcal{Y}_i\mathcal{Y}_j^T\right) \\
= & \operatorname{tr}\left(\sum_i D_{ii}\mathbf{U}^T\mathcal{X}_i\mathbf{V}\mathbf{V}^T\mathcal{X}_i^T\mathbf{U} - \sum_{i,j} W_{ij}\mathbf{U}^T\mathcal{X}_i\mathbf{V}\mathbf{V}^T\mathcal{X}_j^T\mathbf{U}\right) \\
= & \operatorname{tr}\left(\mathbf{U}^T\left(\sum_i D_{ii}\mathcal{X}_i\mathbf{V}\mathbf{V}^T\mathcal{X}_i^T - \sum_{i,j} W_{ij}\mathcal{X}_i\mathbf{V}\mathbf{V}^T\mathcal{X}_j^T\right)\mathbf{U}\right) \\
\doteq & \operatorname{tr}\left(\mathbf{U}^T\left(\mathbf{D}_V - \mathbf{W}_V\right)\mathbf{U}\right)
\end{aligned}
$$

where $\mathbf{D}_V = \sum_i D_{ii}\mathcal{X}_i\mathbf{V}\mathbf{V}^T\mathcal{X}_i^T$, and $\mathbf{W}_V = \sum_{i,j} W_{ij}\mathcal{X}_i\mathbf{V}\mathbf{V}^T\mathcal{X}_j^T$. Likewise, $\frac{1}{2}\sum_{i,j}\left\|\mathbf{U}^T\mathcal{X}_i\mathbf{V} - \mathbf{U}^T\mathcal{X}_j\mathbf{V}\right\|_F^2 W_{ij}^p \doteq \operatorname{tr}\left(\mathbf{U}^T\left(\mathbf{D}_V^p - \mathbf{W}_V^p\right)\mathbf{U}\right)$, where $\mathbf{D}_V^p = \sum_i D_{ii}^p\mathcal{X}_i\mathbf{V}\mathbf{V}^T\mathcal{X}_i^T$, and $\mathbf{W}_V^p = \sum_{i,j} W_{ij}^p\mathcal{X}_i\mathbf{V}\mathbf{V}^T\mathcal{X}_j^T$.

Similarly, $\|\mathcal{A}\|_F^2 = \operatorname{tr}\left(\mathcal{A}^T\mathcal{A}\right)$ for a 2^{nd}-order tensor $\mathcal{A}$, so we also have:

$$
\begin{aligned}
& \frac{1}{2}\sum_{i,j} \left\|\mathbf{U}^T \mathcal{X}_i \mathbf{V} - \mathbf{U}^T \mathcal{X}_j \mathbf{V}\right\|_F^2 W_{ij} \\
= & \frac{1}{2}\sum_{i,j} \left\|\mathcal{Y}_i - \mathcal{Y}_j\right\|_F^2 W_{ij} \\
= & \frac{1}{2}\sum_{i,j} \operatorname{tr}\left(\left(\mathcal{Y}_i - \mathcal{Y}_j\right)^T\left(\mathcal{Y}_i - \mathcal{Y}_j\right)\right) W_{ij} \\
= & \frac{1}{2}\sum_{i,j} \operatorname{tr}\left(\mathcal{Y}_i^T\mathcal{Y}_i + \mathcal{Y}_j^T\mathcal{Y}_j - \mathcal{Y}_i^T\mathcal{Y}_j - \mathcal{Y}_j^T\mathcal{Y}_i\right) W_{ij} \\
= & \operatorname{tr}\left(\sum_i D_{ii}\mathcal{Y}_i^T\mathcal{Y}_i - \sum_{i,j} W_{ij}\mathcal{Y}_i^T\mathcal{Y}_j\right) \\
= & \operatorname{tr}\left(\sum_i D_{ii}\mathbf{V}^T\mathcal{X}_i^T\mathbf{U}\mathbf{U}^T\mathcal{X}_i\mathbf{V} - \sum_{i,j} W_{ij}\mathbf{V}^T\mathcal{X}_i^T\mathbf{U}\mathbf{U}^T\mathcal{X}_j\mathbf{V}\right) \\
= & \operatorname{tr}\left(\mathbf{V}^T\left(\sum_i D_{ii}\mathcal{X}_i^T\mathbf{U}\mathbf{U}^T\mathcal{X}_i - \sum_{i,j} W_{ij}\mathcal{X}_i^T\mathbf{U}\mathbf{U}^T\mathcal{X}_j\right)\mathbf{V}\right) \\
\doteq & \operatorname{tr}\left(\mathbf{V}^T\left(\mathbf{D}_U - \mathbf{W}_U\right)\mathbf{V}\right)
\end{aligned}
$$

where $\mathbf{D}_U = \sum_i D_{ii}\mathcal{X}_i^T\mathbf{U}\mathbf{U}^T\mathcal{X}_i$, and $\mathbf{W}_U = \sum_{i,j} W_{ij}\mathcal{X}_i^T\mathbf{U}\mathbf{U}^T\mathcal{X}_j$. Likewise, $\frac{1}{2}\sum_{i,j}\left\|\mathbf{U}^T\mathcal{X}_i\mathbf{V} - \mathbf{U}^T\mathcal{X}_j\mathbf{V}\right\|_F^2 W_{ij}^p \doteq$

$\mathrm{tr}\left(\mathbf{V}^T\left(\mathbf{D}_U^p - \mathbf{W}_U^p\right)\mathbf{V}\right)$, where $\mathbf{D}_U^p = \sum_i D_{ii}^p \mathcal{X}_i^T \mathbf{U}\mathbf{U}^T \mathcal{X}_i$, and $\mathbf{W}_U^p = \sum_{i,j} W_{ij}^p \mathcal{X}_i^T \mathbf{U}\mathbf{U}^T \mathcal{X}_j$.

Finally, the optimization problem Eq. (1) in this supplementary material can be reformulated as either of the following two optimization problems:

$$\min_{\mathbf{U},\mathbf{V}} \mathrm{tr}\left(\frac{\mathbf{U}^T\left(\mathbf{D}_V - \mathbf{W}_V\right)\mathbf{U}}{\mathbf{U}^T\left(\mathbf{D}_V^p - \mathbf{W}_V^p\right)\mathbf{U}}\right)$$

or

$$\min_{\mathbf{U},\mathbf{V}} \mathrm{tr}\left(\frac{\mathbf{V}^T\left(\mathbf{D}_U - \mathbf{W}_U\right)\mathbf{V}}{\mathbf{V}^T\left(\mathbf{D}_U^p - \mathbf{W}_U^p\right)\mathbf{V}}\right)$$

□

3. Video to Showcase the Experimental Results

The video in the zip file showcase the experimental results with respect to two aspects: i) the effectiveness of our approach; ii) comparison with competing trackers.

In the first part, we use six videos to verify the conclusions in Section 3.1 of the main paper.

- TDT can't handle the challenges in scale variation (*e.g.* **david**, **woman**), background clutter (*e.g.* **football**), fast motion and blur (*e.g.* **animal**) well;
- SSI-VDT still can't meet the challenges caused by the presence of occlusion (*e.g.* **coke11**, **football**, **woman**), fast motion and blur (*e.g.* **animal**) well;
- SSI technique and MI technique are not required in the **coke11** video;
- SSI technique can still enhance TDT in the **sylv** video, although TDT achieves good results in this video.

In the second part, we also use six videos to verify the conclusions in Section 3.2 of the main paper.

- Frag, APG-ℓ_1 and SSOBT are easily confused by the impostor object in the **dollar** video;
- SSI-TDT outperforms the other approaches significantly when heavy occlusion and pose variation appear simultaneously (*e.g.* **woman**, **coke11**, **skating1**);
- IVT can capture appearance variations due to scale change (*e.g.* **david**) and background clutter (*e.g.* **dollar**);
- VTD and VTS achieve good results over the **david**, **skating1** and **sylv** videos, however they achieve higher tracking errors and lower success rates than our approach.