Supplementary Material for "Re-ranking via Metric Fusion for Object Retrieval and Person Re-identification"

Song Bai[1] Peng Tang[2] Philip H.S. Torr[1] Longin Jan Latecki[3]
[1]University of Oxford [2]Huazhong University of Science and Technology [3]Temple University
{songbai.site,tangpeng723}@gmail.com, philip.torr@eng.ox.ac.uk, latecki@temple.edu

This document contains the supplementary material for "Re-ranking via Metric Fusion for Object Retrieval and Person Re-identification". The proofs of two key statements made in the main manuscript are given in Sec. 1. The additional performance evaluation and comparisons are given in Sec. 2.

1. Proofs

Proposition 1. *Eq. (19) converges to the closed-form solution in Eq. (14).*

Proof. By executing the iteration for t times, $\vec{A}^{(t+1)}$ can be expanded as

$$\vec{\mathbf{A}}^{(t+1)} = (\frac{\mathbf{S}}{\Lambda})^t \vec{\mathbf{A}}^{(1)} + \frac{\gamma}{\Lambda}\sum_{i=0}^{t-1}(\frac{\mathbf{S}}{\Lambda})^i \vec{\mathbf{I}}, \tag{1}$$

where

$$\mathbf{S} = \sum_{\mu,\nu=1}^{M} \beta_\mu \beta_\nu (\mathbf{S}^\mu \otimes \mathbf{S}^\nu). \tag{2}$$

It is known that the spectral radius of both $\mathbf{S}^\mu$ and $\mathbf{S}^\nu$ are no larger than 1. According to the spectral property of Kronecker product, all the eigenvalues of $\mathbf{S}^\mu \otimes \mathbf{S}^\nu$ are also in [-1,1]. Hence, the spectral radius of $\mathbf{S}/\Lambda$ is bounded by

$$\frac{1}{\Lambda}\sum_{\mu,\nu=1}^{M} \beta_\mu \beta_\nu = \frac{1}{\Lambda} = \frac{1}{\gamma+1} < 1. \tag{3}$$

Recall that $\gamma > 0$. Then, we have

$$\begin{aligned} &\lim_{t\to\infty}(\frac{\mathbf{S}}{\Lambda})^t = 0, \\ &\lim_{t\to\infty}\sum_{i=0}^{t-1}(\frac{\mathbf{S}}{\Lambda})^i = (\mathbf{I} - \frac{\mathbf{S}}{\Lambda})^{-1}. \end{aligned} \tag{4}$$

As a result, we derive that

$$\begin{aligned} \lim_{t\to\infty}\vec{\mathbf{A}}^{(t+1)} &= \frac{\gamma}{\Lambda}(\mathbf{I} - \frac{\mathbf{S}}{\Lambda})^{-1}\vec{\mathbf{I}} \\ &= \frac{\gamma}{\Lambda}(\mathbf{I} - \frac{1}{\Lambda}\sum_{\mu,\nu=1}^{M}\beta_\mu\beta_\nu \mathbf{S}^\mu \otimes \mathbf{S}^\nu)^{-1}\vec{\mathbf{I}}, \end{aligned} \tag{5}$$

which is equivalent to Eq. (14). The proof is complete. □

Proposition 2. *The minimization in Eq. (20) is equivalent to the maximization in Eq. (21).*

Recall that the objective function of Eq. (20) is

$$\min_{\boldsymbol{\beta}} \boldsymbol{\beta}^{\mathrm{T}}\mathbf{H}\boldsymbol{\beta} + \eta\|\boldsymbol{\beta}\|_2^2, \;\; \textit{s.t. } \boldsymbol{\beta} \in \Delta, \tag{6}$$

and the objective function of Eq. (21) is

$$\max_{\boldsymbol{\beta}} \boldsymbol{\beta}^{\mathrm{T}}\bar{\mathbf{H}}\boldsymbol{\beta}, \;\; \textit{s.t. } \boldsymbol{\beta} \in \Delta, \tag{7}$$

where $\bar{\mathbf{H}} = -\mathbf{H}/2 - \mathbf{H}^{\mathrm{T}}/2 - \eta\mathbf{I} + \mathbf{C}$ and $\mathbf{C} \in \mathbb{R}^{M\times M}$ is a matrix with all its entries equal to the maximum element of $(\mathbf{H}/2 + \mathbf{H}^{\mathrm{T}}/2 + \eta\mathbf{I})$.

Proof. To prove the equivalence, first we have the following preliminary fact

$$\boldsymbol{\beta}^{\mathrm{T}}\frac{\mathbf{H}-\mathbf{H}^{\mathrm{T}}}{2}\boldsymbol{\beta} \equiv 0. \tag{8}$$

It holds, since $(\mathbf{H}/2 - \mathbf{H}^{\mathrm{T}}/2)$ is an antisymmetric matrix.

Then, we have

$$\begin{aligned} &\min_{\boldsymbol{\beta}} \boldsymbol{\beta}^{\mathrm{T}}\mathbf{H}\boldsymbol{\beta} + \eta\|\boldsymbol{\beta}\|_2^2 \\ \Leftrightarrow &\min_{\boldsymbol{\beta}} \boldsymbol{\beta}^{\mathrm{T}}\frac{\mathbf{H}+\mathbf{H}^{\mathrm{T}}}{2}\boldsymbol{\beta} + \boldsymbol{\beta}^{\mathrm{T}}\frac{\mathbf{H}-\mathbf{H}^{\mathrm{T}}}{2}\boldsymbol{\beta} + \eta\|\boldsymbol{\beta}\|_2^2 \\ \Leftrightarrow &\min_{\boldsymbol{\beta}} \boldsymbol{\beta}^{\mathrm{T}}\frac{\mathbf{H}+\mathbf{H}^{\mathrm{T}}}{2}\boldsymbol{\beta} + \eta\|\boldsymbol{\beta}\|_2^2 \\ \Leftrightarrow &\min_{\boldsymbol{\beta}} \boldsymbol{\beta}^{\mathrm{T}}(\mathbf{H}/2 + \mathbf{H}^{\mathrm{T}}/2 + \eta\mathbf{I})\boldsymbol{\beta} \\ \Leftrightarrow &\max_{\boldsymbol{\beta}} \boldsymbol{\beta}^{\mathrm{T}}(-\mathbf{H}/2 - \mathbf{H}^{\mathrm{T}}/2 - \eta\mathbf{I})\boldsymbol{\beta}. \end{aligned} \tag{9}$$

As replicator equation [3] requires non-negative input, we define $\mathbf{C} \in \mathbb{R}^{M\times M}$ is a matrix with all its entries equal to the maximum element of $(\mathbf{H}/2 + \mathbf{H}^{\mathrm{T}}/2 + \eta\mathbf{I})$. It is easy

Baselines	NF	TPF	RED	Ours
B1+B2+B3	3.900	3.854∼3.884	**3.919**	**3.919**
B1+B2+B4	3.822	3.626∼3.876	3.920	**3.922**
B1+B3+B4	3.865	3.626∼3.884	3.927	**3.930**
B2+B3+B4	3.893	3.629∼3.861	3.923	**3.926**
B1+B2+B3+B4	3.907	3.626∼3.884	**3.938**	**3.938**

Table 1. The performance comparison of different fusion methods on the Ukbench dataset.

to see that $\boldsymbol{\beta}^{\mathrm{T}}\mathbf{C}\boldsymbol{\beta}$ is a constant. Then, Eq. (9) is tranformed into

$$\begin{aligned} &\max_{\boldsymbol{\beta}} \boldsymbol{\beta}^{\mathrm{T}}(-\mathbf{H}/2 - \mathbf{H}^{\mathrm{T}}/2 - \eta\mathbf{I} + C)\boldsymbol{\beta} \\ \Leftrightarrow &\max_{\boldsymbol{\beta}} \boldsymbol{\beta}^{\mathrm{T}}\bar{\mathbf{H}}\boldsymbol{\beta}, \end{aligned} \tag{10}$$

which is equivalent to Eq. (21). The proof is complete. □

2. Experiment on Ukbench

Ukbench dataset [2] is a classical and representative benchmark for image retrieval, which is composed of $10,200$ images. The whole dataset has $2,550$ categories with 4 images per category. Each image is used in turn as a query. The performance is measured by the average recall of the top-4 ranked images, referred as N-S score (maximum is 4).

In recent years, the performance on the Ukbench dataset has gradually gotten saturated. Therefore, we do not include the performance comparison in the main manuscript. As can be drawn from Table 1, compared with RED [1], the proposed UED achieves better performance in three settings and the same performance in two settings.

References

[1] S. Bai, Z. Zhou, J. Wang, X. Bai, L. J. Latecki, and Q. Tian. Ensemble diffusion for retrieval. In *ICCV*, pages 774–783, 2017. 2

[2] D. Nistér and H. Stewénius. Scalable recognition with a vocabulary tree. In *CVPR*, pages 2161–2168, 2006. 2

[3] M. Pelillo. Replicator equations, maximal cliques, and graph isomorphism. In *NIPS*, pages 550–556, 1999. 1