# Super-resolution Person Re-identification with Semi-coupled Low-rank Discriminant Dictionary Learning

Xiao-Yuan Jing[1,3], Xiaoke Zhu[1], Fei Wu[1,3], Xinge You[2], Qinglong Liu[1], Dong Yue[3], Ruimin Hu[4], Baowen Xu[1]
[1]State Key Laboratory of Software Engineering, School of Computer, Wuhan University, China
[2]School of Electronic Information and Communications, Huazhong University of Science and Technology, China
[3]College of Automation, Nanjing University of Posts and Telecommunications, China
[4]National Engineering Research Center for Multimedia Software, School of Computer, Wuhan University, China

Person re-identification [2] has been widely studied due to its importance in surveillance and forensics applications. In practice, gallery images are high-resolution (HR) while probe images are usually low-resolution (LR) in the identification scenarios with large variation of illumination, weather or quality of cameras. Person re-identification in this kind of scenarios, which we call super-resolution (SR) person re-identification, has not been well studied.

Motivated by dictionary learning based SR restoration works [5], we propose a semi-coupled low-rank discriminant dictionary learning ($SLD^2L$) approach for SR person re-identification in this paper. Specifically, assume that $C_A$ is a HR pedestrian image set from camera A and $C_B$ is a LR pedestrian image set from camera B, we aim to learn a pair of HR and LR dictionaries and a mapping function between features of HR and LR images, such that the features of LR images in $C_B$ can be converted into discriminating HR features.

To this end, we firstly generate the LR version of $C_A$ by performing down-sampling and smoothing operations, which has the same resolution as $C_B$ and is denoted by $C_A'$. Then we exploit semi-coupled dictionary learning (DL) to learn a pair of HR and LR dictionaries and a mapping matrix between the corresponding features of $C_A$ and $C_A'$. To ensure that the learned dictionaries and mapping matrix have favorable discriminative capability, we require that HR features of images in $C_B$, which are reconstructed using the learned dictionary pair and mapping matrix, should be close to the features of images from the same person in $C_A$, but far away from the features of images from different persons in $C_A$.

In practice, low resolution has different influences on different patches, e.g., patches with pure color suffer little influence, while patches with complex texture suffer more influence. Therefore, learning a common mapping function is not enough to catch all the relationships. Intuitively, we can divide images into patches and group patches into several clusters, and then a pair of HR and LR sub-dictionaries and a more stable mapping function can be learned for each cluster. In this paper, we group patches in $C_A'$ and $C_B$ using K-means algorithm according to the similarity of patch features. Then, the patches in $C_A$ are grouped according to clustering results of the corresponding patches in $C_A'$. We require that each cluster-specific sub-dictionary has good representation ability to the patches from the associated cluster but poor representation ability for other clusters. Denote by $D_H^i$ and $D_L^i$ the HR and LR sub-dictionaries of the $i^{th}$ cluster, respectively. And $V_i$ denotes the mapping of the $i^{th}$ cluster. By separately combining HR and LR sub-dictionaries, we can obtain the structured HR and LR dictionaries, namely $D_H = [D_H^1, D_H^2, ..., D_H^c]$ and $D_L = [D_L^1, D_L^2, ..., D_L^c]$, where $c$ is the number of clusters.

To ensure that the learned sub-dictionary pairs can well characterize the intrinsic feature spaces of HR and LR images, the noises should be separated from patches in the learning process. Considering that patches from the same cluster are linearly correlated, we can employ low-rank matrix recovery technique to separate noises from patches [3, 4]. Figure 1 illustrates the overall flow of $SLD^2L$.

With the learned dictionaries and mapping matrices, features of LR probe images can be converted into discriminative HR features. Then SR person re-identification can be performed using the features of HR gallery images and the converted HR probe features.

Figure 2 reports the matching results of all compared methods on the VIPeR dataset [1] at sampling rate of $1/8$. The matching rates of all competing methods are significantly lower than those provided in the original papers. The reason is that low resolution results in the loss of useful information and these methods cannot work well in this scenario. The experimental results of $SLD^2L$ always outperform these related methods. Experimental results on the i-LIDS and PRID datasets also demonstrate the effectiveness of the proposed approach for SR person re-identification problem.

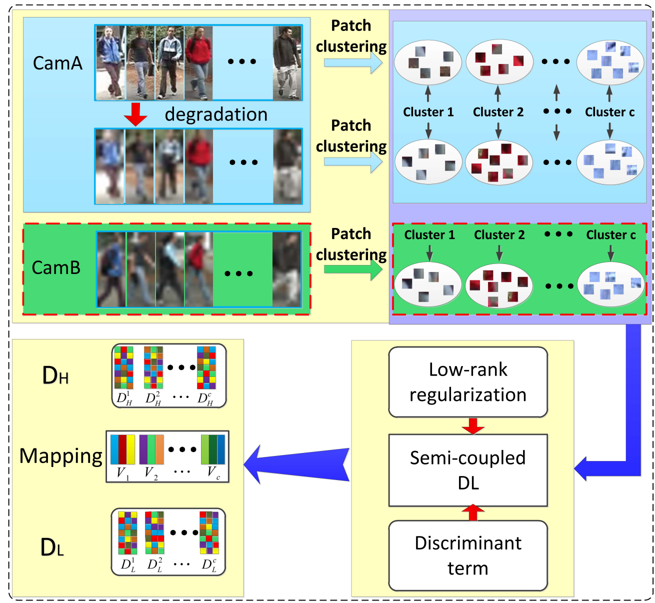


Figure 1: The flowchart of $SLD^2L$.

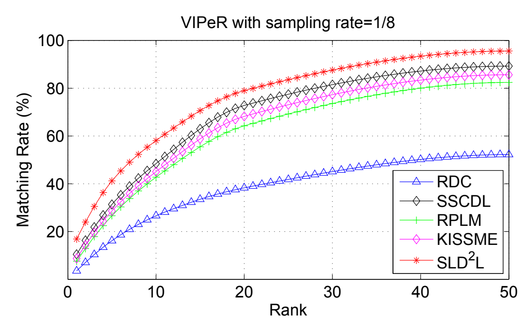


Figure 2: Results on the VIPeR dataset.

This is an extended abstract. The full paper is available at the Computer Vision Foundation webpage.

[1] Douglas Gray, Shane Brennan, and Hai Tao. Evaluating appearance models for recognition, reacquisition, and tracking. In *Performance Evaluation of Tracking and Surveillance, IEEE Workshop on*, 2007.

[2] Xiao Liu, Mingli Song, Dacheng Tao, Xingchen Zhou, Chun Chen, and Jiajun Bu. Semi-supervised coupled dictionary learning for person re-identification. In *CVPR, IEEE Conference on*, pages 3550–3557, 2014.

[3] Long Ma, Chunheng Wang, Baihua Xiao, and Wen Zhou. Sparse representation for face recognition based on discriminative low-rank dictionary learning. In *CVPR, IEEE Conference on*, pages 2586–2593, 2012.

[4] Houwen Peng, Bing Li, Rongrong Ji, Weiming Hu, Weihua Xiong, and Congyan Lang. Salient object detection via low-rank and structured sparse matrix decomposition. In *AAAI*, pages 796–802, 2013.

[5] Shenlong Wang, Lei Zhang, Yan Liang, and Quan Pan. Semi-coupled dictionary learning with applications to image super-resolution and photo-sketch synthesis. In *CVPR, IEEE Conference on*, pages 2216–2223, 2012.