Projective Weight-based Unsupervised Laplacian Graph Learning for Person Re-identification

Bin Gao
and Li Zhao
Key Laboratory of Underwater Acoustic
Signal Processing of Ministry of Education
Southeast University, China
Email: feimaxiao123@gmail.com, zhaoli@seu.edu.cn

Peng Lan
and Fenggang Sun
College of Information Science and Engineering
Shandong Agricultural University, China
Email: lanpeng@sdau.edu.cn, sunfg@sdau.edu.cn

Zhanke Yu
College of Communication Engineering
PLA Army Engineering University, China
Email: jackty_2004@163.com

Xiumei Pang
Department of Fundamental Research
PLA Army Engineering University, China
Email: 463169344@qq.com

***Abstract*—For unsupervised person re-identification, traditional Laplacian regularisation based dictionary learning methods encounter the fixed weight problem and thus impair the match rate. To address this limitation, we develop a novel unsupervised dictionary learning approach to learn a discriminative representation. The proposed approach takes the efficiency of l_2 graph regularization with a closed-form solution into account. Our approaches achieve very promising results on the challenging VIPeR dataset.**

I. Introduction

Machine learning methods [1], [2] can be applied to solve unsupervised person re-identification. Especially, the graph regularized dictionary learning (GRDL) [3] is one of the most representative approaches by considering the Laplacian graph regularization to fully explore the cross-view identity-discriminative information. To reduce the high algorithm complexity and enhance the weight learning efficiency of GRDL, we propose a new method which share the same superiority due to the Lapacian regularized term but with notable differences.

An efficient l_2 graph dictionary learning approach (GRDL2) is proposed to speed up the convergence rate of GRDL. Compared with the iterative feature sign search algorithm in GRDL, the proposed GRDL2 has a closed-form solution, which is not only faster than GRDL, but also achieves better rank-1 recognition rate. We evaluate the performance of GRDL2 and related methods on the public VIPeR dataset [4]. Extensive experimental results demonstrate that our proposed approaches outperform the state-of-the-art methods.

II. Laplacian Graph Learning

Let $\mathbf{X} = [\mathbf{X}^a, \mathbf{X}^b]$ denote the training set, where $\mathbf{X}^a = [\mathbf{x}_1^a, \mathbf{x}_2^a, \cdots, \mathbf{x}_{m_1}^a] \in R^{d\times m_1}$ and $\mathbf{X}^b = [\mathbf{x}_1^b, \mathbf{x}_2^b, \cdots, \mathbf{x}_{m_2}^b] \in R^{d\times m_2}$ are collected under two different views, respectively. Adopting a dictionary learning model, for the dataset $\mathbf{X} \in R^{d\times m}$ with $m = m_1 + m_2$, the target is to learn a shared dictionary $\mathbf{D}$, with which each d-dimensional feature vector can be projected into a lower k-dimensional subspace $\mathbf{Y} \in R^{k\times m}$. The subspace is spanned by the k dictionary atoms (columns of $\mathbf{D} \in R^{d\times k}$) and the dictionary learning problem is formulated as

$$\min_{\mathbf{D},\mathbf{Y}} \|\mathbf{X} - \mathbf{D}\mathbf{Y}\|_F^2 + \lambda_1 \|\mathbf{Y}\|_1 \tag{1}$$

To enhance the discriminativity of unsupervised dictionary learning for person re-identification, Kodirov *et al.* [3] introduced a Lapacian regularization term to preserve the local manifold:

$$\begin{gathered}\min_{\mathbf{D},\mathbf{Y}} \|\mathbf{X} - \mathbf{D}\mathbf{Y}\|_F^2 + \lambda_1 \|\mathbf{Y}\|_1 + \frac{\lambda_2}{2}\sum_{i,j}^{m} w_{ij}\|\mathbf{y}_i - \mathbf{y}_j\|_2^2, \\ \text{s.t.} \quad \|\mathbf{d}_i\|_2^2 \le 1, i = 1, \cdots, k\end{gathered} \tag{2}$$

where w_{ij} is the similarity weight between $\mathbf{x}_i^a$ and $\mathbf{x}_j^b$. However, there exist at least one challenge for conventional Lapalacian graph based methods. That is, the l_1 norm term $\lambda_1\|\mathbf{Y}\|_1$ in (2) is hard to cope with, especially when combined with the Laplacian regularized term.

To mitigate the impact of above problems, we propose GRDL2 approach as follows:

$$\begin{gathered}\min_{\mathbf{D},\mathbf{Y}} \|\mathbf{X} - \mathbf{D}\mathbf{Y}\|_F^2 + \lambda_1 \|\mathbf{Y}\|_F^2 + \frac{\lambda_2}{2}\sum_{i,j}^{m} w_{ij}\|\mathbf{y}_i - \mathbf{y}_j\|_2^2, \\ \text{s.t.} \quad \|\mathbf{d}_i\|_2^2 \le 1, i = 1, \cdots, k\end{gathered} \tag{3}$$

The main contribution is that we boost the coding efficiency by using l_2 norm instead of l_1 for coding coefficients. And we derive a closed-form solution for the Laplacian regularized model (3). The experimental parts demonstrates that the new model can achieve better Rank 1 matching rate with much less computational cost.

III. Proposed Approach

With the aid of the alternating direction method of multipliers (ADMM) [5], we divide the primal learning problem (3) into four sub-problems: fix three variables in $\{\mathbf{D}, \mathbf{Y}\}$ and solve the remaining one. These four sub-problems are solved alternatively and iteratively, and optimization scheme stops at a stationary point to obtain the optimal solutions of $\mathbf{D}, \mathbf{Y}, \mathbf{W}$ and $\mathbf{F}$, which are represented below.

Step1. Fix $\mathbf{Y}$, and update $\mathbf{D}$: (3) degenerates into

$$\min_{\mathbf{D}} \|\mathbf{X} - \mathbf{DY}\|_F^2 \quad \text{s.t.} \quad \|\mathbf{d}_i\|_2^2 \leq 1, i = 1, \cdots, k \tag{4}$$

which can be optimised by the Lagrange dual method and the analytical solution of $\mathbf{D}$ can be computed as: $\mathbf{D} = \mathbf{XY}^T(\mathbf{YY}^T + \mathbf{\Lambda})$, where $\mathbf{\Lambda}$ is a diagonal matrix constructed by all the dual variables.

Step2. Fix $\mathbf{D}$, and update $\mathbf{Y}$: (3) now becomes

$$\min_{\mathbf{Y}} \|\mathbf{X} - \mathbf{DY}\|_F^2 + \lambda_1\|\mathbf{Y}\|_F^2 + \frac{\lambda_2}{2}\sum_{i,j}^{m} w_{ij}\|\mathbf{y}_i - \mathbf{y}_j\|_2^2 \tag{5}$$

Define $q_i = \sum_{j=1}^{m} w_{ij}$, and $\mathbf{Q} = diag(q_1, \cdots, q_m)$, we have

$$\frac{1}{2}\sum_{i,j}^{m} w_{ij}\|\mathbf{y}_i - \mathbf{y}_j\|_2^2 = Tr(\mathbf{YLY}^T)$$

where $\mathbf{L} = (\mathbf{Q} - \mathbf{W})$ is the Laplacian graph. Thus, we can rewrite problem (5) as

$$\min_{\mathbf{Y}} \|\mathbf{X} - \mathbf{DY}\|_F^2 + \lambda_1\|\mathbf{Y}\|_F^2 + \lambda_2 Tr(\mathbf{YLY}^T) \tag{6}$$

For ease of representation, we denote $L(\mathbf{Y}) = \|\mathbf{X} - \mathbf{DY}\|_F^2 + \lambda_1\|\mathbf{Y}\|_F^2 + \lambda_2 Tr(\mathbf{YLY}^T)$ and we get

$$L(\mathbf{y}_i) = \|\mathbf{x}_i - \mathbf{Dy}_i\|_F^2 + \lambda_1\|\mathbf{y}_i\|_2^2 + \lambda_2 l_{ii}\mathbf{y}_i^T\mathbf{y}_i + \mathbf{y}_i^T\mathbf{h}_i$$

with $\mathbf{h}_i = 2\lambda_2\sum_{j\neq i} l_{ij}\mathbf{y}_j$. Taking the derivative of $L(\mathbf{y}_i)$ w.r.t $\mathbf{y}_i$, and setting the derivative to zero, we have

$$\mathbf{y}_i = (\mathbf{D}^T\mathbf{D} + 2\lambda_2 l_{ii}\mathbf{I})^{-1}(\mathbf{D}^T\mathbf{x}_i - 2\lambda_2\sum_{k\neq i}\mathbf{y}_k l_{ik})$$

where l_{ik} is the (i, k) entry of $\mathbf{L}$, $\mathbf{I}$ is the identity matrix. $\mathbf{x}_i$ and $\mathbf{y}_k$ are the ith column of $\mathbf{X}$ and kth column of $\mathbf{Y}$, respectively.

IV. Experimental results

The GRDL2 is validated on the VIPeR dataset, which contains images of 632 pedestrian pairs under two camera views with different viewpoints, with only one image per person in each view and all the images normalized to 128×48 pixels. Suffering from significant viewpoint changes, pose variation, and illumination difference across cameras, it is one of the most challenging databases for person re-identification. The procedure is repeated 10 times to get an average performance. We compare our approaches with two state-of-the-art methods, i.e., eSDC [6] and GRDL.

Fig. 1 shows the rank-1 recognition rate comparisons of the four methods. From the results one can easily confirm that GRDL2 significantly outperforms GRDL by improving the rank-1 accuracy from 26.9% to 32.3% on the VIPeR dataset. As a byproduct, GRDL2 has a much faster convergence rate which is verified in our open source code, which is released on public for further reproducible research (https://github/gaobingaobingaobin/person-ReID3). Thus, the proposed GRDL2 is superior to GRDL.

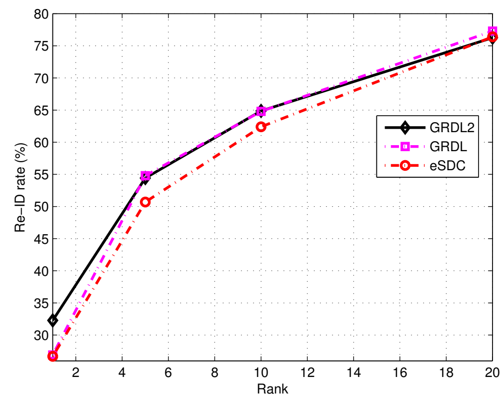

Fig. 1: CMC curves on VIPeR dataset.

V. Conclusion

A novel graph regularized dictionary learning method, i.e., GRDL2, for unsupervised person re-identification has been introduced and validated to reach excellent performance. GRDL2 can be viewed as an improved GRDL version by using l_2 regularization with closed-form solution. Our proposed GRDL2 is very fast for learning, which are important for real applications. Experimental results demonstrate their efficiency on unsupervised person re-identification.

Acknowledgment

This work was supported by the National Natural Science Foundation of China (Grants 61673108, 61571106 and 71401176) and the Natural Science Foundation of Jiangsu Province (Grant BK20170765).

References

[1] L. Xiao, Y. Li, G. Han, G. Liu, W. Zhuang, "PHY-layer Spoofing Detection with Reinforcement Learning in Wireless Networks," *IEEE Trans. Veh. Technol.*, vol. 65, no. 12, pp. 10037–10047, Dec. 2016.

[2] L. Xiao, Y. Li, G. Han, H. Dai, and H. V. Poor, "A Secure Mobile Crowdsensing Game with Deep Reinforcement Learning," *IEEE Trans. Inf. Forensic Secur.*, vol. 136, no. 1, pp. 35-47, Aug. 2017.

[3] E. Kodirov, T. Xiang, S. Gong, "Dictionary learning with iterative laplacian regularisation for unsupervised person re-identification," in *Proc. of the British Machine Vision Conference (BMVC)*, Swansea, Sep. 2015, pp. 44.1–44.12.

[4] D. Gray, S. Brennan, and H. Tao, "Evaluating appearance models for recognition, reacquisition, and tracking," *IEEE Int. Workshop on Performance Evaluation for Tracking and Surveillance (PETS)*, Nov. 2007, pp. 41–47.

[5] B. Gao and F. Ma, "Symmetric alternating direction method with indefinite proximal regularization for linearly constrained convex optimization," *J. Optim. Theory Appl.*, vol. 176, no. 1, pp. 178–204, Jan. 2018.

[6] R. Zhao, W. Ouyang, and X. Wang, "Unsupervised salience learning for person re-identification," in *IEEE Conf. Computer Vision and Pattern Recognition (CVPR)*, Portland, USA, Jun. 2013, pp. 3586–3593.