REGULARIZATION IN METRIC LEARNING FOR PERSON RE-IDENTIFICATION

Jianlou Si, Honggang Zhang, and Chun-Guang Li

School of Information and Communication Engineering
Beijing University of Posts and Telecommunications, Beijing 100876, P. R. China

ABSTRACT

Metric learning plays a critical role in person re-identification problem. Unfortunately, due to the small size of training data, the metric learning used in this scenario suffers from overfitting which leads to degenerated performance. In this paper, we investigate the effect of regularization in metric learning for person re-identification. Concretely we formulate the distance function from three perspectives and hence present four different regularized metric learning methods. Experiments on two popular benchmark data sets VIPeR and CUHK01 validate the effectiveness of our proposed regularization approaches.

Index Terms— Person Re-identification, Metric Learning, Regularization

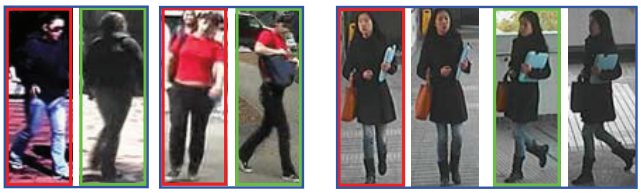
1. INTRODUCTION

Nowadays, video surveillance networks are widely deployed in the areas of city, highway, airports, train stations, and etc. The massive captured surveillance videos contain important cues for the public security. Therefore, automatically verifying the identity of a person from non-overlapping surveillance camera views becomes one of the most critical tasks in video content analysis. This is the problem termed as *Person Re-Identification* (Re-ID) [1]. Although it is easy to understand that the process of Re-ID is to match a probe image against a gallery set, it is indeed quite challenging, due to the large variations in illumination and pose, and complicated background cluttering. We display some paired images in Fig. 1 from two benchmark data sets to show the difficulties in Re-ID problem.

2. RELATED WORK

Over the past a few years, Re-ID problem has received a large amount of attention and dozens of methods have been proposed [4, 5, 1, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19].

This work was partially supported by National Natural Science Foundation of China under Grant Nos. 61175011, 61273217, and 61171193, the 111 project under Grant No.B08004, and the Scientific Research Foundation for the Returned Overseas Chinese Scholars, Ministry of Education of China.



(a) (b)

Fig. 1. Paired images in (a) and (b) are from data sets VIPeR [2] and CUHK01 [3], respectively. Images boxed in red and green are paired candidates need to re-identify.

Roughly, the existing methods can be divided into two categories: *feature representation* based methods [6, 7, 8, 9, 10, 11, 12, 13, 19] and *metric learning* based methods [14, 15, 16, 17, 4, 18, 5].

Extracting features with powerful representative ability is one of the critical components in Re-ID problem. It has been demonstrated that incorporating body parts layout information into feature extraction is very effective for Re-ID. For example, [6, 20] demonstrated that constructing the symmetry-driven accumulation features from body parts and weighting them according to the body symmetric axis was useful, [19] also gave an alternative to extract and weight features via Region-of-Interest, [8, 9, 12] shown that introducing the saliency information to measure the discriminative and representative ability of local features was effective, and [13, 21] validated the usefulness of incorporating the semantic part-based appearance changes. While the aforementioned sophisticated features improving the accuracy in a degree, they all suffered from the requirement of rich priori knowledge of the specific problem. Besides, it is usually difficult to find a suitable metric for the hand-crafted features and hence lead to degenerated performance. Therefore, finding a distance (or similarity) measure which is suitable for the features is another critical component in Re-ID problem. This issue is addressed by *metric learning*, e.g.[22, 23], which has been applied to Re-ID problem recently [14, 4], [15], [17], [16], [5]. The basic idea is to learn a task-specific metric (space) from training data such that the distance between images of the same identity is smaller than others. Zheng et al. [14, 4] proposed to formulate Re-ID problem as a relative distance learning task in which an optimal metric was learnt for pairwise images.

Mignon [15] proposed to learn the distance metric by optimizing a decision function and Li et al. [17] gave a large-margin solution. Kostinger et al. [16] proposed to derive the metric learning from a statistical inference perspective. Xiong et al. [5] presented a kernelized extension for the common metric learning methods, and gained performance improvements. In Re-ID problem, however, the paired training samples are far from sufficient. Because of this, the aforementioned metric learning formulations suffer from serious *over-fitting* which significantly degenerates the generalization ability.

In this paper, we propose to accommodate the regularization strategy to enhance the generalization ability of metric learning methods for Re-ID problem. To be concrete, we investigate the effects of the regularized metric learning with three different metric function formulations. Experiments conducted on two benchmark data sets VIPeR and CUHK01 confirm the effectiveness of our proposals.

3. OUR PROPOSAL: REGULARIZED METRIC LEARNING

In this section, we introduce three different formulations of the metric function and then present our proposal – regularized metric learning. For clarity, we give some symbol notations here: the bold lowercase $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^d$ are the feature vectors extracted from images, and $\ell_i \in \{1, 2, 3, ..., c\}$ indicates the identity of i-th feature vector, and each pair of feature vectors $(\boldsymbol{x}_i, \boldsymbol{x}_j)$ or $(\boldsymbol{x}_i, \boldsymbol{y}_j)$ has a matching score $g_{ij} = +1$ if $\ell_i = \ell_j$ otherwise $g_{ij} = -1$.

3.1. Different Formulations for Metric Function

Metric learning is to learn, by using training data, a task-specific metric function, which can be formulated as follows:

$$d^2_{\mathbf{M}}(\boldsymbol{x}, \boldsymbol{y}) = (\boldsymbol{x} - \boldsymbol{y})^T \mathbf{M} (\boldsymbol{x} - \boldsymbol{y}), \tag{1}$$

where $\mathbf{M} \in \mathbb{S}^d_+$ in which $\mathbb{S}^d_+$ is the set of symmetry positive semi-definite (PSD) matrices. Note that with the constraint $\mathbf{M} \in \mathbb{S}^d_+$, we can express $\mathbf{M}$ as $\mathbf{L}^T\mathbf{L}$, where $\mathbf{L} \in \mathbb{R}^{k \times d}$ is of rank k, and thus reformulate the metric function as follows:

$$d^2_{\mathbf{M}}(\boldsymbol{x}, \boldsymbol{y}) = (\mathbf{L}\boldsymbol{x} - \mathbf{L}\boldsymbol{y})^T (\mathbf{L}\boldsymbol{x} - \mathbf{L}\boldsymbol{y}). \tag{2}$$

By using the reformulation (2), we observe the connection between supervised metric leaning and supervised linear dimensionality reduction.

Considering of that the feature vectors $\boldsymbol{x}$ and $\boldsymbol{y}$ for Re-ID are extracted in images from different cameras, we further modify the formulation (2) to tolerant camera-specific projections as follows:

$$d^2_{\mathbf{M}}(\boldsymbol{x}, \boldsymbol{y}) = (\mathbf{L}\boldsymbol{x} - \mathbf{H}\boldsymbol{y})^T (\mathbf{L}\boldsymbol{x} - \mathbf{H}\boldsymbol{y}), \tag{3}$$

where $\mathbf{L}, \mathbf{H} \in \mathbb{R}^{k \times d}$.

The three formulations in (1), (2) and (3) are used in three different metric learning approaches. In the next subsection, we investigate regularization technique for each of them.

3.2. Regularized Metric Learning

Regularization is an effective strategy to tackle with an ill-posed problem. In this paper, we attempt to adopt the *regularization* into metric learning methods for Re-ID. To be specific, we modify three different metric learning approaches – Large Margin Nearest Neighbors (LMNN) [22], Linear Discriminant Analysis (LDA), and Decision Function Learning (DFL) [15, 17] – into their regularized formulations, and then solve each of them efficiently.

3.2.1. Regularized Large Margin Nearest Neighbors

Define a target neighbors set for each input $\boldsymbol{x}$ as those data points that are desired to be close to $\boldsymbol{x}$, and an impostors set as those points are closer than target neighbors but labeled differently. LMNN seeks a PSD matrix $\mathbf{M}$ by pulling target neighbors together and pushing imposters away simultaneously, which can be formulated as follows:

$$\begin{aligned} \min_{\mathbf{M} \in \mathbb{S}^d_+} (1-\mu) \sum_{i, j \rightsquigarrow i} d^2_{\mathbf{M}}(\boldsymbol{x}_i, \boldsymbol{x}_j) + \mu \sum_{i, j \rightsquigarrow i, l} (1 - g_{il}) \xi_{ijl}, \\ \text{s.t. } d^2_{\mathbf{M}}(\boldsymbol{x}_i, \boldsymbol{x}_l) - d^2_{\mathbf{M}}(\boldsymbol{x}_i, \boldsymbol{x}_j) \geq 1 - \xi_{ijl}, \;\; \xi_{ijl} \geq 0, \end{aligned} \tag{4}$$

where $0 \leq \mu \leq 1$ is the weighting parameter to balance pulling and pushing effects, $j \rightsquigarrow i$ denotes that $\boldsymbol{x}_j$ belongs to the target neighbors of $\boldsymbol{x}_i$, and $d^2_{\mathbf{M}}(\cdot, \cdot)$ is described as Eq. (1).

While problem (4) could be solved with Semi-Definite Programming (SDP) solver, the lack of training data makes LMNN prone to over-fitting. To this end, we propose to add a nuclear norm $\|\mathbf{M}\|_*$ as a regularizer to problem (4) and hence lead to a nuclear norm Regularized LMNN (nuLMNN) problem as follows:

$$\begin{aligned} \min_{\mathbf{M} \in \mathbb{S}^d_+} \sum_{i, j \rightsquigarrow i} (1-\mu) d^2_{\mathbf{M}}(\boldsymbol{x}_i, \boldsymbol{x}_j) + \sum_{i, j \rightsquigarrow i, l} \mu (1 - g_{il}) \xi_{ijl} + \lambda \|\mathbf{M}\|_* \\ \text{s.t. } d^2_{\mathbf{M}}(\boldsymbol{x}_i, \boldsymbol{x}_l) - d^2_{\mathbf{M}}(\boldsymbol{x}_i, \boldsymbol{x}_j) \geq 1 - \xi_{ijl}, \;\; \xi_{ijl} \geq 0, \end{aligned} \tag{5}$$

where λ is regularization parameter. Note that for a PSD matrix $\mathbf{M}$, we have $\mathrm{tr}(\mathbf{M}) = \|\mathbf{M}\|_*$ and thus we also term problem (5) as trace regularized LMNN (trLMNN). Another way to regularize problem (4) is by using a LogDet divergence term [23], which is the differential relative entropy between two multivariate Gaussians under constraints on two covariance matrices, as follows:

$$D_{ld}(\mathbf{M}, \mathbf{M}_0) = \mathrm{tr}\left(\mathbf{M}\mathbf{M}_0^{-1}\right) - \log\det\left(\mathbf{M}\mathbf{M}_0^{-1}\right) - d, \tag{6}$$

where $\mathbf{M}, \mathbf{M}_0 \in \mathbb{S}^d_+$, and $\mathbf{M}_0$ is usually set as an identity matrix $\mathbf{I}$. Similarly, by adding $\lambda D_{ld}(\mathbf{M}, \mathbf{M}_0)$ into problem (4), we obtain a LogDet regularized LMNN (ldLMNN). Note

that, to solve the regularized LMNN problems (trLMNN and ldLMNN), we need only modify the definitions of the gradient of the objective function in trLMNN and ldLMNN, respectively as follows:

$$\mathbf{G}_t = (1-\mu)\sum_{i,j\rightsquigarrow i}\mathbf{C}_{ij} + \mu\sum_{i,j\rightsquigarrow i,l}(\mathbf{C}_{ij}-\mathbf{C}_{il}) + \lambda\mathbf{I} \quad (7)$$

$$\mathbf{G}_t = (1-\mu)\sum_{i,j\rightsquigarrow i}\mathbf{C}_{ij} + \mu\sum_{i,j\rightsquigarrow i,l}(\mathbf{C}_{ij}-\mathbf{C}_{il}) + \lambda(\mathbf{I} - \det(\mathbf{M}_t)\mathbf{M}_t^{-1}) \quad (8)$$

where $\mathbf{C}_{ij} = (\boldsymbol{x}_i - \boldsymbol{x}_j)(\boldsymbol{x}_i - \boldsymbol{x}_j)^T$. Since matrix $\mathbf{M}_t$ may not be invertible, we instead use $\hat{\mathbf{M}}_t = (1-\alpha)\mathbf{M}_t + \frac{\alpha}{N}\operatorname{tr}(\mathbf{M}_t)\mathbf{I}$ where $0 \leq \alpha \leq 1$ and N is the number of samples.

3.2.2. Regularized Linear Discriminant Analysis

LDA is the most popular supervised dimensionality reduction method, and has many variations , e.g. Local Fisher Discriminant Analysis (LFDA) [24] and Marginal Fisher Analysis (MFA) [5]. One of the advantages of these methods is that the optimal solution can be found by solving a generalized eigenvalue problem. The objective optimization of these methods are usually defined as follows:

$$\max_{\mathbf{L}\in\mathbb{R}^{k\times d}} \quad \frac{\operatorname{tr}(\mathbf{L}\mathbf{S}^b\mathbf{L}^T)}{\operatorname{tr}(\mathbf{L}\mathbf{S}^w\mathbf{L}^T)}, \quad (9)$$

where $\mathbf{S}^w \in \mathbb{R}^{d\times d}$ is the intra-class scatter matrix and $\mathbf{S}^b \in \mathbb{R}^{d\times d}$ is inter-class scatter matrix. For the objective function as above, the intra-class scatter matrix $\mathbf{S}^w$ should be invertible to guarantee the feasibility of optimization. However, if the sample size N is smaller than the feature dimensionality d, $\mathbf{S}^w$ will be singular. In this case, we have a regularized term instead of $\mathbf{S}^w$ as follows

$$\hat{\mathbf{S}}^w = (1-\alpha)\mathbf{S}^w + \frac{\alpha}{N}\operatorname{tr}(\mathbf{S}^w)\mathbf{I}, \quad (10)$$

where $0 \leq \alpha \leq 1$ and N is the number of samples. We call it stable Linear Discriminant Analysis (sLDA).

3.2.3. Regularized Decision Function Learning

For a linearly separable classification problem, one can find an optimal hyperplane $f(\cdot) = 0$ so that all positive samples lie on one side and all negative samples lie on another side. The ultimate goal of Re-ID is to distinguish whether input pairwise example $(\boldsymbol{x}_i, \boldsymbol{x}_j)$ represents the same individual or not, so it can also be formulated to learn a decision hyperplane $f(\cdot,\cdot) = 0$. The decision function is formulated as $f(\boldsymbol{x}_i, \boldsymbol{x}_j) = d^2_{\mathbf{M}}(\boldsymbol{x}_i, \boldsymbol{x}_j) - t$, where t is the bias.

The previous metric learning approaches for Re-ID usually assume that elements of input pairs come from the same feature space, which is denoted as $(\boldsymbol{x}_i, \boldsymbol{x}_j)$. However, the pairwise samples in Re-ID are captured from different non-overlapping cameras, which can be denoted as $(\boldsymbol{x}_i, \boldsymbol{y}_j)$. For this reason, the projection matrices, i.e. $\mathbf{L}, \mathbf{H} \in \mathbb{R}^{k\times d}$, should be different. To this end, we reformulate the metric function with camera-specific projections as in Eq. (3).

In this paper, we consider $\boldsymbol{x}_i$ and $\boldsymbol{y}_j$ come from different feature spaces, and t takes a quadratic form, (which is slightly different with [17]), as follows:

$$\begin{aligned} t(\boldsymbol{x}_i, \boldsymbol{y}_j) = &\frac{1}{2}\boldsymbol{x}_i^T\tilde{\mathbf{A}}\boldsymbol{x}_i + \frac{1}{2}\boldsymbol{y}_j^T\tilde{\mathbf{B}}\boldsymbol{y}_j + \boldsymbol{x}_i^T\tilde{\mathbf{C}}\boldsymbol{y}_j \\ &+ \boldsymbol{w}^T(\boldsymbol{x}_i + \boldsymbol{y}_j) + \boldsymbol{w}_0, \end{aligned} \quad (11)$$

where $\tilde{\mathbf{A}} \in \mathbb{S}^d$, $\tilde{\mathbf{B}} \in \mathbb{S}^d$, $\tilde{\mathbf{C}} \in \mathbb{R}^{d\times d}, \boldsymbol{w} \in \mathbb{R}^d, \boldsymbol{w}_0 \in \mathbb{R}$, and $\mathbb{S}^d$ are symmetry matrices. Now the decision function can be rewritten as follows

$$\begin{aligned} &f(\boldsymbol{x}_i, \boldsymbol{y}_j) = d^2_{\mathbf{M}}(\boldsymbol{x}_i, \boldsymbol{y}_j) - t(\boldsymbol{x}_i, \boldsymbol{y}_j) \\ &= \frac{1}{2}\boldsymbol{x}_i^T\mathbf{A}\boldsymbol{x}_i + \frac{1}{2}\boldsymbol{y}_j^T\mathbf{B}\boldsymbol{y}_j - \boldsymbol{x}_i^T\mathbf{C}\boldsymbol{y}_j - \boldsymbol{w}^T(\boldsymbol{x}_i + \boldsymbol{y}_j) - \boldsymbol{w}_0, \end{aligned} \quad (12)$$

where $\mathbf{A} = 2\mathbf{L}^T\mathbf{L} - \tilde{\mathbf{A}}$, $\mathbf{B} = 2\mathbf{H}^T\mathbf{H} - \tilde{\mathbf{B}}$, and $\mathbf{C} = 2\mathbf{L}^T\mathbf{H} + \tilde{\mathbf{C}}$. It is easy to show that $\mathbf{A} \in \mathbb{S}^d$ and $\mathbf{B} \in \mathbb{S}^d$ are both symmetry matrices and $\mathbf{C} \in \mathbb{R}^{d\times d}$ is not necessary to be symmetric.

The optimal hyperplane can be solved via a regularized Decision Function Learning (rDFL), i.e., minimizing the loss function with a regularizatioin term as follows:

$$\min_{\boldsymbol{\Theta}} \quad E(\boldsymbol{\Theta}) = \sum_i\sum_j h_\beta\left(g_{ij}f(\boldsymbol{x}_i, \boldsymbol{y}_j)\right) + R(\boldsymbol{\Theta}), \quad (13)$$

where $\boldsymbol{\Theta} = \{\mathbf{A}, \mathbf{B}, \mathbf{C}, \boldsymbol{w}\}$, $R(\boldsymbol{\Theta}) = \lambda_1||\mathbf{A}||_F^2 + \lambda_2||\mathbf{B}||_F^2 + \lambda_3||\mathbf{C}||_F^2$, and $h_\beta(x) = \frac{1}{\beta}\log(1 + e^{\beta x})$ is the smooth approximation of the hinge loss. This optimization problem can be solved via a gradient descent, which minimizes with respect to $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$, and $\boldsymbol{w}$ alternately.

4. EXPERIMENTS

4.1. Experimental Settings

Feature Extraction. We extract the local features from sliding windows. We normalize each image into 128×48 and then slide a 16×16 window with 8 pixels overlapping. The 8-bins histogram and 3-dimensional color moments from the HSV and YUV color channels, as well as the 10-dimensional rotation invariant uniform Local Binary Patterns (LBP) from gray-scale image, are extracted from each patch. And then we obtain a 5700-dimensional feature vector by concatenating them together. Finally PCA is used to reduce the features into dimension 250.[1]

Data Set Preparing. We evaluate our proposed regularized approaches on two widely used benchmark data sets:

[1]Note that we reduce separately the features extracted from HSV and YUV channels and LBP features into dimensions 160, 40, 50 by PCA.

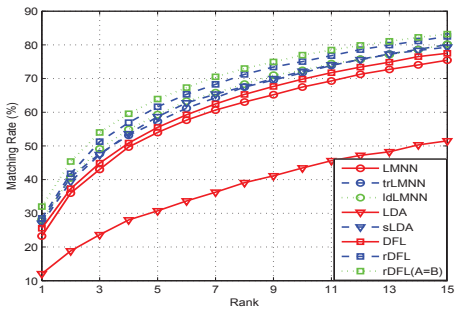


Fig. 2. Performance of proposed methods on VIPeR.

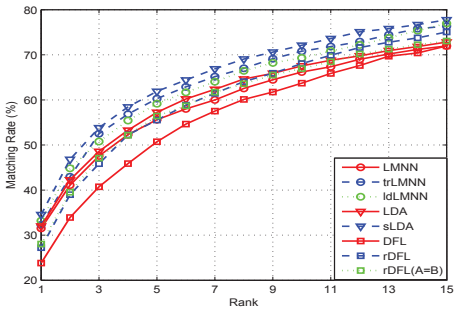


Fig. 3. Performance of proposed methods on CUHK01.

Table 1. Matching Rate on VIPeR.

Method	Rank1	Rank5	Rank10	Rank20
SCNCD[12]	**33.7%**	62.7%	74.8%	85%
SDALF[6]	19.9%	38.9%	49.4%	65.7%
SDC[9]	26.7%	50.7%	62.4%	76.4%
kMFA[5]	31.1%	**65.2%**	**79.6%**	**90.2%**
LFDA[5]	21.5%	49.6%	64.6%	79.1%
LADF[5]	30.1%	63.2%	**77.4%**	88.1%
LMNN	23.2%	54.0%	67.5%	80.8%
trLMNN	27.8%	57.3%	71.7%	84.0%
ldLMNN	28.2%	59.3%	72.4%	84.6%
LDA	12.1%	30.7%	43.5%	57.7%
sLDA	27.0%	58.7%	72.2%	83.8%
DFL	25.6%	55.5%	69.8%	81.9%
rDFL	28.5%	61.7%	75.0%	87.0%
rDFL(A=B)	**32.0%**	**63.9%**	77.0%	**88.6%**

VIPeR [2] and CUHK01 [3]. Some sample pairs are displayed in Fig. 1. Both of the two data sets contain images captured from two disjoint cameras with different illumination, low resolution and changing poses. Data set VIPeR contains 632 pairs of pedestrian images captured in outdoor academic environment. Each pair include one image from camera-A and one image from camera-B. Since trLMNN, ldLMNN, sLDA, and other metric learning methods, except for DFL, do not need the camera-specific projections, so we randomly choose one of the pairwise images to compose the gallery set and leave the other one as a probe. However, the camera labels are useful for DFL, so we regard images from camera-A as probe set and regard images from camera-B as gallery set for DFL and the regularized DFL. Data set CUHK01 contains 971 pedestrians images (which is a multi-shot to multi-shot data set). Each pedestrian has four images, including two captured from camera-A and two captured from camera-B. The partition protocol of gallery and probe set on this data set is the same as VIPeR, except that only one image from each camera was randomly chose for every pedestrian to comprise gallery or probe set.

Evaluation Protocol. We evaluate our proposed regularized metric learning methods on VIPeR and CUHK01, by calculating Cumulative Matching Characteristic (CMC) curve. We randomly split the data into 50% for training and 50% for testing, and record the average matching rate over 10 trials.

Parameter Settings. For trLMNN and ldLMNN, we set $\lambda = 10^4$. For sLDA and ldLMNN, we set $\alpha = 0.5$. And for rDFL, $\lambda_1 = 0.6, \lambda_2 = 0.6, \lambda_3 = 0.9$.

4.2. Results

We show the CMC curves of our proposed methods on data sets VIPeR and CUHK01 in Fig. 2 and Fig. 3, respectively. Note that, for DFL the pairwise samples consist of two images from different cameras, whereas the pairwise samples can be selected from any cameras (the same one or not) for other metric learning methods. Thus the number of training samples for DFL is much less than the number for others. While the lack of training sample lowers the matching rate of DFL, it still shows that regularization can improve the matching rate significantly. In particular, for matching rate on rank 1, the average improvement is about 4% of all these four regularized methods.

We also compare our approaches with three popular metric learning methods, i.e., kMFA [5], LFDA [24], LADF [17]. Since that our feature representation method is similar to [5], we directly cite the results from [5]. In addition, we also list other three state-of-art non-metric learning methods, i.e., SCNCD [12], SDALF [6], SDC [9]. The comparison is listed in Table 1. Notice that the methods with regularization consistently improve the Re-ID performance. In particular, for rank 1 matching rate on VIPeR, trLMNN and ldLMNN improve the rate by $3\% \sim 5\%$ from LMNN, sLDA gets 15% improvement from LDA. Furthermore, rDFL catches up with the best metric learning method in Re-ID at 32.0%.

5. CONCLUSION

In this paper, we investigated the effect of regularization in metric learning for person re-identification. Concretely we formulated the distance function from three perspectives and hence presented four different regularized metric learning methods. Experiments conducted on two benchmark data sets VIPeR and CUHK01 confirmed the effectiveness of our proposals.

6. REFERENCES

[1] Shaogang Gong and Tao Xiang, *Person Re-identification*, Springer, 2014.

[2] Douglas Gray, Shane Brennan, and Hai Tao, "Evaluating appearance models for recognition, reacquisition, and tracking," in *PETS*, 2007.

[3] Wei Li, Rui Zhao, and Xiaogang Wang, "Human reidentification with transferred metric learning," in *ACCV*, 2013.

[4] Wei-Shi Zheng, Shaogang Gong, and Tao Xiang, "Reidentification by relative distance comparison," *TPAMI*, vol. 35, no. 3, pp. 653–668, 2013.

[5] Fei Xiong, Mengran Gou, Octavia Camps, and Mario Sznaier, "Person re-identification using kernel-based metric learning methods," in *ECCV*, 2014.

[6] Michela Farenzena, Loris Bazzani, Alessandro Perina, Vittorio Murino, and Marco Cristani, "Person re-identification by symmetry-driven accumulation of local features," in *CVPR*, 2010.

[7] Igor Kviatkovsky, Amit Adam, and Ehud Rivlin, "Color invariants for person reidentification," *TPAMI*, vol. 35, no. 7, pp. 1622–1634, 2013.

[8] Rui Zhao, Wanli Ouyang, and Xiaogang Wang, "Person re-identification by salience matching," in *ICCV*, 2013.

[9] R. Zhao, W. L. Ouyang, and X. G. Wang, "Unsupervised salience learning for person re-identification," in *CVPR*, 2013.

[10] Yang Hu, Shengcai Liao, Zhen Lei, Dong Yi, and Stan Z. Li, "Exploring structural information and fusing multiple features for person re-identification," in *CVPR Workshop*, 2013.

[11] Rui Zhao, Wanli Ouyang, and Xiaogang Wang, "Learning mid-level filters for person re-identification," in *CVPR*, 2014.

[12] Yang Yang, Jimei Yang, Junjie Yan, Shengcai Liao, Dong Yi, and Stan Z Li, "Salient color names for person re-identification," in *ECCV*, 2014.

[13] Ziming Zhang, Yuting Chen, and Venkatesh Saligrama, "A novel visual word co-occurrence model for person re-identification," in *ECCV Workshop*, 2014.

[14] W. S. Zheng, S. G. Gong, and T. Xiang, "Person re-identification by probabilistic relative distance comparison," in *CVPR*, 2011.

[15] Alexis Mignon and Frdric Jurie, "Pcca: A new approach for distance learning from sparse pairwise constraints," in *CVPR*. 2012, IEEE.

[16] M Kostinger, Martin Hirzer, Paul Wohlhart, Peter M Roth, and Horst Bischof, "Large scale metric learning from equivalence constraints," in *CVPR*, 2012.

[17] Z. Li, S. Y. Chang, F. Liang, T. S. Huang, L. L. Cao, and J. R. Smith, "Learning locally-adaptive decision functions for person verification," in *CVPR*, 2013.

[18] Dong Yi, Zhen Lei, Shengcai Liao, and Stan Z Li, "Deep metric learning for person re-identification," in *ICPR*, 2014.

[19] Jianlou Si, Honggang Zhang, and Chun-Guang Li, "Person re-identification via region-of-interest based features," in *VCIP*, 2014.

[20] Loris Bazzani, Marco Cristani, and Vittorio Murino, "Symmetry-driven accumulation of local features for human characterization and re-identification," *CVIU*, vol. 117, no. 2, pp. 130–144, 2013.

[21] Ziming Zhang and Venkatesh Saligrama, "Person re-identification via structured prediction," *arXiv preprint arXiv:1406.4444*, 2014.

[22] Kilian Q Weinberger and Lawrence K Saul, "Distance metric learning for large margin nearest neighbor classification," *JMLR*, vol. 10, pp. 207–244, 2009.

[23] Jason V Davis, Brian Kulis, Prateek Jain, Suvrit Sra, and Inderjit S Dhillon, "Information-theoretic metric learning," in *ICML*, 2007.

[24] Sateesh Pedagadi, James Orwell, Sergio Velastin, and Boghos Boghossian, "Local fisher discriminant analysis for pedestrian re-identification," in *CVPR*, 2013.