Neural Processing Letters
https://doi.org/10.1007/s11063-019-10000-4

Hessian Regularized Distance Metric Learning for People Re-Identification

Guanhua Feng[1] · Weifeng Liu[1] · Dapeng Tao[2] · Yicong Zhou[3]

Abstract

Distance metric learning is a vital issue in people re-identification. Although numerous algorithms have been proposed, it is still challenging especially when the labeled information is few. Manifold regularization can take advantage of labeled and unlabeled information and achieve promising performance in a unified metric learning framework. In this paper, we propose Hessian regularized distance metric learning for people re-identification. Particularly, the second-order Hessian energy prefers functions whose values vary linearly with respect to geodesic distance. Hence Hessian regularization allows us to preserve the geometry of the intrinsic data probability distribution better and then promotes the performance when there is few labeled information. We conduct extensive experiments on the popular VIPeR dataset, CUHK Campus dataset and CUHK03 dataset. The encouraging results suggest that manifold regularization can boost distance metric learning and the proposed Hessian regularized distance metric learning algorithm outperforms the traditional manifold regularized distance metric learning algorithms including graph Laplacian regularization algorithm.

Keywords Metric learning · Person re-identification · Hessian energy · Manifold regularization

✉ Weifeng Liu
liuwf@upc.edu.cn

Guanhua Feng
axjlgh@163.com

Dapeng Tao
dapeng.tao@gmail.com

Yicong Zhou
yicongzhou@um.edu.mo

1 College of Information and Control Engineering, China University of Petroleum (East China), Qingdao 266580, China

2 School of Information Science and Engineering, Yunnan University, Kunming 650091, China

3 Faculty of Science and Technology, University of Macau, Macau 999078, China

Published online: 06 February 2019

1 Introduction

People re-identification, which aims to match the same person across disjoint camera views, has attracted extensive attention in the past few years. It has many important application fields, such as human–computer interaction, security monitoring and so on. Generally speaking, people re-identification is composed of two vital stages. The first stage is to extract discriminative visual features for robust representation. The second stage is to use distance metric for conducting similarity measure to seek correct matches. Then the methods for people re-identification can be divided into two types i.e. feature representation and distance metric learning. The feature representation methods mainly focus on the extraction of appearance features. However, it is hard to compute distinctive and reliable features because people's appearances often undergo large variations across non-overlapping camera views because of significant changes in viewpoint, lighting, background and occlusion. On the other hand, methods of feature representation mostly apply standard distance measure including Manhattan distance, Euclidean distance and Cosine distance. Choosing standard distance measure means treating all features equally. In other words, it won't automatically choose good features or discard bad features, which is inappropriate in some particular situations.

Learning a good metric in feature space is crucial to people re-identification. Dozens of distance metric learning methods have been developed which aim to make use of prior information in form of labels to conduct similarity measures. These methods can be traditionally classified into three categories: supervised form with class labels [1]; Unsupervised form without label information [2]; semi-supervised form with both information [3].

In supervised distance metric learning, the label information of data is assigned in the general form of pairwise constraints. The metric is learned by minimizing the distance of data pairs within the equivalence constraints meanwhile maximizing the distance of data pairs within the inequivalence constraints. For example, Zheng et al. [4] proposed Probabilistic Relative Distance Comparison (PRDC) model to maximize likelihood of true matches which has a comparatively smaller distance than that of the wrong matches. Kostinger et al. [5] developed keep it simple and straightforward (KISS) to learn a metric in the perspective of statistical inference. Later two improved algorithms called Regularized Smoothing KISS (RS-KISS) [6] and QR Decomposition (QRKISS) [7] respectively were further introduced. Mignon et al. [8] proposed pairwise constraint component analysis (PCCA) to cope with high-dimensional input space by learning a projection into a low-dimensional space. Lu et al. [9] proposed a new framework to learn multi-metrics with multiple kernels embedding. Hu et al. [10] designed a two-step metric learning strategy to improve the accuracy of Support Vector Machine (SVM).

In unsupervised distance metric learning, the gist is to learn an underlying low-dimensional subspace in which the geometric distance between the majority of the observed data are preserved. Kodirov E et al. [11] proposed regularized dictionary learning based model to learn cross-view discriminative information from unlabelled data and simultaneously preserve the data relationship in the subspace spanned by the learned dictionary bases. Peng et al. [12] introduced unsupervised cross-dataset transfer learning framework to obtain source-dataset-shared while target-dataset-biased representation which achieved cracking performance. Yu et al. [13] developed clustering-based asymmetric metric learning approach to construct view-specific projection, which not only learns the asymmetric metric but also seeks optimal cluster separations.

In semi-supervised distance metric learning, the majority of methods are based on the manifold assumption which means the sample points are concentrated upon a low-dimensional submanifold instead of being filled in the whole feature space. Liu et al. [14] proposed Semi-supervised Sparse Metric Learning (SSSML) through employing unlabeled samples in Affinity-Preserving Principle and ensuring sparsity of metrics. Niu et al. [15] employed an information-theoretic method named Semi-supervised metric learning paradigm with hyper-sparsity (SERAPH) by maximizing and minimizing the entropy of probability on labeled data and unlabeled data respectively within the principle of entropy regularization. Hoi et al. [16] proposed Laplacian Regularized Metric Learning (LRML) by integrating both labeled data and unlabeled data information through an effective graph regularization framework. In this algorithm, the Laplacian of adjacency graph is computed in an unsupervised manifold using Laplacian Eigenmap with both labeled and unlabeled samples. The data manifold can be approximated with the graph Laplacian.

Although Laplacian regularized metric learning has achieved good performance. However, this kind of semi-supervised algorithm based on the graph Laplacian suffers from the fact that the solution is biased towards a constant and the lack of extrapolating power. Aiming to solve the above problems, we choose Hessian regularization to assist distance metric learning and apply it for person re-identification. Compared to the well-known Laplacian regularization, Hessian has two predominant merits. The first property is that Hessian has good extrapolating power, that is, the values of functions are not restricted by the scope of training outputs. This extrapolation capability can bring about significant improvements especially when the labeled points are insufficient. The second advantage is that Hessian has richer null space. The outputs of the functions contained in Hessian vary by linearity with the geodesic distance along the underlying manifold. This advantage makes it particularly suited for semi-supervised tasks, where the goal is to build user-defined embedding based on the given labels. To evaluate the proposed algorithm, we conduct extensive experiments on publicly available VIPeR dataset, CUHK Campus dataset and CUHK03 dataset for person re-identification. The results demonstrate that our framework is superior to the related algorithms in terms of accuracy and computing complexity.

The rest of this paper is arranged as follows: In Sect. 2, we briefly review the related works including traditional distance metric learning and Hessian regularization. In Sect. 3, we present the proposed Hessian regularized distance metric learning. In Sect. 4, we describe our extensive experiments and discuss the experimental results. In Sect. 5, we conclude the paper with some discussions.

2 Related Works

In this section, we briefly review the related works including distance metric learning and Hessian regularization.

2.1 Distance Metric Learning

Suppose we are given a collection of labeled training set $C = \{x_1, x_2 \cdots x_n\}$, where $x_i \in R^d$. For the examples in C, two sets of pairwise constraints are available:

$$S = \{(x_i, x_j) \mid x_i \text{ and } x_j \text{ are labeled to be similar}\}$$

$$D = \{(x_i, x_j) \mid x_i \text{ and } x_j \text{ are labeled to be dissimilar}\}$$

where S is the set of equivalence constraints and D is the set of inequivalence constraints. For any two points $x_i$and x_j, distance between them can be expressed as follows:

$$\begin{aligned} d_M(x_i, x_j) &= \left\| x_i - x_j \right\|_M \\ &= \sqrt{(x_i - x_j)^T M (x_i - x_j)} \\ &= \sqrt{tr\left(M(x_i - x_j)(x_i - x_j)^T\right)} \end{aligned} \tag{1}$$

In order to guarantee that this be a metric, which satisfies the property of non-negativity and the triangle inequality. We require M is a positive semi-definite matrix and tr is the trace of the matrix. Furthermore, we can assume that there is a kind of corresponding linear mapping P: $\mathbb{R}^m \rightarrow \mathbb{R}^r$, where $P = [p_1, \cdots p_r] \in \mathbb{R}^{d\times r}$, for a possible metric M, the distance between two examples can also be described as:

$$\begin{aligned} d(x_i, x_j) &= \left\| P^T (x_i - x_j)^2 \right\| \\ &= (x_i - x_j)^T PP^T (x_i - x_j) \\ &= (x_i - x_j)^T M (x_i - x_j) \end{aligned} \tag{2}$$

In other words, learning M is equivalent to learning a linear projective mapping P in the feature space. Following the principle of keeping similar data points close and dissimilar data points well-separated. The optimization framework is expressed as follows:

$$\begin{aligned} &\min_{M\geq 0} \sum_{(x_i, x_{j\in S})} \left\| x_i - x_j \right\|_M^2 \\ &s.t. \sum_{(x_i, x_{j\in D})} \left\| x_i - x_j \right\|_M^2 \geq 1 \end{aligned} \tag{3}$$

This formulation attempts to find the metric M by minimizing the sum of squared distances between the similar data points and meanwhile enforcing the sum of distances between the dissimilar data points larger than 1. The above method has been shown effective for cluster tasks, but it is likely to be overfitting. To enhance the generalization and robustness performance, regularization technique has been extensively used in distance metric learning. In Regularized distance metric learning (RDML) [17], a Frobenius norm as the regularization term is added to the overall objective function to improve the generalization and control the model complexity. The method is efficient especially for dealing high dimensional data.

But in most cases, the labeled data is insufficient and unlabeled data is very beneficial to DML task. In order to take advantage of all the given labeled and unlabeled information, Laplacian manifold regularization is utilized in Manifold Regularized Transfer Distance Metric Learning (MTDML) [18] framework. The Laplacian of the adjacency graph is computed in an unsupervised manner using Laplacian Eigenmap with both labeled and unlabeled samples. The linear mapping P is smoothed along the data manifold which can be approximated with the graph Laplacian. Mathematically, the regularization is written as follows:

$$
\begin{aligned}
R(M) &= \frac{1}{2}\sum_{i,j=1}^{N}\left\|Px_i - Px_j^2\right\|W_{ij} \\
&= \sum_{l=1}^{m} p_l^T X(D-W)X^T p_l \\
&= \sum_{l=1}^{m} p_l^T XLX^T p_l = tr\left(P^T XLX^T P\right) \\
&= tr\left(XLX^T PP^T\right) = tr\left(XLX^T M\right)
\end{aligned}
\tag{4}
$$

where W_{ij} is the edge weight between two samples x_i and x_j, If nodes i and j are connected, $W_{ij} = e^{\frac{-\left\|x_i - x_j^2\right\|}{t}}$, $t \in \mathbb{R}$; Otherwise,$W_{ij} = 0$. D is a diagonal matrix whose diagonal elements are equal to the sum of the row entries of W, and $L = D - W$ is known as the Laplacian matrix. Combining the basic loss function with this regularization term, the framework is formulated as follows:

$$
\begin{aligned}
&\min_{M} \gamma_s \sum_{\left(x_i, x_{j\in S}\right)} \left\|x_i - x_j\right\|_M^2 - \gamma_d \sum_{\left(x_i, x_{j\in D}\right)} \left\|x_i - x_j\right\|_M^2 + tr\left(XLX^T M\right) \\
&s.t. M \succeq 0
\end{aligned}
\tag{5}
$$

2.2 Hessian Regularization

It has been proved that the geodesic function in full space of Laplacian is no other than a constant which implies that Laplacian regularization (LR) biases the solution towards a constant function and then leads to poor extrapolation capability [19]. In contrast to Laplacian, Hessian has richer null space and drives the solution varying smoothly along the manifold. In other words, Hessian regularization (HR) is more preferable for exploiting the local geometry than LR. The Hessian of a function f is defined by evaluating Eells-energy $S_{Eells}(f)$ in normal coordinates. $S_{Eells}(f)$ is written for real-valued functions, f $: M \to \mathbb{R}$, as:

$$
S_{Eells}(f) = \int \nabla_a \nabla_b f^2_{T_x^* M \otimes T_x^* M} dV(x) \tag{6}
$$

where M means the m-dimensional data sub-manifold in $\mathbb{R}^d$, $\nabla_a \nabla_b f$ is the second covariant derivative of f, $T_{x_i} M$ is the local tangent space on the point X_i (seen as an m-dimensional affine subspace of $\mathbb{R}^d$). $dV(x)$ is the volume element. Normal coordinates at a given point X_i are coordinates on M such that the manifold looks as Euclidean as possible (up to second order) around X_i. Thus in normal coordinates x_r centered at point X_i:

$$
\left\|\nabla_a \nabla_b f\right\|^2_{T_{x_i}^* M \otimes T_{x_i}^* M} = \sum_{r,s=1}^{m} \left(\frac{\partial^2 f}{\partial x_r \partial x_s}\right)^2 \tag{7}
$$

Then an operator B is defined to give the value $f\left(X_j\right)$ on $N_k\left(X_i\right)$ for estimating Hessian of f at X_i.

$$\frac{\partial^2 f}{\partial x_r \partial x_s}|_{X_i} = \sum_{j=1}^{k} B_{rsj}^{(i)} f(X_j) \tag{8}$$

So we obtain

$$\left\|\nabla_a \nabla_b f\right\|^2 = \sum_{r,s=1}^{m} \left(\sum_{\alpha=1}^{k} B_{rs\alpha}^{(i)} \mathbf{f}_\alpha \right)^2 = \sum_{\alpha,\beta=1}^{k} \mathbf{f}_\alpha \mathbf{f}_\beta H_{\alpha\beta}^{(i)} \tag{9}$$

where $H_{\alpha\beta}^{(i)} = \sum_{r,s=1}^{n} B_{rs\alpha}^{(i)} B_{rs\beta}^{(i)}$, $\mathbf{f} \in \mathbb{R}^k$ and $\mathbf{f}_j = f\left(X_j\right)$ with $X_j \in N_k\left(X_i\right)$. $N_k\left(X_i\right)$ denotes the set of k nearest neighbors (NN) of point X_i. At point X_i, the norm of the second covariant derivative is just the Frobenius norm of the Hessian of f in normal coordinates. The final norm is formulated as follows by summing all the points:

$$\begin{aligned} S_{Hess}(f) &= \sum_{i=1}^{n} \sum_{r,s=1}^{m} \left(\frac{\partial^2 f}{\partial x_r \partial x_s}|_{X_i} \right)^2 \\ &= \sum_{i=1}^{n} \sum_{\alpha \in N_k(X_i)} \sum_{\beta \in N_k(X_i)} \mathbf{f}_\alpha \mathbf{f}_\beta H_{\alpha\beta}^{(i)} \\ &= \langle f, Hf \rangle \end{aligned} \tag{10}$$

Here, the resulting function is called the Hessian regularization $S_{Hess}(f)$ and H is called Hessian matrix.

3 Hessian Regularized Distance Metric Learning

In this section, we introduce our proposed Hessian Regularized distance metric learning. We are given h unlabeled examples $\mathcal{H} = \left\{ \left(x_1, x_2, \ldots, x_h \right) \right\}$ and l labeled examples to make up two sets of pairwise constraints S and D. Based on the above information, we formulate the distance metric learning problem into the following optimization framework:

$$\begin{aligned} \tilde{M} &= \min_M f(M, S, D, \mathcal{H}) \\ & s.t. M \succeq 0 \end{aligned} \tag{11}$$

where f. is some objective function defined over the given data and $\tilde{M}$ is the desired distance metric. Following the regularized loss minimization principle, we define our regularized loss function as:

$$f(M, S, D, \mathcal{H}) = L(M, S, D) + J(M, S, D, \mathcal{H}) \tag{12}$$

or like this:

$$f(M, S, D, \mathcal{H}) = \gamma_a L(S, M) + \gamma_b L(D, M) + \gamma_c J(M, S, D, \mathcal{H}) \tag{13}$$

where γ_a and γ_b and γ_c are the parameters to balance the different terms. $L(S, M)$ and $L(D, M)$ are the loss functions defined on the sets of similar and dissimilar constraints. In this paper, we adopt the sum of squared distances expressions to define two loss functions in terms of effectiveness and efficiency:

$$\begin{aligned} L(S,M) &= \sum_{(x_i,x_{j\in S})} \left\| x_i - x_j \right\|_M^2 \\ &= \sum_{(x_i,x_{j\in S})} \left(x_i - x_j\right)^T M \left(x_i - x_j\right) \\ &= tr(S \cdot M) \end{aligned} \tag{14}$$

$$\begin{aligned} L(D,M) &= \sum_{(x_i,x_{j\in D})} \left\| x_i - x_j \right\|_M^2 \\ &= \sum_{(x_i,x_{j\in D})} \left(x_i - x_j\right)^T M \left(x_i - x_j\right) \\ &= tr(D \cdot M) \end{aligned} \tag{15}$$

As for $J(M, S, D, \mathcal{H})$, the additional regularization term is expressed as follows:

$$J(M,S,D,\mathcal{H}) = \langle f, Hf \rangle = \sum_{\alpha,\beta=1}^{k} H_{\alpha\beta}\left(f_\alpha - f_\beta\right)^2 = \sum_{\alpha,\beta=1}^{k} H_{\alpha\beta}\left\| Px_\alpha - Px_\beta \right\|^2 = tr\left(XHX^T M\right) \tag{16}$$

By summing the energy over all points, the squared norm of the Hessian is actually weighted with local density of the points leading to a stronger penalization of the Hessian in densely sampled regions. Substituting the loss functions and regularization term, the formula of Hessian regularized distance metric learning can now be immediately stated. We have to solve:

$$\begin{aligned} &\arg\min_M \gamma_s tr(S \cdot M) - \gamma_d tr(D \cdot M) + \gamma_c tr\left(XHX^T M\right) \\ &s.t. M \succeq 0,\ \gamma_s \geq 0,\ \gamma_d \geq 0,\ \gamma_c \geq 0 \end{aligned} \tag{17}$$

The optimization obviously is a form of Convex Programming, which can be solved efficiently by using traditional Newton–Raphson method. The details are given:

Optimization algorithm
Step 1. Initialize M with M_0, set $k = 0$ and $0 < \varepsilon \ll 1$.
Step 2. Compute g_k(gradient)andH_k(Hessian). $g_k = \nabla f\left(M_k\right) = \frac{\partial f}{\partial M_k}$, $H_k = \nabla^2 f\left(M_k\right) = \frac{\partial^2 f}{\partial M_k^2}$.
Step 3. If $g_k < \varepsilon$, stop iterating;
Else compute $d_k = -H_k^{-1} \cdot g_k$
Step 4. compute $M_{k+1} \sim = M_k + d_k$
Step 5. $k \sim = k + 1$, Jump to Step 2.

3.1 Complexity Analysis

Suppose we are given n samples. The time complexity of optimization algorithm is O (n^3) which is rather high because it needs to compute H_k and H_k^{-1}. For computing Hessian regularization, the first step is to find the k-nearest neighbors N_k of sample X_i and centralize the neighborhood which determines the time complexity of this process is O ($n \times k$). Compared with computing Laplacian regularization, firstly we have to compute adjacency matrix W in order to compute Laplacian matrix L which means we need to compute the weights of each sample with other samples. In this process, each sample will traverse n times. So the time complexity is O (n^2). Since the number of nearest neighbors of X_i is generally smaller than the number of samples n, there is less computation complexity in computing Hessian regularization than Laplacian regularization.

4 Experiments

We evaluate the proposed Hessian regularized DML on the VIPeR dataset、CUHK Campus dataset and CUHK03 dataset, which are widely used for evaluating people re-identification. By the average cumulative match characteristics (CMC) curves, we obtain the ranked matching rates over 10 trials to evaluate the person re-identification performance of the learned distance metric.

4.1 Dataset Description

To evaluate the performance of the proposed approach, we implement extensive experiments on VIPeR dataset [20], CUHK Campus dataset [21] and CUHK03 dataset [22] which are widely used for evaluation. The VIPeR dataset was captured by two cameras in the open environment. It is a very challenging dataset widely used for benchmark evaluation which suffers from viewpoint changes, pose variations and illumination between two images of a person. It consists of 632 pedestrian pairs. Each pair is composed of images of the same person from different viewpoints. All images are normalized to 128×48 pixels. The CUHK Campus dataset was acquired in two disjoint camera views in a campus environment. It contains 971 pairs of pedestrian images. The images in this dataset have higher resolution than those in VIPeR dataset. Image regions are scaled to 160×60 pixels. The CUHK03 dataset was built under six surveillance cameras with 13,164 images of 1360 pedestrians. Each pedestrian in the dataset is observed in two disjoint views with an average of 4.8 images in one view. What's more, the dataset provides manually cropped person images as well as samples detected with a pedestrian detector. Figure 1 exhibits some samples of VIPeR dataset, CUHK Campus dataset and CUHK03 dataset.

4.2 Feature Selection

In this paper, we employ three types of low-visual feature including Local Binary Pattern (LBP) descriptors, HSV histograms and Lab histograms to represent each normalized image. The features are extracted from 8×16 overlapping blocks and step on every

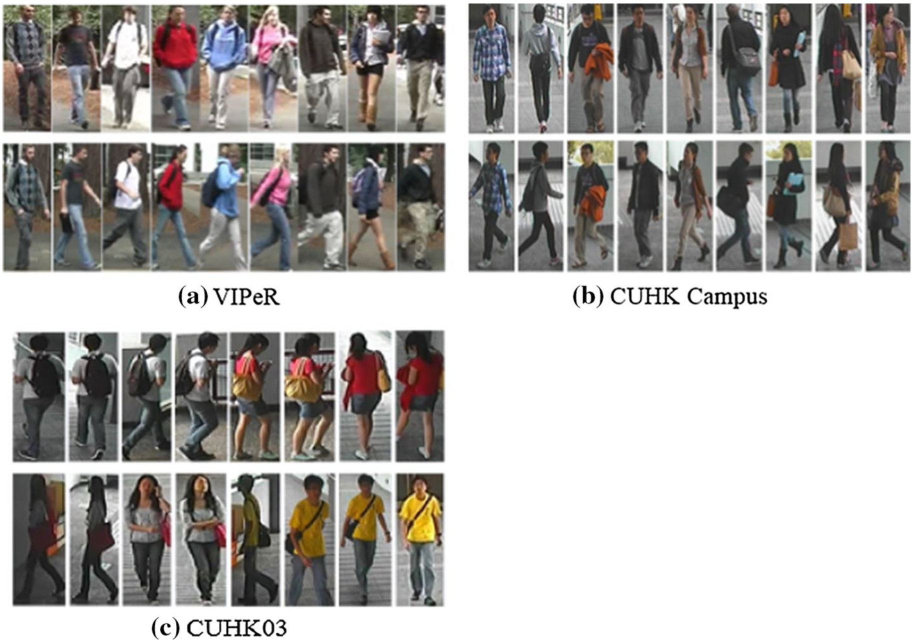

Fig. 1 Some instances from the VIPeR dataset、CUHK Campus dataset and CUHK03 dataset

image is 8×8. Ultimately we link all the feature descriptors and conduct dimensionality reduction by PCA to speed up learning process and reduce interference factors.

4.3 Experimental Settings

We follow the widely adopted evaluation procedure [23] in our experiments on VIPeR dataset, which means randomly choose 316 pairs of person images as training set and the remaining as testing set. Experiments on the CUHK Campus dataset also apply the same evaluation procedure. We randomly select 485 pairs of person images for training and the other 486 pairs for testing. Experiments on the CUHK03 dataset follow the setting of [22] which means the dataset is divided into a training set of 1160 persons and a test set of 100 persons. All processes are repeated for ten times. In the training stage, we use two images of the same person as similar pairs and make up dissimilar image pairs by choosing two images of different individuals. In the test stage, we randomly divide the test set into a gallery set and a probe set. A single image from the probe set is then selected and matched with all images from the gallery set. Parameters $\gamma_s, \gamma_d, \gamma_c$ are tuned from candidate set $\{1 \times 10^e | e = -4, -3, \dots, 3\}$ by cross validation.

4.4 Experimental Results and Analysis

4.4.1 Experiments on the VIPeR Dataset

In this section, we firstly compare the proposed method Hessian regularization DML with some popular methods on the VIPeR dataset. The cumulative matching scores at

Table 1 Comparison of several popular approaches on the VIPeR dataset (M = 316)

Method	Rank 1	Rank 5	Rank 10	Rank 20
Zhao et al. [24]	30.73	52.34	65.8	–
Zeng et al. [25]	31.46	54.65	68.45	81.85
Mid-level filter [26]	29.13	52.54	65.91	79.96
Region-based salience [27]	27.07	50.42	62.67	76.32
LDFV [28]	26.56	56.4	70.93	84.64
Zhao et al. [29]	24.1	49.65	61.3	74.27
Leng et al. [30]	21.72	50.7	70.24	81.45
EIML [31]	20.95	46.57	61.76	76.85
Hessian (Our method)	20.31	41.35	52.16	68.18
SDALF [31]	19.8	39.45	48.64	64.8
KISS [5]	18.35	48.26	62.18	77.06
PRDC [4]	15.33	38.3	53.75	70.09
Mahal distance	17.72	37.82	49.68	62.97
LMNN [33]	17.09	41.61	54.27	68.04
ITML [34]	15.51	36.39	52.06	68.67
Laplacian [16]	13.65	28.8	39.69	51.22
Baseline [35]	6.96	17.03	24.05	32.28

rank-1, rank-5, rank-10 and rank-20 are listed in Table 1. We can see that our proposed method obtains inspired performance. Then we assign 5%, 10%, 20%, 50% of training data as labeled data and the rest as unlabeled data to conduct semi-supervised experiments. The results in Fig. 2 show that the matching rate appears to increase when there are more labeled samples and our proposed method Hessian regularization DML significantly boosts performance than the compared algorithms, thereby validating the effectiveness of Hessian regularization for encoding the marginal distribution information of local geometry. Finally, we conduct the parameter sensitivity analysis of scale-balance parameters γ_d and γ_c which are the most important parameters. The influence of balanced parameters γ_d and γ_c to the Rank 1 matching rates is showed in Fig. 3.

4.4.2 Experiments on the CUHK Campus Dataset

In this section, we compare our proposed method Hessian regularization DML with some popular methods on the CUHK Campus dataset. Table 2 shows different approaches of the cumulative matching scores at rank-1, rank-5, rank-10 and rank-20. As shown in Table 2, our approach outperforms most benchmarking approaches especially be superior to Laplacian regularization DML. Figure 4 reveals the influence of parameter k which is the number of neighbors for computing Hessian to the Rank 1 matching rates. We can see that there is an upper trend of matching rates with the increase of k. When k reaches 80, the matching rates slowly remain stable. Meanwhile, the average training time of six different DML methods is shown in Table 3. We perform the training on a notebook PC with an Intel (R) i5-4200 M @ 2.5GHZ CPU.

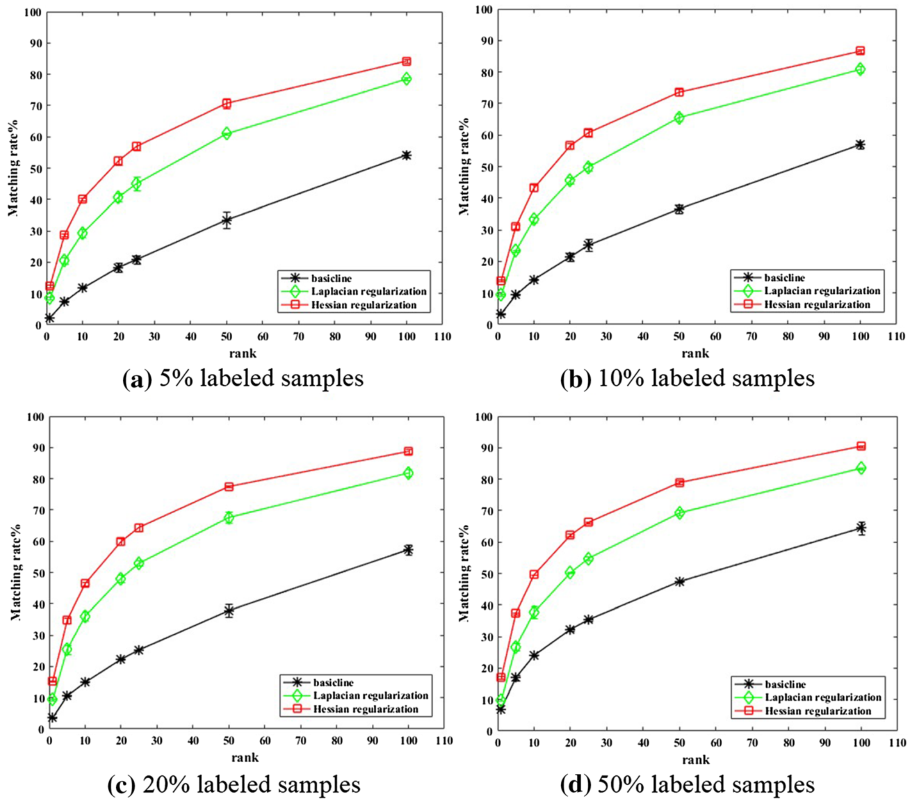

Fig. 2 Matching rate versus different number of labeled samples. The red legend 'Hessian regularization' represents our proposed method

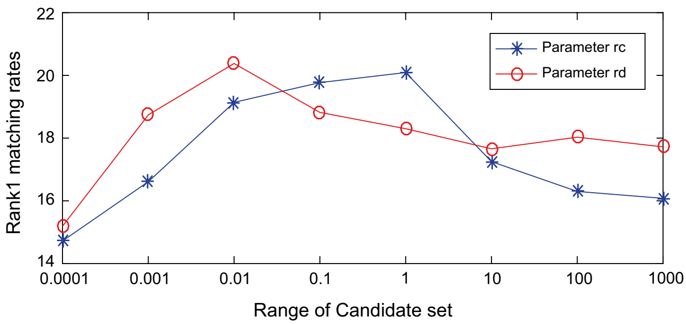

Fig. 3 Performance of Hessian regularization when the parameter γ_d and γ_c varies. Measured by Rank-1 matching rates for VIPeR dataset

4.4.3 Experiments on the CUHK03 Dataset

In this section, we compare our proposed method Hessian regularization DML with some popular methods on manually cropped person images of CUHK03 dataset. The cumulative matching rates at rank-1, rank-5, rank-10 and rank-20 are shown in Table 4.

Table 2 Comparison of several popular approaches on the CUHK Campus dataset (M = 485)

Method	Rank 1	Rank 5	Rank 10	Rank 20
Mid-level filter [26]	32.3	53.2	63.77	72.15
Zhao et al. [24]	28.45	45.87	55.76	67.91
Hessian (Our method)	21.96	40.82	52.03	62.27
Genericmetric [21]	20	43.7	55.13	69.22
Laplacian [16]	18.56	35.88	44.85	53.71
eSDC [29]	19.67	22.51	40	50.64
ITML [34]	15.98	33.54	45.63	59.72
LMNN [33]	13.45	31.56	42.37	53.04
L1-norm distance	10.33	20.84	24.2	31.73
SDALF [32]	9.90	21.25	30.07	39.84
L2-norm distance	9.84	19.89	25.1	32.5
Baseline [35]	8.96	17.11	21.55	28.87

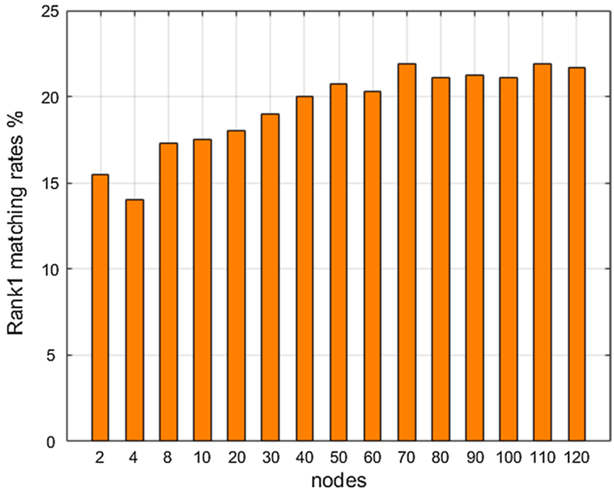

Fig. 4 Performance of Hessian regularization when the parameter *k* varies. Measured by Rank-1 matching rates for CUHK Campus dataset

Table 3 Comparison of training time (s) of six different algorithms

Method	Baseline	ITML	Laplacian	KISSME	LMNN	Hessian
Time (s)	0.97	8.60	1.6	0.01	27.56	1.25

It is obvious that the average matching rates are lower than those in the first two datasets. But compared with related classical algorithms, our method still has a certain degree of improvement.

5 Conclusions

Distance metric learning is critically important for the surveillance task person re-identification. Many semi-supervised algorithms based on manifold regularization have been successfully applied to it. However, the most representative method based on Laplacian Regularization (LR) which suffers from the null space is the constant function along the compact support of the marginal distribution and the lack of extrapolating power particularly when

Table 4 Comparison of several popular approaches on the CUHK03 dataset (M = 100)

Method	Rank 1	Rank 5	Rank 10	Rank 20
FPNN [22]	19.89	49.91	63.57	78.69
Hessian(Our method)	12.41	37.46	47.66	65.87
KISS [5]	11.70	36.78	50.22	67.58
LDM [36]	10.92	34.85	48.15	66.32
Rank [37]	8.52	29.62	42.34	59.68
eSDC [29]	7.68	24.58	35.17	51.76
Laplacian [16]	7.36	26.37	36.28	48.78
LMNN [33]	6.25	20.12	29.88	46.54
ITML [34]	5.14	19.53	29.26	43.23
SDALF [31]	4.87	23.64	34.86	50
L2-norm distance	4.94	17.52	26.33	42.76

the number of labeled samples is small. Therefore, we present Hessian Regularization to tackle these problems for distance metric learning. The proposed method can naturally combine both loss function and Hessian regularization to boost learning performance. Extensive experiments on the VIPeR dataset 、CUHK Campus dataset and CUHK03 dataset demonstrate that the proposed Hessian regularization DML significantly outperforms LR-based and other algorithms.

Acknowledgements This work is supported the National Natural Science Foundation of China under Grant 61671480, in part by the Fundamental Research Funds for the Central Universities, China University of Petroleum (East China) under Grant 18CX07011A.

References

1. Xiong F, Gou M, Camps O et al (2014) Person re-identification using Kernel-based metric learning methods. European Conference on Computer Vision. Springer, Cham, pp 1–16
2. Zhao R, Oyang W, Wang X (2016) Person re-identification by saliency learning. IEEE Trans Pattern Anal Mach Intell 39(2):356–370
3. Chen F, Chai J, Ren D et al (2017) Semi-supervised distance metric learning for person re-identification. In: International conference on computer and information science. IEEE, pp 733–738
4. Zheng W S, Gong S, Xiang T (2011) Person re-identification by probabilistic relative distance comparison. In: Computer vision and pattern recognition. IEEE, pp 649–656
5. Koestinger M, Hirzer M, Wohlhart P et al (2012) Large scale metric learning from equivalence constraints. In: Computer vision and pattern recognition. IEEE, pp 2288–2295
6. Tao D, Jin L, Wang Y et al (2013) Person re-identification by regularized smoothing kiss metric learning. IEEE Trans Circuits Syst Video Technol 23(10):1675–1685
7. Zhao C, Chen Y, Wei Z et al (2018) QRKISS: A two-stage metric learning via QR-decomposition and KISS for person re-identification. In: Neural Processing Letters, pp 1–24
8. Mignon A, Jurie F (2012) Pcca: a new approach for distance learning from sparse pairwise constraints. In: Computer vision and pattern recognition. IEEE, pp 2666–2672
9. Lu X, Wang Y, Zhou X et al (2016) A method for metric learning with multiple-kernel embedding. Neural Process Lett 43(3):905–921
10. Hu L, Hu J, Ye Z et al (2018) Performance analysis for SVM combining with metric learning. In: Neural processing letters, pp 1–22
11. Kodirov E, Xiang T, Gong S (2015) Dictionary learning with iterative Laplacian regularisation for unsupervised person re-identification. In: British machine vision conference, pp 44.1–44.12
12. Peng P, Xiang T, Wang Y et al (2016) Unsupervised cross-dataset transfer learning for person re-identification. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 1306–1315

13. Yu H X, Wu A, Zheng W S (2017) Cross-view asymmetric metric learning for unsupervised person re-identification. In: IEEE international conference on computer vision. IEEE, pp 994–1002
14. Liu W, Ma S, Tao D et al (2010) Semi-supervised sparse metric learning using alternating linearization optimization. In: Proceedings of the 16th ACM SIGKDD international conference on knowledge discovery and data mining. ACM, pp 1139–1148
15. Niu G, Dai B, Yamada M et al (2012) Information-theoretic semi-supervised metric learning via entropy regularization. Neural Comput 26(8):1717–1762
16. Hoi SCH, Liu W, Chang SF (2008) Semi-supervised distance metric learning for collaborative image retrieval. In: Computer Vision and Pattern Recognition. IEEE, pp 1–7
17. Jin R, Wang S, Zhou Y (2009) Regularized distance metric learning: theory and algorithm. In: Advances in neural information processing systems, pp 862–870
18. Shi H, Luo Y, Xu C et al (2015) Manifold Regularized Transfer Distance Metric Learning. In: British machine vision conference, pp 158.1–158.11
19. Kim K I, Steinke F, Hein M (2009) Semi-supervised regression using Hessian energy with an application to semi-supervised dimensionality reduction. In: Adv Neural Inf Process Syst pp 979–987
20. Gray D, Brennan S, Tao H (2007) Evaluating appearance models for recognition, reacquisition, and tracking. In: Proceedings of IEEE international workshop on performance evaluation for tracking and surveillance (PETS), pp 1–7
21. Li W, Zhao R, Wang X (2012) Human reidentification with transferred metric learning. Asian Conference on Computer Vision. Springer, Berlin, pp 31–44
22. Li W, Zhao R, Xiao T et al (2014) DeepReID: Deep filter pairing neural network for person re-identification. In: Computer vision and pattern recognition. IEEE, pp 152–159
23. Liao S, Hu Y, Zhu X et al (2015) Person re-identification by Local Maximal Occurrence representation and metric learning. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 2197–2206
24. Zhao R, Ouyang W, Wang X (2013) Person re-identification by salience matching. Proc IEEE Int Conf Comput Vision 2013:2528–2535
25. Zeng M, Wu Z, Tian C et al (2015) Efficient person re-identification by hybrid spatiogram and covariance descriptor. In: Proceedings of the IEEE conference on computer vision and pattern recognition workshops, pp 48–56
26. Zhao R, Ouyang W, Wang X (2014) Learning Mid-level Filters for Person Re-identification. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 144–151
27. Geng Y, Hu HM, Zeng G et al (2015) A person re-identification algorithm by exploiting region-based feature salience. J Vis Commun Image Represent 29:89–102
28. Ma B, Su Y (2012) Local descriptors encoded by fisher vectors for person re-identification. European Conference on Computer Vision. Springer, Berlin, pp 413–422
29. Zhao R, Ouyang W, Wang X (2013) Unsupervised salience learning for person re-identification. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp 3586–3593
30. Leng Q, Hu R, Liang C et al (2015) Person re-identification with content and context re-ranking. Multimedia Tools Appl 74(17):6989–7014
31. Hirzer M, Roth PM, Bischof H (2012) Person re-identification by efficient impostor-based metric learning. In: Advanced video and signal-based surveillance (AVSS). IEEE, pp 203–208
32. Farenzena M, Bazzani L, Perina A et al (2010) Person re-identification by symmetry-driven accumulation of local features. In: Computer vision and pattern recognition. IEEE, pp 2360–2367
33. Weinberger KQ, Saul LK (2008) Fast solvers and efficient implementations for distance metric learning. In: Proceedings of the 25th international conference on machine learning. ACM, pp 1160–1167
34. Davis J V, Kulis B, Jain P et al (2007) Information-theoretic metric learning. In: Proceedings of the 24th international conference on machine learning. ACM, pp 209–216
35. Xing E P, Jordan M I, Russell S J et al (2003) Distance metric learning with application to clustering with side-information. In: Advances in neural information processing systems, pp 521–528
36. Guillaumin M, Verbeek J, Schmid C (2009) Is that you? Metric learning approaches for face identification. In: ICCV 2009-international conference on computer vision. IEEE, pp 498–505
37. McFee B, Lanckriet G R (2010) Metric learning to rank. In: Proceedings of the 27th international conference on machine learning (ICML-10), pp 775–782