Maximum Margin Metric Learning
over Discriminative Nullspace
for Person Re-identification

T. M. Feroz Ali[(✉)] and Subhasis Chaudhuri

Indian Institute of Technology Bombay, Mumbai, India
{ferozalitm,sc}@ee.iitb.ac.in

Abstract. In this paper we propose a novel metric learning framework called Nullspace Kernel Maximum Margin Metric Learning (NK3ML) which efficiently addresses the small sample size (SSS) problem inherent in person re-identification and offers a significant performance gain over existing state-of-the-art methods. Taking advantage of the very high dimensionality of the feature space, the metric is learned using a maximum margin criterion (MMC) over a discriminative nullspace where all training sample points of a given class map onto a single point, minimizing the within class scatter. A kernel version of MMC is used to obtain a better between class separation. Extensive experiments on four challenging benchmark datasets for person re-identification demonstrate that the proposed algorithm outperforms all existing methods. We obtain 99.8% rank-1 accuracy on the most widely accepted and challenging dataset VIPeR, compared to the previous state of the art being only 63.92%.

Keywords: Person re-identification · Metric learning
Small sample size problem

1 Introduction

Person re-identification (re-ID) is the task of matching the image of pedestrians across spatially non overlapping cameras, even if the pedestrian identities are *unseen* before. It is a very challenging task due to large variations in illumination, viewpoint, occlusion, background and pose changes. Supervised methods for re-ID generally include two stages: computing a robust feature descriptor and learning an efficient distance metric. Various feature descriptors like SDALF [10], LOMO [23] and GOG [31] have improved the efficiency to represent a person. But feature descriptors are unlikely to be completely invariant to large variations in the data collection process and hence the second stage for person re-identification focusing on metric learning is very important. They learn a discriminative metric space to minimize the intra-person distance while maximizing the inter-person distance. It has been shown that learning a good distance metric can drastically improve the matching accuracy in re-ID. Many efficient metric learning methods have been developed for re-ID in the last few years, for e.g., XQDA [23], KISSME

[19], LFDA [36]. However, most of these methods suffer from the small sample size (SSS) problem inherent in re-ID since the feature dimension is often very high.

Recent deep learning based methods address feature computation and metric learning jointly for an improved performance. However, their performance depends on the availability of manually labeled large training data, which is not possible in the context of re-ID. Hence we refrain from discussing deep learning based methods in this paper, and concentrate on the following problem: given a set of image features, can we design a good discriminant criterion for improved classification accuracy for cases when the number of training samples per class is very minimal and the testing identities are unseen during training. Our application domain is person re-identification.

In this paper we propose a novel metric learning framework called Nullspace Kernel Maximum Margin Metric Learning (NK3ML) which efficiently addresses the SSS problem and provide better performance compared to the state-of-the-art approaches for re-ID. The discriminative metric space is learned using a maximum margin criterion over a discriminative nullspace. In the learned metric space, the samples of distinct classes are separated with maximum margin while keeping the samples of same class collapsed to a single point (i.e., zero intra-class variance) to maximize the separability in terms of Fisher criterion.

1.1   Related Methods

Most existing person re-identification methods try to build robust feature descriptors and learn discriminative distance metrics. For feature descriptors, several works have been proposed to capture the invariant and discriminative properties of human images [10,12,18,23,26,31,52,59]. Specifically, GOG [31] and LOMO [23] descriptors have shown impressive robustness against illumination, pose and viewpoint changes.

For recognition purposes, many metric learning methods have been proposed recently [6,15,19,23,36,51,54,61,62]. Most of the metric learning methods in re-ID originated elsewhere and are applied with suitable modification for overcoming the additional challenges in re-identification. Köstinger et al. proposed an efficient metric called KISSME [19] using log likelihood ratio test of two Gaussian distributions. Hirzer et al. [15] used a relaxed positive semi definite constraint of the Mahalanobis metric. Zheng et al. proposed PRDC [62] where the metric is learned to maximize the probability of a pair of true match having a smaller distance than that of a wrong match pair. As an improvement for KISSME [19], Liao et al. proposed XQDA [23] to learn a more discriminative distance metric and a low-dimensional subspace simultaneously. In [36], Pedagadi et al. successfully applied Local Fisher Discriminant Analysis (LFDA) [44] which is a variant of Fisher discriminant analysis to preserve the local structure.

Most metric learning methods based on Fisher-type criterion suffer from the small sample size (SSS) problem [14,61]. The dimensionality of various efficient feature descriptors like LOMO [23] and GOG [31] are in ten thousands and too high compared to the number of samples typically available for training. This

makes the within class scatter matrix singular. Some methods use matrix regularization [23,25,31,36,51] or unsupervised dimensionality reduction [19,36] to overcome the singularity which makes them less discriminative and suboptimal. Also these methods typically have a number of free parameters to tune.

Recently, Null Foley-Sammon Transform (NFST) [3,14,61] has gained increasing attention in computer vision applications. NFST was proposed in [61] to address the SSS problem in re-ID. They find a transformation which collapses the intra class training samples into a single point. By restricting the between class variance to be non zero, they maximize the Fisher discriminant criterion without the need of using any regularization or unsupervised dimensionality reduction.

In this paper, we first identify a serious limitation of NFST, i.e. though NFST minimizes the intra-class distance to zero for all training data, it fails to maximize the inter class distance and has serious consequences creating suboptimality in generalizing the discrimination for test data samples when the test sample does not map to the corresponding singular points. Secondly, we propose a novel metric learning framework called Nullspace Kernel Maximum Margin Metric Learning (NK3ML). The method learns a discriminative metric subspace to maximize the inter-class distance as well as minimize the intra-class distance to zero. NK3ML efficiently addresses the suboptimality of NFST in generalizing the discrimination to test data samples also. In particular, NK3ML first take advantage of NFST to find a low dimensional discriminative nullspace to collapse the intra class samples into a single point. Later NK3ML utilizes a secondary metric learning framework to learn a discriminant subspace using the nullspace to maximally separate the inter-class distance. NK3ML also uses a nonlinear mapping of the discriminative nullspace into an infinite dimensional space using an appropriate kernel to further increase the maximum attainable margin between the inter class samples. The proposed NK3ML does not require regularization nor unsupervised dimensionality reduction and efficiently addresses the SSS problem as well as the suboptimality of NFST in generalizing the discrimination for test data samples. The proposed NK3ML has a closed from solution and has no free parameters to tune.

We first explain NFST in Sect. 2. Later we present NK3ML in Sect. 3 and the experimental results in Sect. 4.

2   Null Foley-Sammon Transform

2.1   Foley-Sammon Transform

The objective of Foley-Sammon Transform (FST) [34,38] is to learn optimal discriminant vectors $\mathbf{w} \in \mathbb{R}^d$ that maximize the *Fisher criterion* $J_F(\mathbf{w})$ under orthonormal constraints:

$$J_F(\mathbf{w}) = \frac{\mathbf{w}^T \mathbf{S}_b \mathbf{w}}{\mathbf{w}^T \mathbf{S}_w \mathbf{w}}. \tag{1}$$

$\mathbf{S}_w$ represents the *within class scatter matrix* and $\mathbf{S}_b$ the *between class scatter matrix*. $\mathbf{x} \in \mathbb{R}^d$ are the data samples with classes $\mathcal{C}_1, \ldots, \mathcal{C}_c$ where c is the total

number of classes. Let n be the total number of samples and n_i the number of samples in class $\mathcal{C}_i$. FST tries to maximize the between class distance and minimize the within class distance simultaneously by maximizing the Fisher criterion.

The optimal discriminant vectors of FST are generated using the following steps. The first discriminant vector $\mathbf{w}_1$ of FST is the unit vector that maximizes $J_F(\mathbf{w}_1)$. If $\mathbf{S}_w$ is nonsingular, the solution becomes a conventional eigenvalue problem: $\mathbf{S}_w^{-1}\mathbf{S}_b\mathbf{w} = \lambda\mathbf{w}$, and can be solved by the normalized eigenvector of $\mathbf{S}_w^{-1}\mathbf{S}_b$ corresponding to its largest eigenvalue. The ith discriminant vector $\mathbf{w}_i$ of FST is calculated by the following optimization problem with orthonormality constraints:

$$\underset{||\mathbf{w}_i||=1, \mathbf{w}_i^T\mathbf{w}_j=0}{\text{maximize}} \quad \{J_F(\mathbf{w}_i)\} \quad j = 1, \ldots, i-1. \tag{2}$$

A major drawback of FST is that it cannot be directly applied when $\mathbf{S}_w$ becomes singular in small sample size (SSS) problems. The SSS problem occures when $n < d$. Common solutions include adding regularization term to $\mathbf{S}_w$ or reducing the dimensionality using PCA, which makes them suboptimal.

2.2 Null Foley-Sammon Transform

The suboptimality due to SSS problem in FST is overcome in an efficient way using Null Foley-Sammon Transform (NFST). The objective of NFST is to find orthonormal discriminant vectors satisfying the following set of constraints:

$$\mathbf{w}^T\mathbf{S}_w\mathbf{w} = 0, \quad \mathbf{w}^T\mathbf{S}_b\mathbf{w} > 0. \tag{3}$$

Each discriminant vector $\mathbf{w}$ should satisfy zero within-class scatter and positive between-class scatter. This leads to $J_F(\mathbf{w}) \to \infty$ and thus NFST tries to attain the best separability in terms of Fisher criterion. Such a vector $\mathbf{w}$ is called *Null Projecting Direction* (NPD). The zero within-class scatter ensures that the transformation using NPDs collapse the intra-class training samples into a single point.

Obtaining Null Projecting Directions: We explain how to obtain the Null Projecting Direction (NPD) of NFST. The total class scatter matrix $\mathbf{S}_t$ is defined as $\mathbf{S}_t = \mathbf{S}_b + \mathbf{S}_w$. We also have $\mathbf{S}_t = \frac{1}{n}\mathbf{P}_t\mathbf{P}_t^T$, where $\mathbf{P}_t$ consists of zero mean data $\mathbf{x}_1 - \mathbf{m}, \ldots, \mathbf{x}_n - \mathbf{m}$ as its columns. Let $\mathbf{Z}_t$ and $\mathbf{Z}_w$ be the null space of $\mathbf{S}_t$ and $\mathbf{S}_w$ respectively. Let $\mathbf{Z}_t^\perp$ represent orthogonal complement of $\mathbf{Z}_t$. Note the lemmas [14].

Lemma 1: Let $\mathbf{A}$ be a positive semidefinite matrix. Then $\mathbf{w}^T A\mathbf{w} = 0$ iff $\mathbf{A}\mathbf{w} = 0$.

Lemma 2: If $\mathbf{w}$ is an NPD, then $\mathbf{w} \in (\mathbf{Z}_t^\perp \cap \mathbf{Z}_w)$.

Lemma 3: For small sample size (SSS) case, there exists exactly $c - 1$ NPDs, c being the number of classes.

In order to obtain the NPDs, we first obtain vectors from the space $\mathbf{Z}_t^{\perp}$. From this space, we next obtain vectors that also satisfy $\mathbf{w} \in \mathbf{Z}_w$. A set of orthonormal vectors can be obtained from the resultant vectors which form the NPDs.

Based on the lemmas, $\mathbf{Z}_t$ can be solved as:

$$\begin{aligned}
\mathbf{Z}_t &= \{\mathbf{w} \mid \mathbf{S}_t\mathbf{w} = 0\} = \{\mathbf{w} \mid \mathbf{w}^T\mathbf{S}_t\mathbf{w} = 0\} \\
&= \{\mathbf{w} \mid (\mathbf{P}_t^T\mathbf{w})^T(\mathbf{P}_t^T\mathbf{w}) = 0\} = \{\mathbf{w} \mid \mathbf{P}_t^T\mathbf{w} = 0\}.
\end{aligned} \tag{4}$$

Thus $\mathbf{Z}_t$ is the null space of $\mathbf{P}_t^T$. So $\mathbf{Z}_t^{\perp}$ is the row space of $\mathbf{P}_t^T$, which is the column space of $\mathbf{P}_t$. Therefore $\mathbf{Z}_t^{\perp}$ is the subspace spanned by zero mean data. $\mathbf{Z}_t^{\perp}$ can be represented using an orthonormal basis $\mathbf{Q} = (\theta_1, \ldots, \theta_{n-1})$, where n is the total number of samples. The basis $\mathbf{Q}$ can be obtained using Gram-Schmidt orthonormalization procedure. Any vector in $\mathbf{Z}_t^{\perp}$ can hence be represented as:

$$\mathbf{w} = \beta_1\theta_1 + \ldots + \beta_{n-1}\theta_{n-1} = \mathbf{Q}\boldsymbol{\beta}. \tag{5}$$

A vector $\mathbf{w}$, satisfying Eq. (5) for any $\boldsymbol{\beta}$, belongs to $\mathbf{Z}_t^{\perp}$. Now we have to find those specific $\boldsymbol{\beta}$ which ensures $\mathbf{w} \in \mathbf{Z}_w$. They can be found by substituting (5) in the condition for $\mathbf{w} \in \mathbf{Z}_w$ as follows:

$$\begin{aligned}
0 = \mathbf{S}_w\mathbf{w} &= \mathbf{w}^T\mathbf{S}_w\mathbf{w} = (\mathbf{Q}\beta)^T\mathbf{S}_w(\mathbf{Q}\boldsymbol{\beta}) \\
&= \boldsymbol{\beta}^T(\mathbf{Q}^T\mathbf{S}_w\mathbf{Q})\boldsymbol{\beta} = \mathbf{Q}^T\mathbf{S}_w\mathbf{Q}\boldsymbol{\beta}.
\end{aligned} \tag{6}$$

Hence $\boldsymbol{\beta}$ can be solved by finding the null space of $\mathbf{Q}^T\mathbf{S}_w\mathbf{Q}$. The set of solutions $\{\boldsymbol{\beta}\}$ can be chosen orthonormal. Since the dimension of $\mathbf{w} \in (\mathbf{Z}_t^{\perp} \cap \mathbf{Z}_w)$ is $c-1$ [14], we get $c-1$ solutions for $\boldsymbol{\beta}$. The $c-1$ NPDs can now be computed using (5). Since $\mathbf{Q}$ and $\{\boldsymbol{\beta}\}$ are orthonormal, the resulting NPDs are also orthonormal. The projection matrix $\mathbf{W}_N \in \mathbb{R}^{d \times (c-1)}$ of NFST now constitutes of the $c-1$ NPDs as its columns.
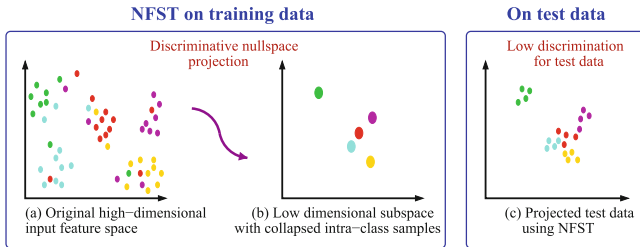

Fig. 1. Illustration of the suboptimality in NFST. Each color corresponds to distinct classes. (Color figure online)

3   Nullspace Kernel Maximum Margin Metric Learning

Methods based on Fisher criterion, in general, learn the discriminant vectors using the training samples so that the vectors generalize well for the test data

also in terms of separability of classes. NFST [3,14] was proposed in [61] to address the SSS problem in re-ID. They find a transformation by collapsing the intra-class samples into a single point. We identify a serious limitation of NFST. Maximizing $J_F(\mathbf{w})$ in Eq. (1) by making the denominator to zero, does not allow to make use of the information contained in the numerator. As illustrated in Fig. 1, the mapped singular points in the NFST projected space for two different classes may be quite close. Thus, when a test data is projected into this NFST nullspace, it no longer maps to the same singular point. Rather, it maps to a point close to the above point. But this projected point may be closer to the singular point for the other class and misclassification takes place. Under the NFST formulation, one has no control on this aspect as one makes $\mathbf{w}^T\mathbf{S}_w\mathbf{w} = 0$, but $\mathbf{w}^T\mathbf{S}_b\mathbf{w}$ may also be very small instead of being large, and the classification performance may be very poor.
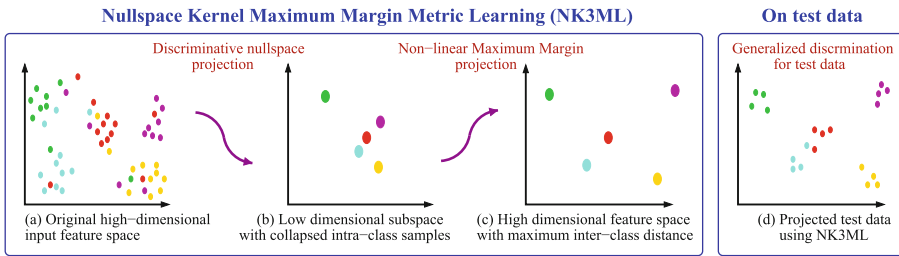

Fig. 2. Illustration of our method NK3ML. Each color corresponds to distinct classes. (Color figure online)

In this paper we propose a metric learning framework, namely, Nullspace Kernel Maximum Margin Metric Learning (NK3ML) to improve the limitation of NFST and better handle the classification of high dimensional data. As shown in Fig. 2, NK3ML first take advantage of NFST to find a low dimensional discriminative nullspace to collapse the intra-class samples into a single point. Later it uses a modified version of Maximum Margin Criterion (MMC) [20] to learn a discriminant subspace using the nullspace to maximally separate the inter-class distance. Further, to obtain the benefit of kernel based techniques, instead of using the MMC, we obtain the *Normalized Kernel Maximum margin criterion* (NKMMC) which is efficient and robust to learn the discriminant subspace to maximize the distances among the classes. NK3ML can efficiently address the suboptimality of NFST in enhancing the discrimination to test data samples also.

3.1   Maximum Margin Criterion

Maximum margin criterion (MMC) [20,21] is an efficient way to learn a discriminant subspace which maximize the distances between classes. For the separability of classes $\mathcal{C}_1, \ldots, \mathcal{C}_c$, the maximum margin criterion is defined as

$$J = \frac{1}{2} \sum_{i=1}^{c} \sum_{j=1}^{c} p_i p_j d(\mathcal{C}_i, \mathcal{C}_j), \tag{7}$$

where the inter-class margin (or distance) of class $\mathcal{C}_i$ and $\mathcal{C}_j$ is defined as

$$d(\mathcal{C}_i, \mathcal{C}_j) = d(\mathbf{m}_i, \mathbf{m}_j) - s(\mathcal{C}_i) - s(\mathcal{C}_j), \tag{8}$$

and $d(\mathbf{m}_i, \mathbf{m}_j)$ represents the squared Euclidean distance between mean vectors $\mathbf{m}_i$ and $\mathbf{m}_j$ of classes $\mathcal{C}_i$ and $\mathcal{C}_j$, respectively. $s(\mathcal{C}_i)$ is the scatter of class $\mathcal{C}_i$, estimated as $s(\mathcal{C}_i) = tr(\mathbf{S}_i)$ where $\mathbf{S}_i$ is the within class scatter matrix of class $\mathcal{C}_i$. The inter-class margin can be solved to get $d(\mathcal{C}_i, \mathcal{C}_j) = tr(\mathbf{S}_b - \mathbf{S}_w)$. A set of r unit linear discriminant vectors $\{\mathbf{v}_k \in \mathbb{R}^d | k = 1, \ldots, r\}$ is learned such that they maximize J in the projected subspace. If $\mathbf{V} \in \mathbb{R}^{d \times r}$ is the projection matrix, the MMC criterion becomes $J(\mathbf{V}) = tr(\mathbf{V}^T(\mathbf{S}_b - \mathbf{S}_w)\mathbf{V})$. The optimization problem can be equivalently written as:

$$\begin{aligned} &\underset{\mathbf{v}_k}{\text{maximize}} && \sum_{k=1}^{r} \mathbf{v}_k^T(\mathbf{S}_b - \mathbf{S}_w)\mathbf{v}_k, \\ &\text{subject to} && \mathbf{v}_k^T \mathbf{v}_k = 1, \qquad k = 1, \ldots, r. \end{aligned} \tag{9}$$

The optimal solutions are obtained by finding the normalized eigenvectors of $\mathbf{S}_b - \mathbf{S}_w$ corresponding to its first r largest eigenvectors.

3.2   Kernel Maximum Margin Criterion

Kernels methods are well known techniques to learn non-linear discriminant vectors. They use an appropriate non-linear function $\Phi(\mathbf{z})$ to map the input data $\mathbf{z}$ to a higher dimensional feature space $\mathcal{F}$ and find discriminant vectors $\mathbf{v}_k \in \mathcal{F}$. Given n training data samples and a kernel function $k(\mathbf{z}_i, \mathbf{z}_j) = \langle \Phi(\mathbf{z}_i), \Phi(\mathbf{z}_j) \rangle$, we can calculate the kernel matrix $\mathbf{K} \in \mathbb{R}^{n \times n}$. The matrix $\mathbf{K}_i \in \mathbb{R}^{n \times n_i}$ for the ith class with n_i samples is $(\mathbf{K}_i)_{pq} := k(\mathbf{z}_p, \mathbf{z}_q^{(i)})$. As every discriminant vector $\mathbf{v}_k$ lies in the span of the mapped data samples, it can be expressed in the form $\mathbf{v}_k = \sum_{j=1}^{n} (\boldsymbol{\alpha}_k)_j \Phi(\mathbf{z}_j)$, where $(\boldsymbol{\alpha}_k)_j$ is the jth element of the vector $\boldsymbol{\alpha}_k \in \mathbb{R}^n$, which constitutes the expansion coefficients of $\mathbf{v}_k$. The optimization problem proposed for Kernel Maximum Margin Criterion (KMMC) [20] is:

$$\begin{aligned} &\underset{\boldsymbol{\alpha}_k}{\text{maximize}} && \sum_{k=1}^{r} \boldsymbol{\alpha}_k^T(\mathbf{M} - \mathbf{N})\boldsymbol{\alpha}_k, \\ &\text{subject to} && \boldsymbol{\alpha}_k^T \boldsymbol{\alpha}_k = 1, \end{aligned} \tag{10}$$

where $\mathbf{N} := \sum_{i=1}^{c} \frac{1}{n} \mathbf{K}_i(\mathbf{I}_{n_i} - \frac{1}{n_i} \mathbf{1}_{n_i} \mathbf{1}_{n_i}^T)\mathbf{K}_i^T$, $\mathbf{I}_{n_i}$ is $(n_i \times n_i)$ identity matrix; $\mathbf{1}_{n_i}$ is n_i dimensional vector of ones and $\mathbf{M} = \sum_{i=1}^{c} \frac{1}{n_i}(\widetilde{\mathbf{m}}_i - \widetilde{\mathbf{m}})(\widetilde{\mathbf{m}}_i - \widetilde{\mathbf{m}})^T$; $\widetilde{\mathbf{m}} := \frac{1}{n} \sum_{i=1}^{c} n_i \widetilde{\mathbf{m}}_i$ and $(\widetilde{\mathbf{m}}_i)_j := \frac{1}{n_i} \sum_{\mathbf{z} \in \mathcal{C}_i} k(\mathbf{z}, \mathbf{z}_j)$. The optimal solutions are the normalized eigenvectors of $(\mathbf{M} - \mathbf{N})$, corresponding to its first r largest eigenvalues.

3.3   NK3ML

The kernalized optimization problem given in (10) obtained by KMMC [20] does not enforce normalization of discriminant vectors in the feature space, but rather uses normalization constraint on eigenvector expansion coefficient vector $\boldsymbol{\alpha}_k$. In NK3ML, we require the discriminant vectors obtained by KMMC to be normalized, i.e., $\mathbf{v}_k^T \mathbf{v}_k = 1$. The normalized discriminant vectors are important to preserve the shape of the distribution of data. Hence we derive *Normalized Kernel Maximum Margin Criterion* (NKMMC) as follows. We rewrite the discriminant vector $\mathbf{v}_k$ as:

$$\mathbf{v}_k = \sum_{j=1}^{n} (\boldsymbol{\alpha}_k)_j \boldsymbol{\Phi}(\mathbf{z}_j) = \Big[\boldsymbol{\Phi}(\mathbf{z}_1) \ \boldsymbol{\Phi}(\mathbf{z}_2) \ \ldots \ \boldsymbol{\Phi}(\mathbf{z}_n) \Big] \boldsymbol{\alpha}_k. \tag{11}$$

Then normalization constraint becomes

$$\Big(\sum_{j=1}^{n} (\boldsymbol{\alpha}_k)_j \boldsymbol{\Phi}(\mathbf{z}_j) \Big)^T \Big(\sum_{j=1}^{n} (\boldsymbol{\alpha}_k)_j \boldsymbol{\Phi}(\mathbf{z}_j) \Big) = 1$$

$$\Rightarrow \qquad \boldsymbol{\alpha}_k^T \mathbf{K} \boldsymbol{\alpha}_k = 1. \tag{12}$$

where $\mathbf{K}$ is the kernel matrix. The optimization problem in (10) can now be reformulated to enforce normalized discriminant vectors as follows.

$$\begin{aligned} \underset{\boldsymbol{\alpha}_k}{\text{maximize}} \quad & \sum_{k=1}^{r} \boldsymbol{\alpha}_k^T (\mathbf{M} - \mathbf{N}) \boldsymbol{\alpha}_k, \\ \text{subject to} \quad & \boldsymbol{\alpha}_k^T \mathbf{K} \boldsymbol{\alpha}_k = 1. \end{aligned} \tag{13}$$

We introduce a Lagrangian to solve the above problem.

$$\mathcal{L}(\boldsymbol{\alpha}_k, \lambda_k) = \sum_{k=1}^{r} \boldsymbol{\alpha}_k^T (\mathbf{M} - \mathbf{N}) \boldsymbol{\alpha}_k + \lambda_k (\boldsymbol{\alpha}_k^T \mathbf{K} \boldsymbol{\alpha}_k - 1), \tag{14}$$

where λ_k is the Lagrangian multiplier. The Lagrangian $\mathcal{L}$ has to be maximized with respect to $\boldsymbol{\alpha}_k$ and the multipliers λ_k. The derivatives of $\mathcal{L}$ with respect to $\boldsymbol{\alpha}_k$ should vanish at the stationary point.

$$\begin{aligned} \frac{\partial \mathcal{L}(\boldsymbol{\alpha}_k, \lambda_k)}{\partial \boldsymbol{\alpha}_k} &= (\mathbf{M} - \mathbf{N} - \lambda_k \mathbf{K}) \boldsymbol{\alpha}_k = 0 \quad \forall \, k = 1, \ldots, r \\ &\Rightarrow (\mathbf{M} - \mathbf{N}) \boldsymbol{\alpha}_k = \lambda_k \mathbf{K} \boldsymbol{\alpha}_k. \end{aligned} \tag{15}$$

This is a generalized eigenvalue problem. λ_k's are the generalized eigenvalues and $\boldsymbol{\alpha}_k$'s the generalized eigenvectors of $(\mathbf{M} - \mathbf{N})$ and $\mathbf{K}$. The objective function at this stationary point is given as:

$$\sum_{k=1}^{r} \boldsymbol{\alpha}_k^T (\mathbf{M} - \mathbf{N}) \boldsymbol{\alpha}_k = \sum_{k=1}^{r} \lambda_k \boldsymbol{\alpha}_k^T \mathbf{K} \boldsymbol{\alpha}_k = \sum_{k=1}^{r} \lambda_k. \tag{16}$$

Hence the objective function in NKMMC is maximized by the *generalized* eigenvectors corresponding to the first r generalized eigenvalues of $(\mathbf{M} - \mathbf{N})$ and $\mathbf{K}$. We choose all the eigenvectors with positive eigenvalues, since they ensure maximum inter-class margin, i.e., the samples of different classes are well separated in the direction of these eigenvectors. It should be noted that our NKMMC has a different solution from that of original KMMC [20], since KMMC uses standard eigenvectors of $\mathbf{M} - \mathbf{N}$.

NFST is first used to learn the discriminant vectors using the training data $\{\mathbf{x}\}$. The discriminants of NFST form the projection matrix $\mathbf{W}_N$. Each training data sample $\mathbf{x} \in \mathbb{R}^d$ is projected as

$$\mathbf{z} = \mathbf{W}_N^T \mathbf{x}. \tag{17}$$

Each projected data sample $\mathbf{z} \in \mathbb{R}^{c-1}$ now lies in the discriminative nullspace of NFST. Now we use all the projected data $\{\mathbf{z}\}$ for learning the secondary distance metric using NKMMC.

Any general feature vector $\widetilde{\mathbf{x}} \in \mathbb{R}^d$ can be projected onto the discriminant vector $\mathbf{v}_k$ of NK3ML in two steps:
Step 1: Project $\widetilde{\mathbf{x}}$ onto the nullspace of NFST to get $\widetilde{\mathbf{z}}$:

$$\widetilde{\mathbf{z}} = \mathbf{W}_N^T \widetilde{\mathbf{x}}. \tag{18}$$

Step 2: Project the $\widetilde{\mathbf{z}}$ onto the discriminant vector $\mathbf{v}_k$ of NKMMC:

$$\mathbf{v}_k^T \boldsymbol{\Phi}(\widetilde{\mathbf{z}}) = \Big(\sum_{j=1}^{n} (\boldsymbol{\alpha}_k)_j \boldsymbol{\Phi}(\mathbf{z}_j) \Big)^T \boldsymbol{\Phi}(\widetilde{\mathbf{z}}) = \sum_{j=1}^{n} (\boldsymbol{\alpha}_k)_j k(\mathbf{z}_j, \widetilde{\mathbf{z}}). \tag{19}$$

The proposed NK3ML does not require any regularization or unsupervised dimensionality reduction and can efficiently address the SSS problem as well as the suboptimality of NFST in generalizing the discrimination for test data samples. The NK3ML has a closed form solution and no free parameters to tune. The only issue to be decided is what kernel to be used. In effect what the proposed method does is to project the data into the NFST nullspace, where the dimensionality of the feature space is reduced to force all points belonging to a given class to a single point. In the second stage, the dimensionality is increased by using an appropriate kernel in conjunction with NKMMC, thereby allowing us to enhance the between class distance. This provides a better margin while classifying the test samples.

4   Experimental Results

Parameter Settings: There are no free parameters to tune in NK3ML, unlike most state-of-the-art methods which have to carefully tune their parameters to attain their best results. In all the experiments, we use the RBF kernel whose kernel width is set to be the root mean squared pairwise distance among the samples.

Table 1. Comparison of NK3ML with baselines on GRID and PRID450S datasets

Methods	GRID		PRID450S	
	Rank1	Rank10	Rank1	Rank10
WHOS + NK3ML	21.20	55.60	50.67	88.09
WHOS + NFST	18.64	52.32	42.58	77.07
WHOS + KNFST	21.12	54.32	45.87	85.78
WHOS + XQDA	18.72	52.56	43.38	77.91
LOMO + NK3ML	18.24	43.76	60.62	91.96
LOMO + NFST	17.04	42.64	58.84	89.42
LOMO + KNFST	14.88	41.28	59.47	91.96
LOMO + XQDA	16.56	41.84	59.78	90.09
GOG + NK3ML	26.96	57.52	68.04	95.07
GOG + NFST	24.88	58.00	67.60	94.18
GOG + KNFST	24.88	53.28	64.80	94.00
GOG + XQDA	24.80	58.40	68.00	94.36

Fig. 3. Sample images of PRID450S dataset. Images with the same column corresponds to the same identities.

Datasets: The proposed NK3ML is evaluated on four popular benchmark datasets: PRID450S [37], GRID [27], CUHK01 [22] and VIPeR [12], respectively contains 450, 250, 971, and 632 identities captured in two disjoint camera views. CUHK01 contains two images for each person in one camera view and all other datasets contain just one image. Quite naturally, these datasets constitute the extreme examples of SSS. Following the conventional experimental setup [1,5,23,31,35,52], each dataset is randomly divided into training and test sets, each having half of the identities. During testing, the probe images are matched against the gallery. In the test sets of all datasets, except GRID, the number of probe images and gallery images are equal. The test set of GRID has additional 775 gallery images that do not belong to the 250 identities. The procedure is repeated 10 times and the average rank scores are reported.

Features: Most existing methods use a fixed feature descriptor for all datasets. Such an approach is less efficient to represent the intrinsic characteristics of each dataset. Hence in NK3ML, we use specific set of feature descriptors for each dataset. We choose from the standard feature descriptors GOG [31] and WHOS [26]. We also use an improved version of LOMO [23] descriptor, which we call *LOMO**. We generate it by concatenating the LOMO features generated using YUV and RGB color spaces separately.

Method of Comparison: We use only the available data in each dataset for training. No separate pre-processing of the features or images (such as domain adaptation/body parts detection), or post-processing of the classifier has been used in the study. There has been some efforts on using even the test data for re-ranking of re-ID results [1,2,63] to boost up the accuracy. But these techniques being not suitable for any real time applications, we refrain from using such supplementary methods in our proposal.

Table 2. Comparison with state-of-the-art results on (a) GRID and (b) PRID450S dataset. The best and second best scores are shown in red and blue, respectively. The methods with a * signifies pre/post-processing based methods

(a) GRID dataset

Methods	Rank1	Rank10	Rank20
MtMCML[28]	14.08	45.84	59.84
KNFST[61]	14.88	41.28	50.88
PolyMap[6]	16.30	46.00	57.60
LOMO+XQDA[23]	16.56	41.84	52.40
MLAPG[24]	16.64	41.20	52.96
KEPLER[30]	18.40	50.24	61.44
DR-KISS[45]	20.60	51.40	62.60
SSSVM[54]	22.40	51.28	61.20
SCSP[5]	24.24	54.08	65.20
GOG+XQDA[31]	24.80	58.40	68.88
NK3ML(Ours)	27.20	60.96	71.04
*SSDAL[43]	22.40	48.00	58.40
*SSM[1]	27.20	61.12	70.56
*OL-MANS[64]	30.16	49.20	59.36

(b) PRID450S dataset

Methods	Rank1	Rank10	Rank20
WARCA[16]	24.58	-	-
SCNCD[52]	41.60	79.40	87.80
CSL[39]	44.40	82.20	89.80
TMA[29]	52.89	85.78	93.33
KNFST[61]	59.47	91.96	96.53
LOMO+XQDA[23]	59.78	90.09	95.29
SSSVM[54]	60.49	88.58	93.60
GOG+XQDA[31]	68.00	94.36	97.64
NK3ML(Ours)	73.42	96.31	98.58
*Semantic[41]	44.90	77.50	86.70
*SSM[1]	72.98	96.76	99.11

4.1 Comparison with Baselines

In Table 1, we compare the performances of NK3ML with the baseline metric learning methods. As NK3ML is proposed as an improvement to address the limitations of NFST, we first compare the performance of NK3ML with NFST. For fair comparison with NFST, we also use its kernalized version KNFST [61]. KNFST is also the state-of-the-art metric learning method applied for LOMO descriptor. For uniformity, all metric learning methods are evaluated using the same standard feature descriptors LOMO [23], WHOS [26] and GOG [31]. We also compare with Cross-view Quadratic Discriminant Analysis (XQDA) [31] which is the state of the art metric learning method for GOG descriptor. XQDA is also successfully applied with LOMO in many cases [23]. We use GRID and PRID450S datasets for comparison with the baselines. GRID is a pretty difficult person re-identification dataset having poor image quality with large variations in pose and illuminations, which makes it very challenging to obtain good matching accuracies. PRID450S is also a challenging dataset due to the partial occlusion, background interference and viewpoint changes. From the results in Table 1, it can be seen that NK3ML provides significant performance gains against all the baselines for all the standard feature descriptors (Fig. 3).

Comparison with NFST: NK3ML provides a good performance gain against NFST. In particular for PRID450S dataset, when compared using WHOS, NK3ML provides an improvement of 8.09% at rank-1 and 11.02% at rank-10. Similar gain can also be seen while using LOMO and GOG features for both GRID and PRID450S datasets.

Table 3. Comparison with state-of-the-art results on CUHK01 dataset using (a) single-shot and (b) multi-shot settings. ** corresponds to deep learning based methods

<table>
<tr><td colspan="4" align="center">(a) single-shot</td><td colspan="4" align="center">(b) multi-shot</td></tr>
<tr><td>Methods</td><td>Rank1</td><td>Rank10</td><td>Rank20</td><td>Methods</td><td>Rank1</td><td>Rank10</td><td>Rank20</td></tr>
<tr><td>MLFL[59]</td><td>34.30</td><td>65.00</td><td>75.00</td><td>l1-Graph[17]</td><td>50.10</td><td>-</td><td>-</td></tr>
<tr><td>LOMO+XQDA[23]</td><td>50.00</td><td>83.40</td><td>89.51</td><td>LOMO+XQDA[23]</td><td>61.98</td><td>89.30</td><td>93.62</td></tr>
<tr><td>KNFST[61]</td><td>52.80</td><td>84.97</td><td>91.07</td><td>CAMEL[53]</td><td>62.70</td><td>-</td><td>-</td></tr>
<tr><td>CAMEL[53]</td><td>57.30</td><td>-</td><td>-</td><td>MLAPG[24]</td><td>64.24</td><td>90.84</td><td>94.92</td></tr>
<tr><td>GOG+XQDA[31]</td><td>**57.89**</td><td>**86.25**</td><td>**92.14**</td><td>SSSVM[54]</td><td>65.97</td><td>-</td><td>-</td></tr>
<tr><td>WARCA[16]</td><td>58.34</td><td>-</td><td>-</td><td>KNFST[61]</td><td>66.07</td><td>91.56</td><td>95.64</td></tr>
<tr><td>NK3ML(Ours)</td><td>**67.09**</td><td>**91.85**</td><td>**95.92**</td><td>GOG+XQDA[31]</td><td>**67.28**</td><td>**91.77**</td><td>**95.93**</td></tr>
<tr><td>*Semantic[41]</td><td>32.70</td><td>64.40</td><td>76.30</td><td>NK3ML(Ours)</td><td>**76.77**</td><td>**95.58**</td><td>**98.02**</td></tr>
<tr><td>*MetricEnsemble[35]</td><td>53.40</td><td>84.40</td><td>90.50</td><td>**DGD[50]</td><td>66.60</td><td>-</td><td>-</td></tr>
<tr><td>**TPC[8]</td><td>53.70</td><td>91.00</td><td>96.30</td><td>*OLMANS[64]</td><td>68.44</td><td>92.67</td><td>95.88</td></tr>
<tr><td>**Quadruplet[7]</td><td>62.55</td><td>89.71</td><td>-</td><td>*SHaPE[2]</td><td>76.00</td><td>-</td><td>-</td></tr>
<tr><td>*DLPAR[56]</td><td>**72.30**</td><td>**94.90**</td><td>**97.20**</td><td>*Spindle[55]</td><td>**79.90**</td><td>**97.10**</td><td>**98.60**</td></tr>
</table>

Comparison with KNFST: In spite of KNFST being the state-of-the-art metric learning method for LOMO descriptor, NK3ML outperforms KNFST with a significant difference. In GRID dataset, NK3ML gains 3.36% in rank-1 and 2.48% in rank-10. Similar improvements are seen for other features also for both datasets.

Comparison with XQDA: For GOG descriptor, XQDA is the state of the art metric learning method. At rank-1, NK3ML gains 2.16% in GRID. Similarly, it gains 7.29% at rank-1 in PRID450S using WHOS descriptor.

Based on the above comparisons, it may be concluded that NK3ML attains a much better margin over NFST as expected from the theory. Also NK3ML outperforms KNFST and XQDA for all aforementioned standard feature descriptors.

4.2   Comparison with State-of-the-Art

In the performance comparison of NK3ML with the state-of-the-art methods, we also report the accuracies of pre/post processing methods on separate rows for completeness. As mentioned previously, direct comparisons of our results with pre/post processing methods are not advisable. However, even if such a comparison is made, we still have accuracies that are best or comparable to the best existing techniques on most of the evaluated datasets. Moreover, our approach is general enough to be easily integrated with the existing pre/post processing methods to further increase their accuracy.

Experiments on GRID Dataset: We use GOG and LOMO* as the feature descriptor for GRID. Table 2a shows the performance comparison of NK3ML. GOG + XQDA [31] reports the best performance of 24.8% at rank-1 till date. NK3ML achieves an accuracy of 27.20% at rank-1, outperforming GOG+XQDA

by 2.40%. At rank-1, NK3ML also outperforms all the post processing methods except OL-MANS [64], which uses the test data and train data together to learn a better similarity function. However, the penalty for misclassification at rank-1, if any, severely affects the rank-N performance for OL-MANS. NK3ML outperforms OL-MANS by 11.76% at rank-10 and 11.68% at rank-20.

Experiments on PRID450S Dataset: GOG and LOMO* are used as the feature descriptor for PRID450S. NK3ML provides the best performances at all ranks, as shown in Table 2b. Especially, it provides an improvement margin of 5.42% in rank-1 compared to the second best method GOG+XQDA [31]. At rank-1, NK3ML also outperforms all the post processing based methods. SSM [1] incorporates XQDA as the metric learning method. As analyzed in Sect. 4.1, since NK3ML outperforms XQDA, it can be anticipated that even the re-ranking methods like SSM can benefit from NK3ML.

Table 4. Comparison with state-of-the-art results on VIPeR dataset. RN means Rank-N accuracy

Methods	Ref	R1	R10	R20	Methods	Ref	R1	R10	R20
ELF[12]	ECCV2008	12.0	44.0	61.0	SSSVM[54]	CVPR2016	42.1	84.3	91.9
PCCA[32]	CVPR2012	19.3	64.9	80.3	**TPC[8]	CVPR2016	47.8	84.8	91.1
KISSME[19]	CVPR2012	19.6	62.2	77.0	GOG+XQDA[31]	CVPR2016	49.7	88.7	94.5
LFDA[36]	CVPR2013	24.2	67.1	-	SCSP[5]	CVPR2016	53.5	91.5	96.7
eSDC[58]	CVPR2013	26.7	62.4	76.4	**SCNN[46]	ECCV2016	37.8	66.9	-
SalMatch[57]	ICCV2013	30.2	-	-	**Shi et al.[40]	ECCV2016	40.9	-	-
MLFL[59]	CVPR2014	29.1	66.0	79.9	$l1$-graph[17]	ECCV2016	41.5	-	-
rPCCA[51]	ECCV2014	22.0	71.0	85.3	**S-LSTM[47]	ECCV2016	42.4	79.4	-
kLFDA[51]	ECCV2014	32.3	79.7	90.9	*SSDAL[43]	ECCV2016	43.5	81.5	89.0
SCNCD[52]	ECCV2014	37.8	81.2	90.4	*TMA[29]	ECCV2016	48.2	87.7	93.5
PolyMap[6]	CVPR2015	36.8	83.7	91.7	*SSM[1]	CVPR2017	53.7	91.5	96.1
LOMO+XQDA[23]	CVPR2015	40.0	80.5	91.1	*Spindle[55]	CVPR2017	53.8	83.2	92.1
*Semantic[41]	CVPR2015	41.6	86.2	95.1	CAMEL[53]	ICCV2017	30.9	-	-
QALF[60]	CVPR2015	30.2	62.4	73.8	*MuDeep	ICCV2017	43.0	85.8	-
CSL[39]	ICCV2015	34.8	82.3	91.8	*OLMANS[64]	ICCV2017	45.0	85.0	93.6
MLAPG[24]	ICCV2015	40.7	82.3	92.4	*DLPAR[56]	ICCV2017	48.7	85.1	93.0
*DCIA[11]	ICCV2015	**63.9**	87.5	-	*PDC[42]	ICCV2017	51.3	84.2	91.5
**DGD[50]	CVPR2016	38.6	-	-	*SHAPE[2]	ICCV2017	62.0	-	-
KNFST[61]	CVPR2016	42.3	82.9	92.1	NK3ML	Ours	**99.8**	**100**	**100**

Experiments on CUHK01 Dataset: We use GOG and LOMO* as the features for CUHK01. Each person of the dataset has two images in each camera view. Hence we report comparison with both single-shot and multi-shot settings in Tables 3a and b. NK3ML provides the state-of-the-art performances in all ranks. For single-shot setting, it outperforms the current best method GOG+XQDA [31] with a high margin of 9.20%. Similarly for multi-shot setting, NK3ML improves the accuracy by 9.49% for rank-1 over GOG+XQDA. At rank-1, NK3ML outperforms almost all of the pre/post processing based

methods also, except DLPAR [56] in single-shot setting, and Spindle [55] and SHaPE [2] for multi-shot setting. However, note that Spindle and DLPAR uses other camera domain information for training, and SHaPE is a re-ranking technique to aggregate scores from multiple metric learning methods. Also note that NK3ML even outperforms the deep learning based methods (see Table 4 also), emphasizing the limitation of deep learning based methods in re-ID systems with minimal training data.

Experiments on VIPeR Dataset: Concatenated GOG, LOMO* and WHOS are used as the features for VIPeR. It is the most widely accepted benchmark for person re-ID. It is a very challenging dataset as it contains images captured from outdoor environment with large variations in background, illumination and viewpoint. An enormous number of algorithms have reported results on VIPeR, with most of them reporting an accuracy below 50% at rank-1, as shown in Table 4. Even with the deep learning and pre/post processing re-ID methods, the best reported result for rank-1 is only 63.92% by DCIA [11]. On the contrary, NK3ML provides unprecedented improvement over these methods and attains a 99.8% rank-1 accuracy. The superior performance of NK3ML is due to its capability to enhance the discriminability even for the test data by simultaneously providing the maximal separation between the classes as well as minimizing the within class distance to the least value of zero.

Table 5. Comparison of execution time (in seconds) on VIPeR dataset

Methods	NK3ML	NFST	KNFST	XQDA	MLAPG	kLFDA	MFA	rPCCA
Training	1.64	1.47	0.37	1.35	12.10	4.10	3.68	23.98
Testing	0.37	0.34	0.33	0.34	0.13	4.13	3.99	3.74

4.3   Computational Requirements

We compare the execution time of NK3ML with other metric learning methods including NFST [61], KNFST [61], XQDA [23,31], MLAPG [24], kLFDA [51], MFA [51] and rPCCA [51] on VIPeR dataset. The details are shown in Table 5. The training time is calculated for the 632 samples in the training set, and the testing time is calculated for all the 316 queries in the test set. The training and testing time are averaged over 10 random trials. All methods are implemented in MATLAB on a PC with an Intel i7-6700 CPU@3.40 GHz and 32 GB memory. The testing time for NK3ML is 0.37 s for the set of 316 query images (0.0012 s per query), which is adequate for real time applications.

4.4   Application in Another Domain

In order to evaluate the applicability of NK3ML on other object verification problems also, we conduct experiments using LEAR ToyCars [33] dataset. It contains a total of 256 images of 14 distinct cars and trucks. The images have
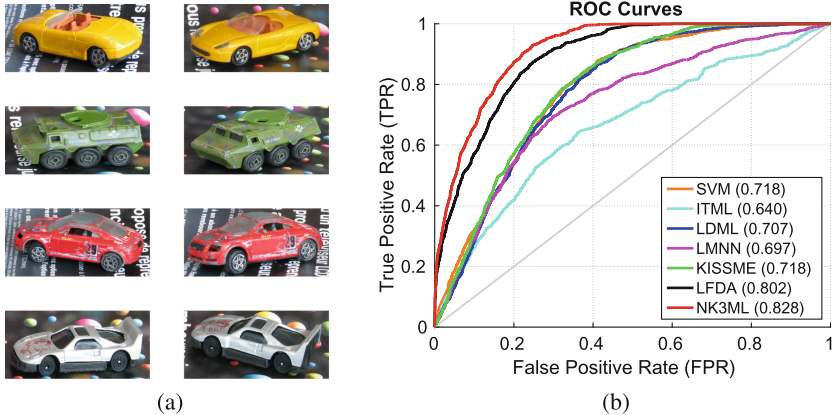
Fig. 4. ToyCars dataset (a) Sample images (b) ROC curves and EER comparisons.

wide variations in pose, illumination and background. The objective is to verify if a given pair of images are similar or not, even if they are *unseen* before. The training set has 7 distinct objects, provided as 1185 similar pairs and 7330 dissimilar pairs. The remaining 7 objects are used in the test set with 1044 similar pairs and 6337 dissimilar pairs. We use the feature representation from [19], which uses LBP with HSV and Lab histograms.

We compare the performance of NK3ML with the state-of-the-art metric learning methods including KISSME [19], ITML [9], LDML [13], LMNN [48,49], LFDA [36,44] and SVM [4]. Note that NK3ML and LMNN need the true class labels (not the similar/dissimilar pairs) for training. The proposed NK3ML learned a six dimensional subspace. For fair comparisons, we use the same features and learn an equal dimensional subspace for all the methods. We plot the Receiver Operator Characteristic (ROC) curves of the methods in Fig. 4, with the Equal Error Rate (EER) shown in parenthesis. NK3ML outperforms all other methods with a good margin. This experiment re-emphasizes that NK3ML is efficient to generalize well for unseen objects. Moreover, it indicates that NK3ML has the potential for other object verification problems also, apart from person re-identification.

5   Conclusions

In this work we presented a novel metric learning framework to efficiently address the small training sample size problem inherent in re-ID systems due to high dimensional data. We identify the suboptimality of NFST in generalizing to the test data. We provide a solution that minimizes the intra-class distance of training samples trivially to zero, as well as maximizes the inter-class distance to a much higher margin so that the learned discriminant vectors are effective in terms of generalization of the classifier performance for the test data also. Experiments on various challenging benchmark datasets show that our method outperforms

the state-of-the-art metric learning approaches. Especially, our method attains near human level perfection in the most widely accepted dataset VIPeR. We evaluate our method on another object verification problem also and validate its efficiency to generalize well to unseen data.

References

1. Bai, S., Bai, X., Tian, Q.: Scalable person re-identification on supervised smoothed manifold. In: CVPR (2017)
2. Barman, A., Shah, S.K.: Shape: a novel graph theoretic algorithm for making consensus-based decisions in person re-identification systems. In: ICCV (2017)
3. Bodesheim, P., Freytag, A., Rodner, E., Kemmler, M., Denzler, J.: Kernel null space methods in novelty detection. In: CVPR (2013)
4. Chang, C.C., Lin, C.J.: LIBSVM: a library for support vector machines. ACM Trans. Intell. Syst. Technol. **2**(3), 27:1–27:27 (2011)
5. Chen, D., Yuan, Z., Chen, B., Zheng, N.: Similarity learning with spatial constraints for person re-identification. In: CVPR (2016)
6. Chen, D., Yuan, Z., Hua, G., Zheng, N., Wang, J.: Similarity learning on an explicit polynomial kernel feature map for person re-identification. In: CVPR (2015)
7. Chen, W., Chen, X., Zhang, J., Huang, K.: Beyond triplet loss: a deep quadruplet network for person re-identification. In: CVPR (2017)
8. Cheng, D., Gong, Y., Zhou, S., Wang, J., Zheng, N.: Person re-identification by multi-channel parts-based CNN with improved triplet loss function. In: CVPR (2016)
9. Davis, J.V., Kulis, B., Jain, P., Sra, S., Dhillon, I.S.: Information-theoretic metric learning. In: ICML (2007)
10. Farenzena, M., Bazzani, L., Perina, A., Cristani, M., Murino, V.: Person re-identification by symmetry-driven accumulation of local features. In: CVPR (2010)
11. Garcia, J., Martinel, N., Micheloni, C., Gardel, A.: Person re-identification ranking optimisation by discriminant context information analysis. In: ICCV (2015)
12. Gray, D., Tao, H.: Viewpoint invariant pedestrian recognition with an ensemble of localized features. In: Forsyth, D., Torr, P., Zisserman, A. (eds.) ECCV 2008. LNCS, vol. 5302, pp. 262–275. Springer, Heidelberg (2008). https://doi.org/10.1007/978-3-540-88682-2_21
13. Guillaumin, M., Verbeek, J., Schmid, C.: Is that you? Metric learning approaches for face identification. In: ICCV (2009)
14. Guo, Y., Wu, L., Lu, H., Feng, Z., Xue, X.: Null Foley-Sammon transform. Pattern Recognit. **39**(11), 2248–2251 (2006)
15. Hirzer, M., Roth, P.M., Köstinger, M., Bischof, H.: Relaxed pairwise learned metric for person re-identification. In: Fitzgibbon, A., Lazebnik, S., Perona, P., Sato, Y., Schmid, C. (eds.) ECCV 2012. LNCS, vol. 7577, pp. 780–793. Springer, Heidelberg (2012). https://doi.org/10.1007/978-3-642-33783-3_56
16. Jose, C., Fleuret, F.: Scalable metric learning via weighted approximate rank component analysis. In: Leibe, B., Matas, J., Sebe, N., Welling, M. (eds.) ECCV 2016. LNCS, vol. 9909, pp. 875–890. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-46454-1_53

17. Kodirov, E., Xiang, T., Fu, Z., Gong, S.: Person re-identification by unsupervised ℓ_1 graph learning. In: Leibe, B., Matas, J., Sebe, N., Welling, M. (eds.) ECCV 2016. LNCS, vol. 9905, pp. 178–195. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-46448-0_11

18. Kviatkovsky, I., Adam, A., Rivlin, E.: Color invariants for person reidentification. IEEE TPAMI **35**(7), 1622–1634 (2013)

19. Kstinger, M., Wohlhart, P., Hirzer, H., Roth, P.M., Bischof, H.: Large scale metric learning from equivalence constraints. In: CVPR (2012)

20. Li, H., Jiang, T., Zheng, K.: Efficient and robust feature extraction by maximum margin criterion. In: NIPS (2004)

21. Li, H., Jiang, T., Zheng, K.: Efficient and robust feature extraction by maximum margin criterion. IEEE TNN **17**(1), 157–165 (2006)

22. Li, W., Zhao, R., Wang, X.: Human reidentification with transferred metric learning. In: Lee, K.M., Matsushita, Y., Rehg, J.M., Hu, Z. (eds.) ACCV 2012. LNCS, vol. 7724, pp. 31–44. Springer, Heidelberg (2013). https://doi.org/10.1007/978-3-642-37331-2_3

23. Liao, S., Hu, Y., Zhu, X., Li, S.Z.: Person re-identification by local maximal occurrence representation and metric learning. In: CVPR (2015)

24. Liao, S., Li, S.Z.: Efficient PSD constrained asymmetric metric learning for person re-identification. In: ICCV (2015)

25. Lisanti, G., Masi, I., Bimbo, A.D.: Matching people across camera views using kernel canonical correlation analysis. In: Proceedings of the International Conference on Distributed Smart Cameras. ACM (2014)

26. Lisanti, G., Masi, I., Bimbo, A.D.: Person re-identification by iterative re-weighted sparse ranking. IEEE TPAMI (2014)

27. Loy, C.C., Xiang, T., Gong, S.: Multi-camera activity correlation analysis. In: CVPR (2009)

28. Ma, L., Yang, X., Tao, D.: Person re-identification over camera networks using multi-task distance metric learning. IEEE TIP **23**(8), 3656–3670 (2014)

29. Martinel, N., Das, A., Micheloni, C., Roy-Chowdhury, A.K.: Temporal model adaptation for person re-identification. In: Leibe, B., Matas, J., Sebe, N., Welling, M. (eds.) ECCV 2016. LNCS, vol. 9908, pp. 858–877. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-46493-0_52

30. Martinel, N., Micheloni, C., Foresti, G.L.: Kernelized saliency-based person re-identification through multiple metric learning. IEEE TIP **24**(12), 5645–5658 (2015)

31. Matsukawa, T., Okabe, T., Suzuki, E., Sato, Y.: Hierarchical Gaussian descriptor for person re-identification. In: CVPR (2016)

32. Mignon, A., Jurie, F.: PCCA: A new approach for distance learning from sparse pairwise constraints. In: CVPR (2012)

33. Nowak, E., Jurie, F.: Learning visual similarity measures for comparing never seen objects. In: CVPR (2007)

34. Okada, T., Tomita, S.: An optimal orthonormal system for discriminant analysis. Pattern Recognit. **18**(2), 139–144 (1985)

35. Paisitkriangkrai, S., Shen, C., van den Hengel, A.: Learning to rank in person re-identification with metric ensembles. In: CVPR (2015)

36. Pedagadi, S., Orwell, J., Velastin, S., Boghossian, B.: Local Fisher discriminant analysis for pedestrian re-identification. In: CVPR (2013)

37. Roth, P.M., Hirzer, M., Köstinger, M., Beleznai, C., Bischof, H.: Mahalanobis distance learning for person re-identification. In: Gong, S., Cristani, M., Yan, S., Loy, C. (eds.) Person Re-Identification. Advances in Computer Vision and Pattern Recognition. Springer, London (2014). https://doi.org/10.1007/978-1-4471-6296-4_12

38. Sammon Jr., J.: An optimal discriminant plane. IEEE Trans. Comput. **100**(9), 826–829 (1970)

39. Shen, Y., Lin, W., Yan, J., Xu, M., Wu, J., Wang, J.: Person re-identification with correspondence structure learning. In: ICCV (2015)

40. Shi, H., et al.: Embedding deep metric for person re-identification: a study against large variations. In: Leibe, B., Matas, J., Sebe, N., Welling, M. (eds.) ECCV 2016. LNCS, vol. 9905, pp. 732–748. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-46448-0_44

41. Shi, Z., Hospedales, T.M., Xiang, T.: Transferring a semantic representation for person re-identification and search. In: CVPR (2015)

42. Su, C., Li, J., Zhang, S., Xing, J., Gao, W., Tian, Q.: Pose-driven deep convolutional model for person re-identification. In: ICCV (2017)

43. Su, C., Zhang, S., Xing, J., Gao, W., Tian, Q.: Deep attributes driven multi-camera person re-identification. In: Leibe, B., Matas, J., Sebe, N., Welling, M. (eds.) ECCV 2016. LNCS, vol. 9906, pp. 475–491. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-46475-6_30

44. Sugiyama, M.: Local Fisher discriminant analysis for supervised dimensionality reduction. In: ICML (2006)

45. Tao, D., Guo, Y., Song, M., Li, Y., Yu, Z., Tang, Y.Y.: Person re-identification by dual-regularized kiss metric learning. IEEE TIP **25**(6), 2726–2738 (2016)

46. Varior, R.R., Haloi, M., Wang, G.: Gated siamese convolutional neural network architecture for human re-identification. In: Leibe, B., Matas, J., Sebe, N., Welling, M. (eds.) ECCV 2016. LNCS, vol. 9912, pp. 791–808. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-46484-8_48

47. Varior, R.R., Shuai, B., Lu, J., Xu, D., Wang, G.: A siamese long short-term memory architecture for human re-identification. In: Leibe, B., Matas, J., Sebe, N., Welling, M. (eds.) ECCV 2016. LNCS, vol. 9911, pp. 135–153. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-46478-7_9

48. Weinberger, K.Q., Blitzer, J., Saul, L.K.: Distance metric learning for large margin nearest neighbor classification. In: NIPS (2006)

49. Weinberger, K.Q., Saul, L.K.: Fast solvers and efficient implementations for distance metric learning. In: ICML (2008)

50. Xiao, T., Li, H., Ouyang, W., Wang, X.: Learning deep feature representations with domain guided dropout for person re-identification. In: CVPR (2016)

51. Xiong, F., Gou, M., Camps, O., Sznaier, M.: Person re-identification using kernel-based metric learning methods. In: Fleet, D., Pajdla, T., Schiele, B., Tuytelaars, T. (eds.) ECCV 2014. LNCS, vol. 8695, pp. 1–16. Springer, Cham (2014). https://doi.org/10.1007/978-3-319-10584-0_1

52. Yang, Y., Yang, J., Yan, J., Liao, S., Yi, D., Li, S.Z.: Salient color names for person re-identification. In: Fleet, D., Pajdla, T., Schiele, B., Tuytelaars, T. (eds.) ECCV 2014. LNCS, vol. 8689, pp. 536–551. Springer, Cham (2014). https://doi.org/10.1007/978-3-319-10590-1_35

53. Yu, H.X., Wu, A., Zheng, W.S.: Cross-view asymmetric metric learning for unsupervised person re-identification. In: ICCV (2017)

54. Zhang, Y., Li, B., Lu, H., Irie, A., Ruan, X.: Sample-specific SVM learning for person re-identification. In: CVPR (2016)

55. Zhao, H., et al.: Spindle net: person re-identification with human body region guided feature decomposition and fusion. In: CVPR (2017)
56. Zhao, L., Li, X., Zhuang, Y., Wang, J.: Deeply-learned part-aligned representations for person re-identification. In: ICCV (2017)
57. Zhao, R., Ouyang, W., Wang, X.: Person re-identification by salience matching. In: ICCV (2013)
58. Zhao, R., Ouyang, W., Wang, X.: Unsupervised salience learning for person re-identification. In: CVPR (2013)
59. Zhao, R., Ouyang, W., Wang., X.: Learning mid-level filters for person re-identification. In: CVPR (2014)
60. Zheng, L., Wang, S., Tian, L., He, F., Liu, Z., Tian, Q.: Query-adaptive late fusion for image search and person reidentification. In: CVPR (2015)
61. Zheng, L., Xiang, T., Gong, S.: Learning a discriminative null space for person re-identification. In: CVPR (2016)
62. Zheng, S., Gong, S., Xiang, T.: Person re-identification by probabilistic relative distance comparison. In: CVPR (2011)
63. Zhong, Z., Zheng, L., Cao, D., Li, S.: Re-ranking person re-identification with k-reciprocal encoding. In: CVPR (2017)
64. Zhou, J., Yu, P., Tang, W., Wu, Y.: Efficient online local metric adaptation via negative samples for person re-identification. In: ICCV (2017)