Neural Process Lett
https://doi.org/10.1007/s11063-018-9820-x

QRKISS: A Two-Stage Metric Learning via QR-Decomposition and KISS for Person Re-Identification

Cairong Zhao[1,2] · Yipeng Chen[1] · Zhihua Wei[1] · Duoqian Miao[1] · Xinjian Gu[3]

Abstract Person re-identification is a challenging task in the field of intelligent video surveillance because there are wide variations between pedestrian images. As a classical metric learning method, Keep It Simple and Straightforward (KISS) has shown good performance for person re-identification. However, when the dimension of data is high, the KISS method may perform poorly because of small sample size problem. A common solution to this problem is to apply dimensionality reduction technologies to original data before the KISS metric learning, such as Principal Component Analysis (PCA) and Linear Discriminant Analysis (LDA). In this paper, to learn a discriminant and robust metric, we propose a novel two-stage metric learning via QR-Decomposition and KISS, named QRKISS. The first stage of QRKISS is to project original data into a lower dimensional space by QR decomposition. In this lower dimensional space, the trace of the covariance matrix of interpersonal differences can reach maximum. Based on KISS method, the second stage of QRKISS obtains a Mahalanobis matrix in the low-dimension space. We conduct thorough validation experiments on the VIPeR, PRID 450S and CUHK01 datasets, which demonstrate that QRKISS method is better than other KISS-based metric learning methods and achieves state-of-the-art performance.

Keywords Person re-identification · Metric learning · Dimension reduction · QR decomposition

Cairong Zhao and Yipeng Chen have contributed equally to this work.

✉ Cairong Zhao
zhaocairong@tongji.edu.cn

✉ Zhihua Wei
zhihua_wei@tongji.edu.cn

1 Department of Computer Science and Technology, Tongji University, Shanghai 201804, China

2 Fujian Provincial Key Laboratory of Information Processing and Intelligent Control (Minjiang University), Fuzhou, China

3 College of Information Science and Technology, Nanjing Agricultural University, Nanjing, China

Published online: 12 March 2018

1 Introduction

In recent years, person re-identification has received growing attention in the field of intelligent video surveillance [1–3]. Re-identification can be simply defined as "To re-identify a particular, then, is to identify it as (numerically) the same particular as one encountered on a previous occasion" [4]. In video surveillance, when being presented with a person-of-interest (query), person re-identification tells whether this person has been observed in another place (time) by another camera. It's a challenging task because there are some appearance changes between two pedestrian images, such as lighting, pose, and viewpoint, etc. Given its research and application significance, the re-ID community is fast growing. The main research of person re-identification can be divided into two categories: feature extraction [5–10] and metric learning [11–18].

Most of exciting studies of feature extraction focus on extracting robust and discriminant features from pedestrian images directly. The extracted features include various hand-crafted color, shape, texture features [5, 6, 8, 9], and some deeply learned features [7, 19, 20]. Regarding feature extraction for image representation, dimension reduction is critical to retain the most effective features for subsequent matching, since most of features are high-dimensional and redundant which would cause high computation complexity and low discriminability. Over the past decade, classical linear dimension reduction algorithm, and the emerging manifold learning algorithms, have enriched our choices for feature selection. The commonly used dimension reduction technologies include Principal Component Analysis (PCA) [21] and Linear Discriminant Analysis (LDA) [22]. PCA is a classical linear algorithm and LDA aims to separate samples drawn from different classes. In person re-identification, the procedure of dimension reduction is often merged into the procedure of metric learning.

As the second stage of person re-identification, metric learning plays an important role in performance improvement which aims to learn a discriminant distance function among the data. A comprehensive survey of the metric learning methods can be accessed in [15]. The metric learning methods can be categorized w.r.t supervised learning versus unsupervised learning, global learning versus local learning, etc. In person re-identification, the majority of works fall into the scope of supervised global distance metric learning. The metric learning algorithms can also be categorized according to their basic theories. For example, based on the theory of Fisher Discriminant Analysis (FDA), some metric learning methods such as, Local Fisher Discriminant Analysis (LFDA) [13], Marginal Fisher Analysis (MFA) [23] and Cross-view Quadratic Discriminant Analysis (XQDA) [6], were proposed. They tried to project descriptors into a more discriminatory subspace based on an eigenvalue resolution. By adding special constraints to the distance functions, some methods like Pairwise Constrained Component Analysis (PCCA) [24], Large Margin Nearest Neighbor (LMNN) [11] and Locally-Adaptive Decision Functions (LADF) [25] were proposed. Different constraints show different performances when they are applied to get a discriminatory metric distance function. In addition, some kernel-based metric learning methods also have been proposed to improve the identification ability of features [23, 26–30].

As a representative classical metric learning algorithm, Keep It Simple and Straightforward (KISS) method [12] was proposed based on maximum likelihood (ML) estimation. The KISS metric learning method is effective by considering a log likelihood ratio of two Gaussian distributions as the distance between a feature pair. The metric matrix of Mahalanobis distance is simply obtained by computing the difference between two inverse of covariance matrices. Thus, it is critical to estimate the covariance matrices accurately to improve the performance of KISS method. However, the estimated covariance matrices are inaccuracy when the dimension of data is much larger than the number of samples, which we call it small sample size (SSS) problem. It also arises in several covariance estimations-based metric learning algorithms, such as LFDA [13] and Null Foley-Sammon Transfer (NFST) [26], causing degradation of the performance. Moreover, when the dimension of data is high, the computation complexity of those covariance estimation-based methods would be high too. To overcome those problem, people always apply dimension reduction technique to map the original high-dimensional data into a low-dimensional space. For instance, there are two extensions of classical KISS method, PCA + KISS [12] and Cross-view Quadratic Discriminant Analysis (XQDA) [6], which combines the dimension reduction technique with classical KISS method. PCA + KISS utilized PCA to reduce the dimension of data before computing the covariance matrix. However, in the procedure of PCA, some discriminative information may be discarded. XQDA, by contrast, utilized LDA to reduce the dimension of data before KISS method, but the calculated transformation matrix of LDA is inaccuracy since the number of intrapersonal samples is much less than the interpersonal samples in person re-identification. Considering the estimation error of the small eigenvalues of the covariance matrices which arose through the SSS problem, researchers also proposed some extensions of classical KISS method based on regularization technique, such as Regularized Smoothing KISS (RS-KISS) [31] and Dual-regularized KISS (DR-KISS) [32], etc. RS-KISS seamlessly integrated smoothing and regularization techniques to estimate covariance matrices. DR-KISS simply regularized the covariance matrices by adding a multiple of identity matrix. However, all the regularization-based methods share a disadvantage that it is difficult to determine an optimal weight of regularization.

In this paper, focusing on metric learning and SSS problem, we propose a two-stage KISS extension, namely QRKISS. Similar to PCA + KISS and XQDA, we want to find an optimal transformation matrix that maps the high-dimensional data into a discriminant low-dimensional space. Instead of the dimension reduction technologies (PCA, LDA) used in PCA + KISS and XQDA, the first stage of QRKISS maps original data into a low-dimensional space via QR decomposition, which is simple and timesaving. In this lower dimensional space, the trace of the covariance matrix of interpersonal differences can reach maximum. Then, in the second stage of QRKISS, the difference between two inverse of covariance matrices is calculated to obtain a robust Mahalanobis matrix in the novel space.

Experiments with four kinds of features on different datasets show that our proposed QRKISS metric learning algorithm can improve the identification accuracy obviously and get better performance than other KISS-based metric learning methods in most of experiments, especially with the lack of training samples.

The main contributions of our work are summarized as the following three points:

1. A novel metric learning method called QRKISS is proposed for person re-identification, which combines the QR decomposition and KISS method.
2. We prove that utilizing the QR decomposition to reduce the dimension of training data can maximize the trace of covariance matrix of interpersonal differences.
3. We have thoroughly compared QRKISS with other KISS-based metric learning methods and the state-of-the-art metric learning methods on three public datasets (VIPeR, PRID 450S and CUHK01). Experimental results demonstrate that our proposed QRKISS metric learning method obtains better rank-1 matching rate and lower time consumption in most of the experiments.

The rest of the paper is organized as follows: in Sect. 2, we briefly review related works for person re-identification and we give an overview of KISS-based methods in Sect. 3. In

Sect. 4, we detail the proposed QRKISS method. Section 5 shows the experimental results on three representative datasets. Finally, we conclude the paper in Sect. 6.

2 Related Works

The proposed methods for person re-identification can be generally categorized into two groups which have been briefly described in Sect. 1. The first group of methods focuses on designing features which are robust to viewpoint change, illumination change and other variations between pedestrian's images. Based on the color, shape and texture information of pedestrian's appearance, several features have been proposed for person re-identification. Yang et al. [5] proposed a novel Salient Color Names Based Color Descriptor (SCNCD) which utilized the probability of sixteen salient color names as feature representation. Liao et al. [6] proposed an effective feature representation method called Local Maximal Occurrence (LOMO) which was composed of HSV histogram and the Scale Invariant Local Ternary Pattern (SILTP) [33] descriptor. There are some other hand-crafted features which have remained more or less the same in the recent years, such as Kernel Canonical Correlation Analysis (kCCA) descriptor [27], Gaussian Of Gaussian (GOG) [9], etc. Moreover, with the development of deep learning technique, more and more researchers used it to extract deep features from pedestrian images [7, 19, 20]. For example, Wu et al. [7] proposed a novel feature extraction model called Feature Fusion Net (FFN) for pedestrian image representation, which utilized the Convolutional Neural Network (CNN) to extract features. However, most of visual features do not capture all the invariant factors under sample variances, thus a good distance metric is critical for re-ID systems.

As mentioned above, the second group of methods for person re-identification focuses on metric learning, in detail it aims to learn a discriminant distance function among the data, keeping all the vectors of the same class closer while pushing vectors of different classes further apart. Several metric learning methods have been proposed and applied to person re-identification successfully. Weinberger et al. [11] proposed the Large-Margin Nearest Neighbor (LMNN) metric to improve the performance of the kNN classification. To avoid the overfitting problems encountered in LMNN, Davis et al. [34] proposed the Information-Theoretic Metric Learning (ITML) as a trade-off between two aspects, namely satisfying the given similarity constraints and ensuring that the learned metric is close to the initial distance function. In [24], an algorithm for learning distance metric called Pairwise Constrained Component Analysis (PCCA) was proposed. It projected the features into a low-dimensional space where the distance between pairs of data points respects the desired constraints. Based on the theory of Fisher Discriminant Analysis (FDA), Local Fisher Discriminant Analysis (LFDA) [13] was proposed to further reduce the dimensionality after an unsupervised PCA dimensionality reduction stage. To deal with the non-linearity in the appearance of pedestrians, Xiong et al. [23] proposed Kernel Local Fisher Discriminant Analysis (kLFDA), which utilized the kernel technique to reduce the influence of non-linearity. Moreover, Köstinger et al. [12] proposed a simplified and efficient metric learning method called Keep It Simple and Straightforward (KISS), which obtained the metric matrix of Mahalanobis distance by computing the difference between two inverses of covariance matrices. Because of its simpleness and effectiveness, more and more researchers pay attentions to KISS method and some improved methods based on KISS have been proposed. In [31], a regularized smoothing KISS metric learning (RS-KISS) was presented which seamlessly integrated smoothing and regularization techniques to estimate covariance matrices. Tao et al. [32] proposed a Dual-

Table 1 The important notations used in the paper

Notations	Descriptions
X / x_i	Training samples
d	Dimension of samples
x_{ij}	$x_i - x_j$
Ω_0	The set of interpersonal differences
Ω_1	The set of intrapersonal differences
Σ_0	The covariance matrices of Ω_0
Σ_1	The covariance matrices of Ω_1
$\mathbf{I}$	The identity matrix
trace()	The trace function

Regularized KISS (DR-KISS) metric learning by regularizing the two covariance matrices. Liong et al. [35] proposed a new Regularized Bayesian Metric Learning (RBML) method to model and regulate the eigen-spectrums of these two covariance matrices in a parametric manner. Liao et al. [6] proposed a metric learning method called Cross-view Quadratic Discriminant Analysis (XQDA), which combined the KISS and LDA algorithms for cross-view metric learning. Although the KISS-based metric learning methods have shown greatly performance for person re-identification, they still face covariance estimation problem and high computational complexity problem, especially when the dimension of data is very high.

In this paper, we focus on the second group of methods, i.e. metric learning, and we aim to explore a novel extension of the KISS method.

3 Overview of KISS Based Methods

In this section, we give a brief overview of classical KISS and its three extensions: PCA + KISS [12], DR-KISS [32] and XQDA [6]. For convenience, we present in Table 1 the important notations used in the paper.

3.1 Classical KISS

Köstinger et al. [12] proposed a simplified and efficient metric learning method called Keep It Simple and Straightforward (KISS), which has acquired the state-of-the-art retrieval performance for real-life applications, such as person re-identification and face recognition.

Given a feature vector pair x_i and x_j which represents two samples. Let H_0 denote the hypothesis that the sample pair (x_i, x_j) is dissimilar (x_i and x_j are extracted from different people), and H_1 denote the hypothesis that the feature pair (x_i, x_j) is similar (x_i and x_j are extracted from the same person). Formula (1) defines the logarithm of the ratio between two posteriors

$$\delta\left(x_i, x_j\right) = \log\left(\frac{p\left(H_0 | \left(x_i - x_j\right)\right)}{p\left(H_1 | \left(x_i - x_j\right)\right)}\right) \tag{1}$$

Let $x_{ij} = x_i - x_j$, formula (1) can be transformed into

$$\delta\left(x_i, x_j\right) = \log\left(\frac{p\left(H_0|x_{ij}\right)}{p\left(H_1|x_{ij}\right)}\right) \tag{2}$$

According to the Bayes formula, formula (2) can be rewritten as

$$\delta\left(x_i, x_j\right) = \log\left(\frac{p\left(x_{ij}|H_0\right)}{p\left(x_{ij}|H_1\right)}\right) + \log\left(\frac{p(H_0)}{p(H_1)}\right) \tag{3}$$

Assuming the difference space is a Gaussian structure, we have

$$p\left(x_{ij}|H_k\right) = \frac{1}{(2\pi)^{d/2}\left|\Sigma_k\right|^{1/2}} \exp\left(-\frac{1}{2}x_{ij}^T\Sigma_k^{-1}x_{ij}\right), \quad (k = 0, 1) \tag{4}$$

$$\Sigma_0 = \frac{1}{N_0}\sum_{y_{ij}=0} x_{ij}x_{ij}^T = \frac{1}{N_0}\sum_{y_{ij}=0}\left(x_i - x_j\right)\left(x_i - x_j\right)^T \tag{5}$$

$$\Sigma_1 = \frac{1}{N_1}\sum_{y_{ij}=1} x_{ij}x_{ij}^T = \frac{1}{N_1}\sum_{y_{ij}=1}\left(x_i - x_j\right)\left(x_i - x_j\right)^T \tag{6}$$

where y_{ij} denotes the indicated variable with respect to the two samples x_i and x_j, namely $y_{ij} = 1$ if x_i and x_j are extracted from the same person, otherwise $y_{ij} = 0$. Moreover, N_0 denotes the number of interpersonal differences, N_1 denotes the number of intrapersonal differences. Given formula (4), formula (3) can be rewritten as

$$\delta\left(x_i, x_j\right) = \frac{1}{2}x_{ij}^T\left(\Sigma_1^{-1} - \Sigma_0^{-1}\right)x_{ij} + \frac{1}{2}\log\left(\frac{|\Sigma_1|}{|\Sigma_0|}\right) + \log\left(\frac{p(H_0)}{p(H_1)}\right) \tag{7}$$

The smaller the value of $\delta(x_i, x_j)$ is, the bigger the probability that x_i and x_j belong to the same person. By dropping the constant terms, we have

$$\delta\left(x_i, x_j\right) = x_{ij}^T\left(\Sigma_1^{-1} - \Sigma_0^{-1}\right)x_{ij} \tag{8}$$

It is similar with the Mahalanobis distance in form, and thus the distance between two samples x_i and x_j is

$$d\left(x_i, x_j\right) = \delta\left(x_i, x_j\right) = \left(x_i - x_j\right)^T \mathbf{M}\left(x_i - x_j\right) \tag{9}$$

$$\mathbf{M} = \Sigma_1^{-1} - \Sigma_0^{-1} \tag{10}$$

3.2 Extensions of Classical KISS

Although the KISS metric learning method for discriminative distance metric learning has been shown to be effective for the person re-identification, the estimation of the inverse of a covariance matrix is unstable and indeed may not exist when the scale of the training set is small, resulting in poor performance. Several extensions including PCA + KISS [12], DR-KISS [32] and XQDA [6] were proposed in the past to deal with the singularity problems as follows.

3.2.1 PCA + KISS

A common way to deal with the singularity problem is to apply a dimension reduction algorithm, such as PCA, to reduce the dimension of the original data before classical KISS is applied. This is known as PCA + KISS. In this two-stage algorithm, the first stage aims to compute an optimal transformation matrix W that solves the following optimization problem:

$$W^* = \arg \max_{W^T W = I} \operatorname{trace}\left(W^T \bar{X} \bar{X}^T W\right) \tag{11}$$

where $\bar{X}$ denotes the centered data of X which has mean 0. The solution can be obtained by solving the eigenvalue problem on $\bar{X}\bar{X}^T$ [21].

In the second stage, KISS algorithm is applied in the novel low-dimensional space, and the distance between a feature vector pair (x_i, x_j) can be rewritten as

$$d\left(x_i, x_j\right) = \delta\left(x_i, x_j\right) = \left(x_i - x_j\right)^T W^* \left(\tilde{\Sigma}_1^{-1} - \tilde{\Sigma}_0^{-1}\right) W^{*T} \left(x_i - x_j\right) \tag{12}$$

where $\tilde{\Sigma}_0 = W^{*T} \Sigma_0 W^*$ and $\tilde{\Sigma}_1 = W^{*T} \Sigma_1 W^*$.

PCA + KISS algorithm has been used successfully in several models of person re-identification [5, 12, 14]. However, PCA is a time consuming process because of the huge computing requirement of SVD. Meanwhile, the optimal dimension of transformation matrix is difficult to determined, which may lead to some useful information lost in the dimension reduction stage.

3.2.2 DR-KISS

To deal with the singularity problems of two covariance matrices, Tao et al. [32] proposed a simple method called DR-KISS by regularizing the two covariance matrices. The two regularized covariance matrices can be obtained by following formulas:

$$\Sigma_{0,\gamma_0} = (1 - \gamma_0)\Sigma_0 + \gamma_0 \alpha_0 \mathbf{I} \tag{13}$$

$$\Sigma_{1,\gamma_1} = (1 - \gamma_1)\Sigma_1 + \gamma_1 \alpha_1 \mathbf{I} \tag{14}$$

where $\alpha_0 = (1/d)\operatorname{trace}(\Sigma_0)$, $\alpha_1 = (1/d)\operatorname{trace}(\Sigma_1)$, $0 < \gamma_0 < 1$ and $0 < \gamma_1 < 1$.

DR-KISS improves on KISS by reducing overestimation of large eigenvalues of the two estimated covariance matrices and, in doing so, guarantees that the covariance matrix is reversible. Furthermore, [32] provided theoretical analyses for supporting the motivations and proved why the regularization is necessary. A limitation of DR-KISS is that the optimal values of the two parameters γ_0 and γ_1 are difficult to determine.

3.2.3 XQDA

Similar to PCA + KISS, Liao et al. [6] try to reduce the dimension of the original data before classical KISS is applied, and they proposed a novel metric learning method called Cross-view Quadratic Discriminant Analysis (XQDA). XQDA is different with PCA + KISS in the procedure of dimension reduction, since XQDA aims to solve the following optimization problem:

$$W^* = \arg \max_{W^T W = I} \operatorname{trace}\left(\frac{W^T \Sigma_0 W}{W^T \Sigma_1 W}\right) \tag{15}$$

This optimization problem can be solved by the generalized eigenvalue decomposition problem as similar in LDA.

After obtaining the optimal transformation matrix W^*, KISS algorithm is applied in the novel low-dimensional space, and the distance between a feature vector pair (x_i, x_j) can be rewritten as

$$d\left(x_i, x_j\right) = \delta\left(x_i, x_j\right) = \left(x_i - x_j\right)^T W^* \left(\tilde{\Sigma}_1^{-1} - \tilde{\Sigma}_0^{-1}\right) W^{*^T} \left(x_i - x_j\right)$$

where $\tilde{\Sigma}_0 = W^{*T}\Sigma_0 W^*$ and $\tilde{\Sigma}_1 = W^{*T}\Sigma_1 W^*$.

XQDA have been successfully used in the application of person re-identification, however, the calculated transformation matrix W^* may have contained some inaccuracy, since the number of intrapersonal samples is much less than the interpersonal samples in person re-identification.

4 Proposed Method

Classical KISS metric learning method based on covariance matrix estimation may performs poorly especially when the dimension of data is large, since the inverse of covariance matrices are inaccurate or non-existent. In view of this, we want to find an optimal transformation matrix W that can transform the high-dimensional original data into a low-dimensional space, in where the estimated covariance matrix would be more accuracy. Aiming at this issue, in this section, we propose a two-stage metric learning via QR decomposition and KISS, namely, QRKISS.

4.1 QRKISS: A Two-Stage Metric Learning

The procedure of QRKISS algorithm can be divided into two steps: (1) Obtaining an optimal transformation matrix W via QR Decomposition. (2) Expanding the classical KISS algorithm in the novel low-dimensional space.

According to classical KISS method, formula (5) and formula (6) represent the way of computing two covariance matrices (Σ_0 and Σ_1) respectively. They can be rewritten as

$$\Sigma_k = \frac{1}{N_k} X\left(D^k - M^k\right) X^T, \quad (k = 0, 1) \tag{16}$$

where

$$M_{ij}^0 = \begin{cases} 0 & if\ y_i = y_j \\ 1 & if\ y_i \neq y_j \end{cases} \tag{17}$$

$$M_{ij}^1 = \begin{cases} 1 & if\ y_i = y_j\ and\ i \neq j \\ 0 & otherwise \end{cases} \tag{18}$$

D^k denotes a diagonal matrix (Each element sums all affinity values over the columns of M^k, such that $D_{ii}^k = \sum_{j=1}^N M_{ij}^k$), N_k denotes the number of value '1' in the matrix M^k.

Given a data matrix $X \in \Re^{d \times N}$, where d denotes the dimension of data, and N denotes the number of samples. We consider finding a linear transformation $W \in \Re^{d \times l}$ that maps each column x_i of X, for $1 \le i \le N$, to a l-dimensional space as $W^T x_i \in \Re^l (l < d)$. Thus, the novel low-dimensional data can be represented by $W^T X$. Therefore, the formula for computing the covariance matrices in the classical KISS algorithm can be rewritten as

$$\Sigma_k = W^T X\left(D^k - M^k\right) X^T W, \quad (k = 0, 1) \tag{19}$$

Considering the number of intrapersonal samples is much less than the interpersonal samples in person re-identification, we try to obtain the optimal transformation matrix W^* without utilizing the covariance matrix of intrapersonal differences Σ_1. Thus, the first stage of QRKISS aims to compute the optimal transformation matrix W^* that can maximize the trace of novel covariance matrix in the low-dimensional space as follows

$$W^* = \arg \max_{W^T W = I} \operatorname{trace}\left(W^T X \left(D^0 - M^0\right) X^T W\right). \tag{20}$$

According to the QR decomposition [36], the data matrix X can be decomposed to the product of two matrices, such as

$$X = QR \tag{21}$$

where $Q \in \Re^{d\times N}$ has orthonormal columns, $R \in \Re^{N\times N}$ is upper triangular.

Let's multiply both sides by Q^T, we can obtain

$$Q^T X = Q^T QR \tag{22}$$

Considering the matrix Q has orthonormal columns, the formula (22) can be rewritten as

$$Q^T X = IR = R \tag{23}$$

where I denotes a N-dimensional identify matrix.

Comparing the formula (23) with the dimension reduction formula $W^T X$, we can find that both formulas have similar forms. Thus we can consider the matrix Q as the linear transformation W, and the matrix R as the novel low-dimensional data matrix $W^T X$. Next, we will give a prove to show $W^* = Q$ solves the optimization problem in formula (20).

Lemma 1 *Let $A \in \Re^{n\times n}$ be positive semidefinite and $W \in \Re^{n\times q}$ have orthogonal columns, where $q \leq n$. The following inequality holds*

$$\operatorname{trace}\left(W^T AW\right) \leq \operatorname{trace}(A).$$

Proof Let $\tilde{W} \in \Re^{n\times(n-q)}$ be the matrix such that $\left[W, \tilde{W}\right]$ is orthogonal. That is,

$$\left[W, \tilde{W}\right] \cdot \left[W, \tilde{W}\right]^T = WW^T + \tilde{W}\tilde{W}^T = I_n \tag{24}$$

where $I_n \in \Re^{n\times n}$ is the identity matrix.

It follows that

$$\begin{aligned} \operatorname{trace}\left(W^T AW\right) &= \operatorname{trace}\left(AWW^T\right) = \operatorname{trace}(A) - \operatorname{trace}\left(A\tilde{W}\tilde{W}^T\right) \\ &= \operatorname{trace}(A) - \operatorname{trace}\left(\tilde{W}^T A\tilde{W}\right) \leq \operatorname{trace}(A) \end{aligned} \tag{25}$$

where the last inequality follows since $\tilde{W}^T A\tilde{W}$ is positive semidefinite. This completes the proof of the Lemma 1. □

Theorem 1 *Let $X = QR$ be the QR Decomposition of X, where $Q \in \Re^{d\times N}$ has orthonormal columns and $R \in \Re^{N\times N}$ is upper triangular. Then,*

$$W^* = QM,$$

for any orthogonal matrix M, solves the optimization problem in formula (20).

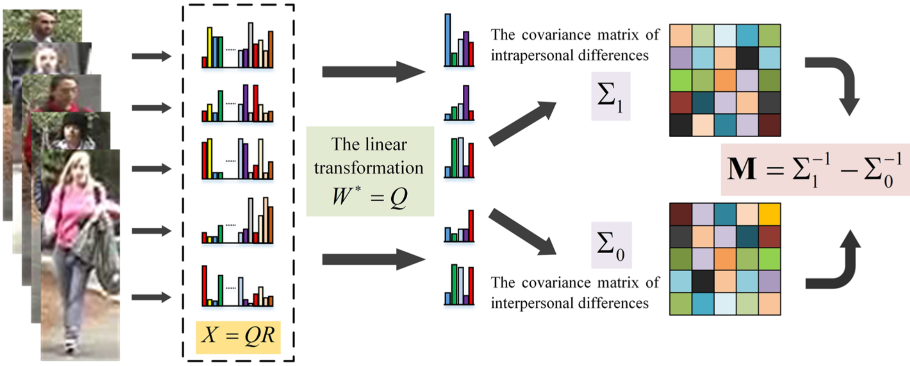

Fig. 1 The flow chart of QRKISS metric learning method

Proof Let $X = QR$ be the QR Decomposition of X. Then, we have

$$X(D^0 - M^0)X^T = QR(D^0 - M^0)R^T Q^T = QEQ^T \tag{26}$$

where $E = R(D^1 - M^1)R^T$.

For any W with orthonormal columns, it is clear that

$$\begin{aligned} \mathrm{trace}(W^T X(D^0 - M^0)X^T W) &\leq \mathrm{trace}(X(D^0 - M^0)X^T) = \mathrm{trace}(QEQ^T) \\ &= \mathrm{trace}(EQ^T Q) = \mathrm{trace}(E) \end{aligned} \tag{27}$$

where the first inequality from Lemma 1 mentioned above. Thus, trace(E) is an upper bound for the optimization in formula (20).

Next, we show that the upper bound is achieved by choosing $W^* = QM$ for any orthogonal M. By the property of trace and the fact that Q has orthonormal columns, we have

$$\begin{aligned} \mathrm{trace}\left(\left(W^*\right)^T X\left(D^0 - M^0\right) X^T W^*\right) &= \mathrm{trace}\left(M^T Q^T QR\left(D^0 - M^0\right) R^T Q^T QM\right) \\ &= \mathrm{trace}\left(M^T R\left(D^0 - M^0\right) R^T M\right) = \mathrm{trace}\left(M^T EM\right) \\ &= \mathrm{trace}\left(EMM^T\right) = \mathrm{trace}(E) \end{aligned} \tag{28}$$

This completes the proof of the theorem. □

Note the choice of orthogonal matrix M is arbitrary since $trace(W^T X(D^0 - M^0)X^T W) = trace(M^T W^T X(D^0 - M^0)X^T WM)$, for any orthogonal matrix M. For convenience, M is set to be the identity matrix and finally we can obtain the optimal transformation matrix $W^* = Q$.

After obtaining the optimal transformation matrix $W^* = Q$, we can compute the distance between a pair vector x_i and x_j according to the theory of KISS, and the distance function is

$$d\left(x_i, x_j\right) = \left(x_i - x_j\right)^T Q\left(\tilde{\Sigma}_1^{-1} - \tilde{\Sigma}_0^{-1}\right) Q^T \left(x_i - x_j\right) \tag{29}$$

$$\tilde{\Sigma}_k = Q^T X\left(D^k - M^k\right) X^T Q, \quad (k = 0, 1) \tag{30}$$

Figure 1 shows the flow chart of QRKISS metric learning method and the complete algorithm is summarized in Algorithm 1.

Algorithm 1 QRKISS

Input: Training data X, Training labels Y and testing samples $(z_i)_{i=1}^n$

Output: The distance between two test samples z_i and z_j

Stage1

1: Calculate the D^0, D^1, M^0, M^1 in Eq. (16) according to the training labels Y;

2: Calculate Q according to Eq. (21) (QR Decomposition);

Stage2

3: Calculate $\tilde{\Sigma}_0$ and $\tilde{\Sigma}_1$ according to Eq. (30);

4: Calculate the distance between two test samples z_i and z_j according to Eq. (29)

4.2 The Differences Between QRKISS, PCA + KISS and XQDA

Same as QRKISS, both PCA + KISS and XQDA try to find a transformation matrix to reduce the dimension of data. But there are some differences among them.

QRKISS aims to compute an optimal transformation matrix W that solves the following optimization problem:

$$W_{QR} = \arg \max_{W^T W = I} \text{trace}\left(W^T \Sigma_0 W\right) \tag{31}$$

The solution can be obtained by utilizing the QR Decomposition for the original data X. It is a simple and quick step that leads QRKISS to be less time-consuming than other KISS-based methods. Moreover, QRKISS can cope with the SSS problem in person re-identification better, since QR Decomposition can reduce the dimension of features effectively.

By contrast, KISS + PCA aims to compute an optimal transformation matrix W_{PCA} that solves the following optimization problem:

$$W_{PCA} = \arg \max_{W^T W = I} \text{trace}\left(W^T \bar{X} \bar{X}^T W\right) \tag{32}$$

where $\bar{X}$ denotes the centered data of X which has mean 0. The solution can be obtained by solving the eigenvalue problem on XX^T. It's easy to prove that the solution of optimization problem of PCA + KISS method W_{PCA} is not suitable for the optimization problem of QRKISS, and $\text{trace}(W_{PCA}^T \Sigma_k W_{PCA}) \leq \text{trace}(W_{QR}^T \Sigma_k W_{QR}), \quad (k = 0, 1)$.

XQDA aims to compute an optimal transformation matrix W_{XQDA} that solves the following optimization problem:

$$W_{XQDA} = \arg \max_{W^T W = I} \text{trace}\left(\frac{W^T \Sigma_0 W}{W^T \Sigma_1 W}\right) \tag{33}$$

which can be solved by the generalized eigenvalue decomposition problem as similar in LDA. According to the theory of Fisher Discriminant Analysis (FDA), it seems that the value of $\text{trace}(W^T \Sigma_1 W)$ should be as small as possible. However, in the real application of person re-identification, the number of intrapersonal samples is much less than interpersonal samples, which would cause the inaccuracy of solution when the optimization problem of XQDA is solved by the generalized eigenvalue decomposition. Compared with XQDA method, QRKISS can cope with the SSS problem better, since it doesn't utilize the imprecise covariance matrix of the intrapersonal differences Σ_1.

Fig. 2 Example images from VIPeR dataset [37]

Moreover, compared from the aspect of time complexity, QRKISS is more efficient than KISS + PCA and XQDA, since the procedures of eigenvalue decomposition of PCA and LDA are time-consumed.

In conclusion, comparing the two classical KISS-based metric learning methods (PCA + KISS and XQDA), our proposed QRKISS metric learning method has mainly two advantages: (1) QRKISS can cope with the SSS problem in person re-identification better. (2) QRKISS is less time consuming. Both advantages have been verified in the next section.

5 Experiments

In this section, we conducted thorough validation experiments on three challenging person re-identification datasets to demonstrate the effectiveness of our proposed QRKISS method.

5.1 Datasets and Settings

5.1.1 Datasets

Three widely used challenging datasets, VIPeR [37], PRID 450S [14], and CUHK01 [38], were used for experiments.

VIPeR [37] is a challenging person re-identification dataset which has been widely used for performance evaluation. The dataset contains 1264 images of 632 pedestrians totally, each pedestrian has two images captured form two different cameras. All the images are cropped and scaled to a resolution of 128×48. Figure 2 shows some example images from this dataset. The VIPeR dataset is randomly divided into two parts, namely one half for training and the other half for testing in the experiments.

PRID 450S dataset [14] consists of 450 images pairs of pedestrians with significant differences in background, viewpoint, and illumination. It has been recently released and is regarded as more realistic. Some example images from PRID 450S dataset are shown in

Fig. 3 Example images from PRID 450S dataset [14]

Fig. 3. Same as the experimental protocol on the VIPeR dataset, the PRID 450S dataset was randomly divided into two parts, namely one half for training and the other half for testing in the experiments.

CUHK01 [38] is another challenging person re-identification dataset which contains 971 pedestrians. Each pedestrian has four images captured from two different camera views in a campus environment. Camera A captures the frontal view or back view of people while camera B captures the side view. Large inter-camera variations are observed in this dataset which makes person re-identification challenging. Figure 4 shows some example images from this dataset. Unlike the experimental settings of VIPeR dataset and PRID 450S dataset, the experiments of CUHK01 dataset can be divided into CUHK01(M = 1) and CUHK01(M = 2) because each pedestrian has more than two images. The experimental setting of CUHK01(M = 1) is the single shot setting and CUHK01(M = 2) is the multi shot setting, which are common to [39] and [6] respectively.

5.1.2 Feature Representations

In our experiments, four kinds of features, LOMO [6], FFN [7], SCNCD [5] and ELF18 [8], were used for all metric learning methods in this experiment. Table 2 shows the comparison of four different features.

5.1.3 Evaluation Metrics

We used Cumulated Matching Characteristics (CMC) curve to evaluate the performance of person re-identification methods for all datasets in this paper. All the experiments were repeated 20 times to record an average performance. Because the complexity of the re-identification problem, the top n-ranked matching rate was considered (n is a small value). In this paper, 1-ranked, 5-ranked, 10-ranked and 20-ranked matching rates were selected for compared.

Fig. 4 Example images from CUHK01 dataset [38]

Table 2 The comparison of four different features used in our experiments

Features	LOMO	FFN	SCNCD	ELF18
Dimension	26960	8064	70	8064
Type	Hand-crafted	Deep learned	Hand-crafted	Hand-crafted
Year	2015	2016	2014	2015

5.2 Experimental Results and Analysis

5.2.1 Experiments on VIPeR Dataset

In this section, we compare our proposed QRKISS metric learning method with other state-of-the-art metric learning methods, including five KISS-based methods (XQDA [6], DR-KISS [32], RBML [35], RS-KISS [31] and PCA+KISS [12]) and five distance metric learning-

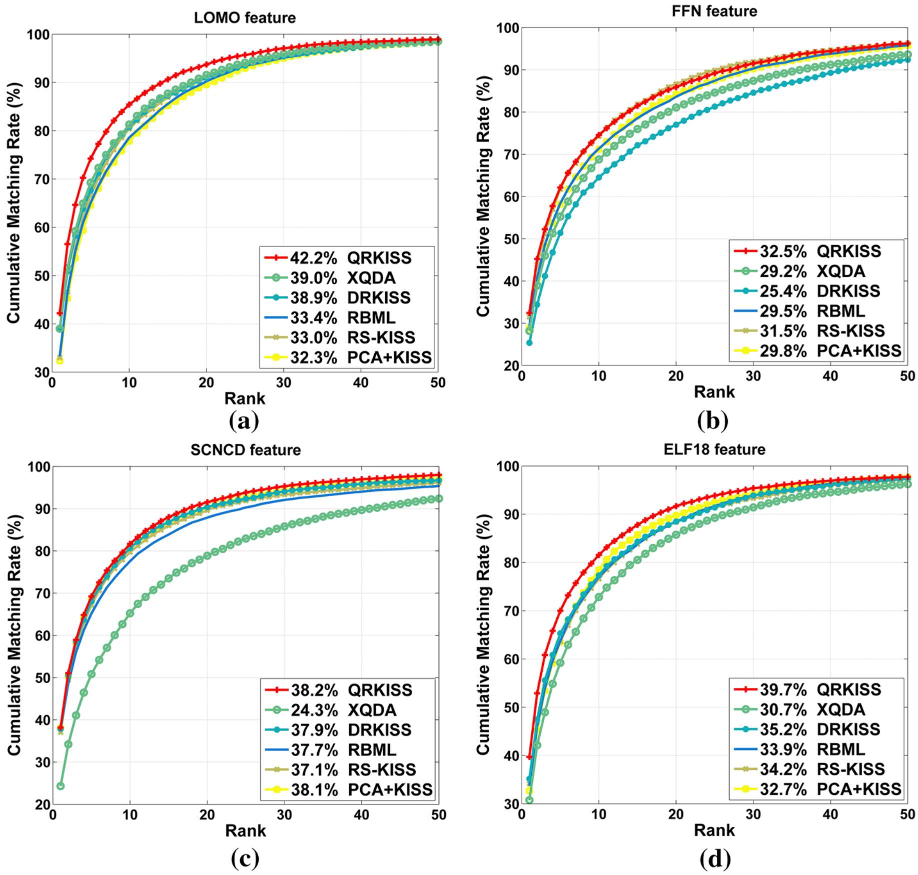

Fig. 5 CMC curves on the VIPeR dataset (P = 316). **a** Comparison of KISS based metric learning methods by using the LOMO feature. **b** Comparison of KISS based metric learning methods by using the FFN feature. **c** Comparison of KISS based metric learning methods by using the SCNCD feature. **d** Comparison of KISS based metric learning methods by using the ELF18 feature

based methods (NFST [26], MLAPG [40], MFA [23], kLFDA [23] and LFDA [13]). Four kinds of features (LOMO [6], FFN [7], SCNCD [5] and ELF18 [8]) were used for all metric learning methods in the experiments.

Figure 5 shows the comparison of our proposed QR-KISS metric learning with five KISS-based methods. In each subfigure, the x-coordinate is the rank score, and the y-coordinate is the matching rate. Only the top 50 ranking positions are shown in the figure. Among the compared KISS-based methods, XQDA and PCA + KISS are discriminative subspace learning-based methods, whilst the others (DR-KISS, RBML, RS-KISS) are regularization-based methods. As we can see from Fig. 5, the matching accuracies of QRKISS have significantly increased in all experiments, which demonstrate that the QR Decomposition-based subspace learning play a distinctive role to improve the performance of classical KISS method.

Moreover, the cumulative matching scores at rank 1, 5, 10 of all compared state-of-the-art metric learning methods are listed in Table 3. Bold and italics numbers are the best and second-best results, respectively. The experimental results in Table 3 show that our

Table 3 Comparison of different metric learning methods on the VIPeR database (P = 316)

Method	Feature												Reference
	LOMO (d = 26960)			FFN (d = 4096)			SCNCD (d = 70)			ELF18 (d = 8064)			
	r = 1	r = 5	r = 10	r = 1	r = 5	r = 10	r = 1	r = 5	r = 10	r = 1	r = 5	r = 10	
QRKISS	**42.2**	**74.2**	**85.4**	**32.5**	**62.2**	**74.6**	**38.2**	**69.3**	**81.7**	**39.7**	*70.0*	81.6	Ours
NFST [26]	*40.1*	*70.3*	82.3	*31.8*	*62.0*	*74.5*	36.7	67.6	80.0	*37.8*	**70.3**	**84.0**	2016 CVPR
MLAPG [40]	39.5	70.0	*82.4*	12.3	30.9	42.0	25.8	57.6	73.1	18.0	44.7	61.5	2015 ICCV
XQDA [6]	39.0	69.3	81.3	29.2	55.3	68.8	24.3	50.9	65.2	30.7	59.2	72.9	2015 CVPR
DR-KISS [32]	38.9	67.7	80.8	25.4	51.4	64.5	37.9	68.0	80.7	35.2	65.3	77.3	2015 IEEE
MFA [23]	35.4	67.9	81.0	29.7	59.4	73.3	37.9	*68.9*	*81.4*	36.8	**70.3**	*83.0*	2014 ECCV
kLFDA [23]	36.3	65.1	77.1	23.0	43.4	53.1	36.9	68.5	81.1	32.6	60.7	71.7	2014 ECCV
RBML [35]	33.4	65.1	78.6	29.5	58.4	71.3	37.7	65.2	77.6	33.9	63.8	77.1	2014 ECCV
LFDA [13]	37.6	68.7	81.3	30.7	56.7	69.5	16.3	29.8	36.6	33.1	63.3	76.8	2013 CVPR
RS-KISS [31]	33.0	66.7	80.4	31.5	61.7	74.3	37.1	67.2	79.7	34.2	63.6	76.6	2013 IEEE
PCA + KISS [12]	32.3	64.6	77.9	29.8	58.2	71.2	*38.1*	68.3	80.8	32.7	63.5	78.5	2012 CVPR

The cumulative matching scores (%) at rank 1, 5, and 10 are listed. Bold and italics numbers are the best and second-best results, respectively

Table 4 Training time (s) of metric learning algorithms

Methods	QRKISS	NFST	MLAPG	XQDA	DR-KISS	MFA
Time (s)	2.91	2.98	40.90	3.86	1383.45	3.74
Methods	kLFDA	RBML	LFDA	RS-KISS	KISS + PCA	
Time (s)	3.24	9.47	229.34	8.03	7.41	

proposed QRKISS metric learning method achieves the best performance in the experiments of three features (LOMO, FFN, and SCNCD), and obtains the best rank-1 accuracy in the experiment of ELF18 feature, which demonstrates that the performance of QRKISS metric learning method achieves the state-of-the-art level on VIPeR dataset. It is worth mentioning that the FFN feature is extracted by deep learning method, and the other three features are hand-crafted. The experimental results demonstrate that QRKISS metric learning method is effective for both hand-crafted feature and deep feature.

Meanwhile, the average training time comparison of metric learning methods is shown in Table 4 (including subspace learning time). The training was performed on a notebook PC with an Intel i5-3210 @2.50 GHz CPU. Table 4 shows that QRKISS is the quickest metric learning method in all compared state-of-the-art methods. This is because in the procedure of QRKISS, the original data would be projected into a lower dimensional space by QR decomposition, while the QR decomposition is not time-consuming.

5.2.2 Experiments on PRID 450S Dataset

In this section, we compared our proposed method with five KISS-based metric learning methods and some other state-of-the-art metric learning methods on the PRID 450S dataset [14]. The features we used in this experiment were the same with the experiment on VIPeR dataset. The compared metric learning methods includes NFST, MLAPG, XQDA, DR-KISS, MFA, kLFDA, RBML, LFDA, RS-KISS and PCA + KISS.

Figure 6 shows the comparison of our proposed QR-KISS metric learning with other KISS-based methods. As we can see from the Fig. 6, our proposed QRKISS method achieves better rank-1 matching rate than other KISS-based methods, which demonstrates that QRKISS is an enhancement of the classical KISS method, with better performance and lower time consumption compared with state-of-the-art methods.

Table 5 shows the experimental results of comparison of different metric learning methods. The experimental results in Table 5 show that our proposed QRKISS metric learning method achieves the best rank-1 accuracy in the experiments of two features (LOMO and ELF18) and obtains the second-best performance in the experiments of FFN feature and SCNCD feature, which demonstrates that QRKISS metric learning method is effective for both hand-crafted feature and deep feature, and its performance has reached state-of-the-art. As a kernel-based metric learning method, the NFST method achieves the top performance in the experiments of LOMO feature and FFN feature because of the non-linearity in the appearance of pedestrian. Despite all this, QRKISS method still obtains better rank-1 accuracy than NFST in the experiment of LOMO feature, which is considered to be the most important performance index for person re-identification.

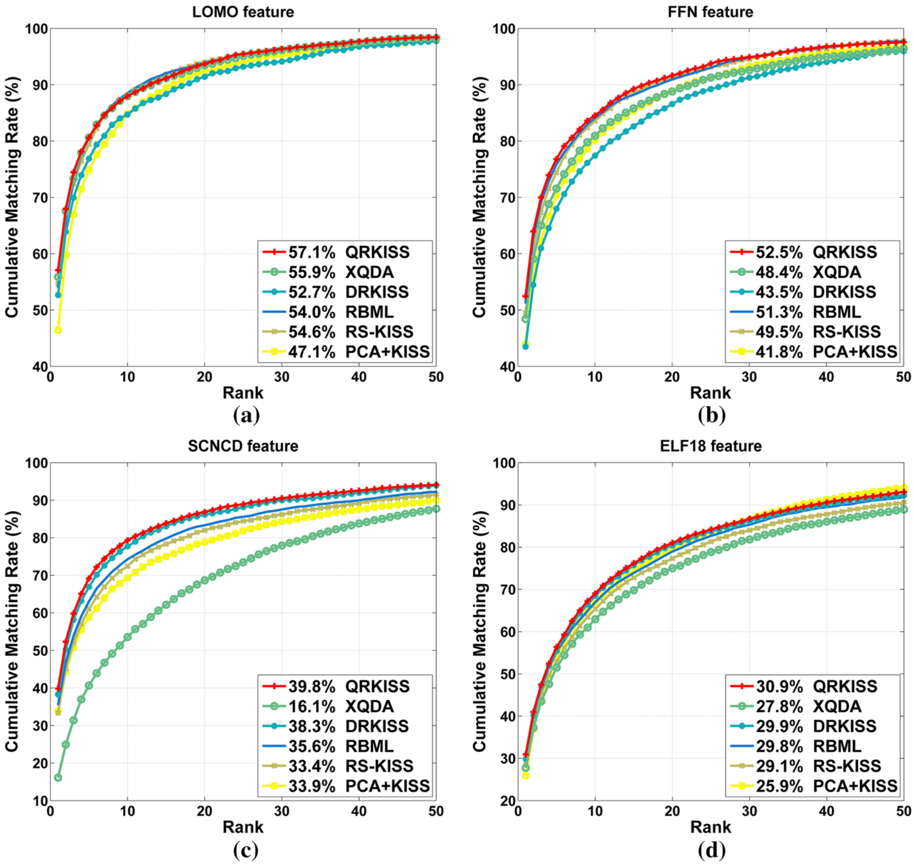

Fig. 6 CMC curves on the PRID 450S dataset (P = 225). **a** Comparison of KISS based metric learning methods by using the LOMO feature. **b** Comparison of KISS based metric learning methods by using the FFN feature. **c** Comparison of KISS based metric learning methods by using the SCNCD feature. **d** Comparison of KISS based metric learning methods by using the ELF18 feature

5.2.3 Experiments on CUHK01 Dataset

In this section, we compared our proposed method with five KISS based metric learning methods and some other state-of-the-art metric learning methods on CUHK01 dataset. The features we used in this experiment include LOMO [6] and FFN [7].

The experimental matching rate curves of all KISS-based methods are shown in Fig. 7 and the experimental results of all state-of-the-art metric learning methods are shown in Table 6. As we can see in Fig. 7 and Table 6, our proposed QRKISS method achieves good performance in the experiments of LOMO feature, but doesn't perform well in the experiments of FFN feature. There are two reasons for that, one of which is that the dimension of FFN feature (4096) is less than most of features. Another reason is that there are more training images and more interpersonal feature pairs in CUHK01 dataset, thus the SSS problem causes less impact on each metric learning method when the FFN feature is used. To prove this point, we reduce the number of training pedestrian image pairs from 486 to 200, and the new experimental results of all state-of-the-art metric learning methods are shown in Table 7. In

Table 5 Comparison of different metric learning methods on the PRID 450S database (P = 225)

Method	Feature												Reference
	LOMO (d = 26960)			FFN (d = 4096)			SCNCD (d = 70)			ELF18 (d = 8064)			
	r = 1	r = 5	r = 10	r = 1	r = 5	r = 10	r = 1	r = 5	r = 10	r = 1	r = 5	r = 10	
QRKISS	**57.1**	*80.7*	88.0	*52.5*	*76.8*	*84.6*	*39.8*	*69.2*	*79.3*	**30.9**	**56.4**	*69.0*	Ours
NFST [26]	*56.0*	**81.9**	**89.1**	**53.5**	**78.8**	**87.4**	32.0	57.7	68.6	29.5	54.1	67.4	2016 CVPR
MLAPG [40]	55.5	80.1	88.3	22.7	46.5	57.9	30.4	60.3	73.6	8.8	23.8	35.2	2015 ICCV
XQDA [6]	55.9	80.6	88.0	48.4	71.6	80.9	16.1	40.7	53.6	27.8	51.5	63.0	2015 CVPR
DR-KISS [32]	52.7	76.9	84.7	43.5	68.0	77.4	38.3	66.9	77.7	29.9	*55.4*	68.5	2015 IEEE
MFA [23]	50.8	78.7	87.7	47.9	75.1	*84.6*	**41.6**	**70.5**	**80.6**	30.0	57.2	**69.7**	2014 ECCV
kLFDA [23]	51.3	78.8	88.2	48.6	75.7	84.5	20.9	39.3	49.4	9.1	23.9	32.5	2014 ECCV
RBML [35]	54.0	*80.7*	*88.6*	51.3	75.9	84.1	35.6	63.0	74.3	29.8	55.3	67.0	2014 ECCV
LFDA [13]	54.1	79.4	87.8	49.9	75.0	83.8	28.9	50.9	62.1	*30.5*	55.1	66.9	2013 CVPR
RS-KISS [31]	54.6	79.5	88.1	49.5	74.4	83.6	33.4	61.1	72.4	29.1	52.9	65.4	2013 IEEE
PCA + KISS [12]	47.1	75.5	85.6	41.8	69.4	79.7	33.9	58.9	69.3	25.9	53.6	66.7	2012 CVPR

The cumulative matching scores (%) at rank 1, 5 and 10 are listed. The best and the second-best scores are respectively shown in bold and italics

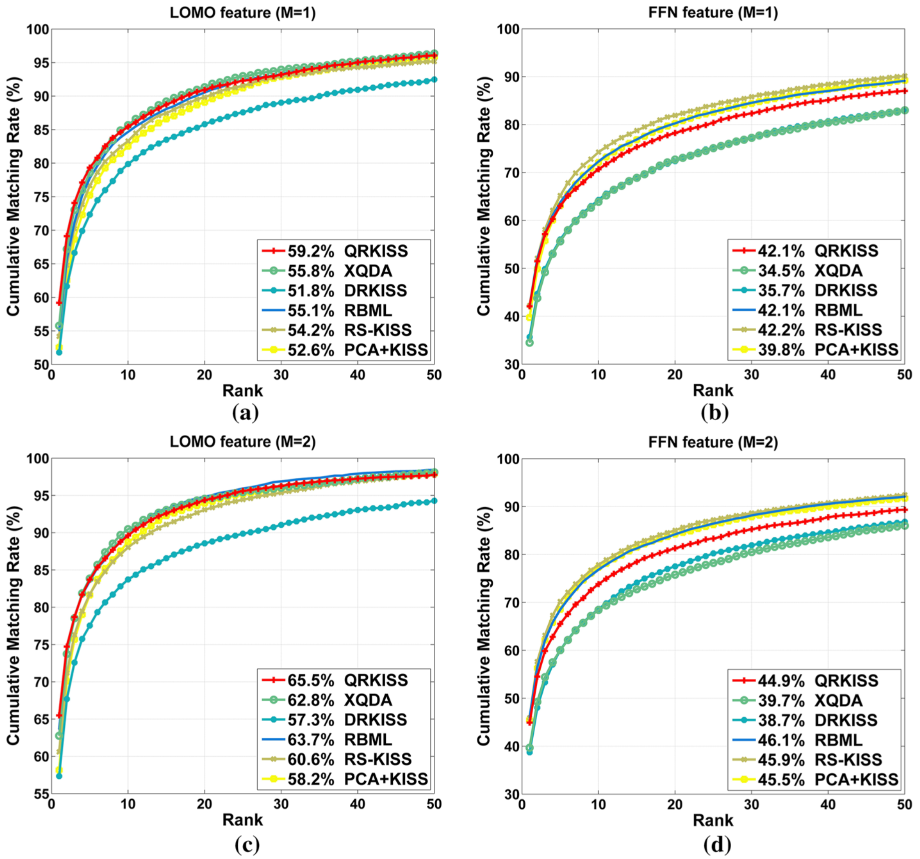

Fig. 7 CMC curves on the CUHK01 dataset (P = 486). **a** Comparison of KISS based metric learning methods with the single shot setting by using the LOMO feature. **b** Comparison of KISS based metric learning methods with the single shot setting by using the FFN feature. **c** Comparison of KISS based metric learning methods with the multi shot setting by using the LOMO feature. **d** Comparison of KISS based metric learning methods with the multi shot setting by using the FFN feature

Table 7, compared with other state-of-the-art metric learning methods, the QRKISS method achieves the best performance. The experimental results indicate that the QRKISS metric learning method is not only suit for the single-shot cases in person re-identification, but also suit for the multi-shot cases, especially with the lack of training samples.

6 Conclusions and Future Works

In this paper, we have presented a novel two-stage metric learning method for person re-identification called QRKISS. Compared with the classical KISS method, the QRKISS method can improve the performance of KISS method by utilizing QR decomposition to reduce the dimension of data, and thus the computation complexity of QRKISS method would be much lower than KISS method. The experiments on three publicly datasets, VIPeR, PRID 450S and CUHK01, show that the QRKISS method is better than state-of-the-art met-

Table 6 Comparison of different metric learning methods on the CUHK01 database (P = 486)

Method	Feature												Reference
	M = 1						M = 2						
	LOMO (d = 26960)			FFN (d = 4096)			LOMO (d = 26960)			FFN (d = 4096)			
	r = 1	r = 5	r = 10	r = 1	r = 5	r = 10	r = 1	r = 5	r = 10	r = 1	r = 5	r = 10	
QRKISS	*59.2*	79.3	85.4	*42.1*	63.0	70.7	**65.5**	83.7	89.6	44.9	65.6	73.8	Ours
NFST [26]	**60.0**	*80.8*	86.7	39.4	60.6	69.0	*65.0*	85.0	89.9	43.4	66.5	74.4	2016 CVPR
MLAPG [40]	58.4	79.0	85.5	24.8	46.5	56.9	64.7	**86.6**	*91.6*	26.8	50.3	61.3	2015 ICCV
XQDA [6]	55.8	78.6	85.7	34.5	55.6	63.9	62.8	83.9	90.5	39.7	60.1	68.4	2015 CVPR
DR-KISS [32]	51.8	72.3	79.9	35.7	56.0	64.3	57.3	77.6	83.7	38.7	60.1	68.7	2015 IEEE
MFA [23]	58.7	**81.2**	**88.1**	39.6	*64.2*	*73.9*	63.8	*85.7*	**91.9**	44.7	**70.6**	**79.9**	2014 ECCV
kLFDA [23]	58.9	*80.8*	*86.9*	32.0	53.8	64.2	64.7	84.0	90.4	34.1	58.5	69.0	2014 ECCV
RBML [35]	55.1	77.8	84.7	*42.1*	63.7	72.3	63.7	83.5	89.8	**46.1**	68.5	76.9	2014 ECCV
LFDA [13]	55.6	77.9	85.1	39.3	63.6	73.1	62.9	83.5	89.7	44.8	68.7	77.7	2013 CVPR
RS-KISS [31]	54.2	76.7	83.3	**42.2**	**65.2**	**74.3**	60.6	81.8	88.0	*45.9*	*70.2*	*77.9*	2013 IEEE
PCA + KISS [12]	52.6	75.2	82.5	39.8	63.1	72.1	58.2	81.7	88.8	45.5	68.6	77.3	2012 CVPR

The cumulative matching scores (%) at rank 1, 5 and 10 are listed. The best and the second-best scores are respectively shown in bold and italics

Table 7 Comparison of different metric learning methods on the CUHK01 database (P = 200)

Method	Feature						Reference
	FFN (d = 4096)						
	M = 1			M = 2			
	r = 1	r = 5	r = 10	r = 1	r = 5	r = 10	
QRKISS	**33.7**	**55.7**	**64.4**	**37.5**	**60.2**	**69.6**	Ours
NFST [26]	29.7	52.2	62.1	34.0	57.9	68.1	2016 CVPR
MLAPG [40]	15.9	34.2	43.2	18.2	37.3	48.3	2015 ICCV
XQDA [6]	28.0	47.5	56.5	31.8	52.9	61.9	2015 CVPR
DR-KISS [32]	29.8	50.4	58.4	32.2	52.7	61.2	2015 IEEE
MFA [23]	26.9	48.8	58.8	32.0	57.4	67.3	2014 ECCV
kLFDA [23]	22.4	42.2	52.2	25.2	47.5	58.5	2014 ECCV
RBML [35]	30.3	51.2	60.6	33.7	56.2	65.8	2014 ECCV
LFDA [13]	*30.5*	*52.9*	*62.6*	*34.7*	*58.6*	*68.3*	2013 CVPR
RS-KISS [31]	28.2	49.2	59.3	33.0	56.7	66.3	2013 IEEE
PCA + KISS [12]	24.6	45.6	55.3	30.3	56.4	63.7	2012 CVPR

The cumulative matching scores (%) at rank 1, 5 and 10 are listed. The best and the second-best scores are respectively shown in bold and italics

ric learning methods with the top performance and the lowest time consumption, especially with the small size samples. It would be interesting to see that the QRKISS method can be applied to other cross-view matching problem, such as face recognition.

Acknowledgements The authors would like to thank the anonymous reviewers for their critical and constructive comments and suggestions. This work is supported by China National Natural Science Foundation under Grant No. 61673299, 61203247 and The Hong Kong Polytechnic University (Project account: G-UC42). This work is also partially supported by China National Natural Science Foundation under Grant No. 61573259, 61573255, 61375012. It is also supported by the Fundamental Research Funds for the Central Universities (Grant No. 0800219327). It is also partially supported by Fujian Provincial Key Laboratory of Information Processing and Intelligent Control (Minjiang University) and by the Open Project Program of Key Laboratory of Intelligent Perception and Systems for High-Dimensional Information of Ministry of Education under Grant No. 30920130122005. It is also partially supported by the program of Further Accelerating the Development of Chinese Medicine Three Year Action of Shanghai Grant No. ZY3-CCCX-3-6002.

References

1. Vezzani R, Baltieri D, Cucchiara R (2013) People reidentification in surveillance and forensics: a survey. ACM Comput Surv CSUR 46(2):29
2. Bedagkar-Gala A, Shah SK (2014) A survey of approaches and trends in person re-identification. Image Vis Comput 32(4):270–286
3. Zheng L, Yang Y, Hauptmann AG (2016) Person re-identification: past, present and future. arXiv preprint arXiv:1610.02984
4. Plantinga A (1961) Things and persons. Rev Metaphys 14(3):493–519
5. Yang Y, Yang J, Yan J et al (2014) Salient color names for person re-identification. In: European conference on computer vision. Springer, Cham, pp 536–551

6. 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
7. Wu S, Chen Y C, Li X et al (2016) An enhanced deep feature representation for person re-identification. In: 2016 IEEE Winter Conference on Applications of Computer Vision (WACV). IEEE, pp 1–8
8. Chen YC, Zheng WS, Lai J (2015) Mirror representation for modeling view-specific transform in person re-identification. In: Twenty-fourth international joint conference on artificial intelligence
9. Matsukawa T, Okabe T, Suzuki E et al (2016) Hierarchical gaussian descriptor for person re-identification. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 1363–1372
10. Shen F, Yang Y, Li Liu, Liu W, Tao D, Shen H (2017) Asymmetric binary coding for image search. IEEE Trans Multimed TMM 19(9):2022–2032
11. Weinberger KQ, Saul LK (2009) Distance metric learning for large margin nearest neighbor classification. J Mach Learn Res 10(Feb):207–244
12. Koestinger M, Hirzer M, Wohlhart P et al (2012) Large scale metric learning from equivalence constraints. In: 2012 IEEE conference on computer vision and pattern recognition (CVPR). IEEE, pp 2288–2295
13. Pedagadi S, Orwell J, Velastin S et al (2013) Local fisher discriminant analysis for pedestrian re-identification. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 3318–3325
14. Roth P M, Hirzer M, Koestinger M et al (2014) Mahalanobis distance learning for person re-identification. In: Person re-identification. Springer London, pp 247–267
15. Yang L, Jin R (2006) Distance metric learning: a comprehensive survey. Mich State Univ 2(2):1–51
16. Shen F, Shen C, Shi Q et al (2015) Hashing on nonlinear manifolds. IEEE Trans Image Process 24(6):1839–1851
17. Shen F, Zhou X, Yang Y et al (2016) A fast optimization method for general binary code learning. IEEE Trans Image Process 25(12):5610–5621
18. Lai Z, Wong W, Xu Y, Zhao C, Sun M (2014) Sparse alignment for robust tensor learning. IEEE Trans Neural Netw Learn Syst 25(10):1779–1792
19. Li W, Zhao R, Xiao T et al (2014) Deepreid: Deep filter pairing neural network for person re-identification. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 152–159
20. Varior RR, Shuai B, Lu J et al (2016) A siamese long short-term memory architecture for human re-identification. In: European conference on computer vision. Springer, Berlin, pp 135–153
21. Wold S, Esbensen K, Geladi P (1987) Principal component analysis. Chemom Intell Lab Syst 2(1–3):37–52
22. Mika S, Ratsch G, Weston J et al (1999) Fisher discriminant analysis with kernels. In: Neural networks for signal processing IX, 1999. Proceedings of the 1999 IEEE signal processing society workshop. IEEE, pp 41–48
23. Xiong F, Gou M, Camps O et al (2014) Person re-identification using kernel-based metric learning methods. In: European conference on computer vision. Springer, Cham, pp 1–16
24. Mignon A, Jurie F (2012) PCCA: a new approach for distance learning from sparse pairwise constraints. In: 2012 IEEE conference on computer vision and pattern recognition (CVPR), pp 2666–2672
25. Li Z, Chang S, Liang F et al (2013) Learning locally-adaptive decision functions for person verification. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 3610–3617
26. Zhang L, Xiang T, Gong S (2016) Learning a discriminative null space for person re-identification. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 1239–1248
27. Lisanti G, Masi I, Del Bimbo A (2014) Matching people across camera views using kernel canonical correlation analysis. In: Proceedings of the international conference on distributed smart cameras. ACM, p 10
28. Wang W, Taalimi A, Duan K et al (2016) Learning patch-dependent kernel forest for person re-identification. In: 2016 IEEE winter conference on applications of computer vision (WACV). IEEE, pp 1–9
29. Lisanti G, Karaman S, Masi I (2017) Multichannel-kernel canonical correlation analysis for cross-view person reidentification. ACM Trans Multi Comput Commun Appl TOMM 13(2):13
30. Liu H, Qi M, Jiang J (2015) Kernelized relaxed margin components analysis for person re-identification. IEEE Signal Process Lett 22(7):910–914
31. 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
32. Tao D, Guo Y, Song M et al (2016) Person re-identification by dual-regularized kiss metric learning. IEEE Trans Image Process 25(6):2726–2738

33. Liao S, Zhao G, Kellokumpu V et al (2010) Modeling pixel process with scale invariant local patterns for background subtraction in complex scenes. In: 2010 IEEE conference on computer vision and pattern recognition (CVPR), pp 1301–1306. IEEE
34. Davis JV, Kulis B, Jain P et al (2007) Information-theoretic metric learning. In: Proceedings of the 24th international conference on machine learning, pp 209–216
35. Liong V E, Lu J, Ge Y (2014) Regularized Bayesian metric learning for person re-identification. In: ECCV workshops (3), pp 209–224
36. Golub GH, Van Loan CF (2012) Matrix computations, vol 3. JHU Press, Baltimore
37. Gray D, Brennan S, Tao H (2007) Evaluating appearance models for recognition, reacquisition, and tracking. In: Proceedings of the IEEE international workshop on performance evaluation for tracking and surveillance (PETS), vol 3, No 5. Citeseer, pp 1–7
38. Li W, Zhao R, Wang X (2012) Human reidentification with transferred metric learning. In: Asian conference on computer vision. Springer, Berlin, pp 31–44
39. Paisitkriangkrai S, Shen C, van den Hengel A (2015) Learning to rank in person re-identification with metric ensembles. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 1846–1855
40. Liao S, Li SZ (2015) Efficient psd constrained asymmetric metric learning for person re-identification. In: Proceedings of the IEEE international conference on computer vision, pp 3685–3693