Kernel Hierarchical PCA for Person Re-Identification

Raphael Felipe Prates and William Robson Schwartz
Smart Surveillance Interest Group, Computer Science Department
Universidade Federal de Minas Gerais, Minas Gerais, Brazil

***Abstract*—Person re-identification (Re-ID) maintains a global identity for an individual while he moves along a large area covered by multiple cameras. Re-ID enables a multi-camera monitoring of individual activity that is critical for surveillance systems. However, the low-resolution images combined with the different poses, illumination conditions and camera viewpoints make person Re-ID a challenging problem. To reach a higher matching performance, state-of-the-art methods map the data to a nonlinear feature space where they learn a cross-view matching function using training data. Kernel PCA is a statistical method that learns a common subspace that captures most of the variability of samples using a small number of vector basis. However, Kernel PCA disregards that images were captured by distinct cameras, a critical problem in person Re-ID. Differently, Hierarchical PCA (HPCA) captures a consensus projection between multiblock data (e.g, two camera views), but it is a linear model. Therefore, we propose the Kernel Hierarchical PCA (Kernel HPCA) to tackle camera transition and dimensionality reduction in a unique framework. To the best of our knowledge, this is the first work to propose a kernel extension to the multiblock HPCA method. Experimental results demonstrate that Kernel HPCA reaches a matching performance comparable with state-of-the-art nonlinear subspace learning methods at PRID450S and VIPeR datasets. Furthermore, Kernel HPCA reaches a better combination of subspace learning and dimensionality requiring significantly lower subspace dimensions.**

I. Introduction

Person re-identification (Re-ID) aims at maintaining a unique identity for an individual while he moves across a large area covered by multiple cameras with non-overlapping field-of-views (FOV). Re-ID matches a probe image captured by a camera A with stored images captured by a surveillance camera B (gallery set). Thus, person Re-ID allows the supervision of subjects actions at indoor and outdoor environments using a reduced number of surveillance cameras. Despite its relevance for security and safety management, Re-ID is still an open problem that has attracted researches attention in the recent years [1], [2].

The infrastructure of installed surveillance cameras provides low-resolution images where biometric cues, such as face and iris, are unreliable. Therefore, most of the works in literature use clothing appearance as the unique information to perform person Re-ID. However, the same individuals captured by two distinct cameras may have a dissimilar appearances due the variations in resolution, pose, illumination and camera viewpoint. These camera transition problems make person Re-ID a challenging problem.

Re-ID methods address the camera transition problem using a cross-view matching function [3]–[11] and robust feature descriptors [12]–[17]. Works that focus on the feature representation seek for subtle characteristics that are robust to different camera conditions and discriminative, usually hand-crafted descriptors. Then, they match probe and gallery images using a similarity function computed without labeled image pairs from both cameras (training data), which is known as the *unsupervised Re-ID* setting. On the other hand, the *supervised Re-ID* methods use training data to learn a deep representation [18], a metric distance [3], [4], [7] or projections to a common subspace [5], [6], [19]. Particularly, recent approaches have reported improved performance by considering a nonlinear mapping to a high-dimensional feature space through a kernel function [6], [19], which can be explained by the impossibility of linear models to deal with appearance changes caused by camera transitions.

Principal Component Analysis (PCA) is a statistical method for unsupervised dimensionality reduction that linearly project the data to a common latent subspace that explains most of the variance among samples using a smaller set of vector basis [20]. Classical applications of PCA in computer vision include eigenfaces [21], eigengait [22] and eigentracking [23]. However, while many computer vision problems present a strong nonlinear behavior, PCA is a linear model. Therefore, Schölkopf et *al.* [20] proposed an extension of PCA that nonlinearly maps the data at a high-dimensional feature space and computes the principal components using a "kernel trick", which is known as Kernel PCA (KPCA).

Despite the success of PCA, there are few works that address person Re-ID problem using PCA-based methods [24], [25]. Martinel and Micheloni [25] use PCA to project image dissimilarities onto a common subspace where they learn a binary classifier to discriminate between same and not-same classes. Differently, Yang et *al.* [24] learn for each gallery image a nonlinear eigenspace appearance representation using Kernel PCA and pixelwise color and spatial information. Then, they project the probe image at each gallery-specific subspace and compute the similarity using latent scores. However, the direct comparison between probe and gallery images disregards that these images were captured by distinct cameras. In fact, the small number of samples at each class and the camera transition problem compose a challenging scenario to the application of conventional PCA-based methods in person Re-ID problem.

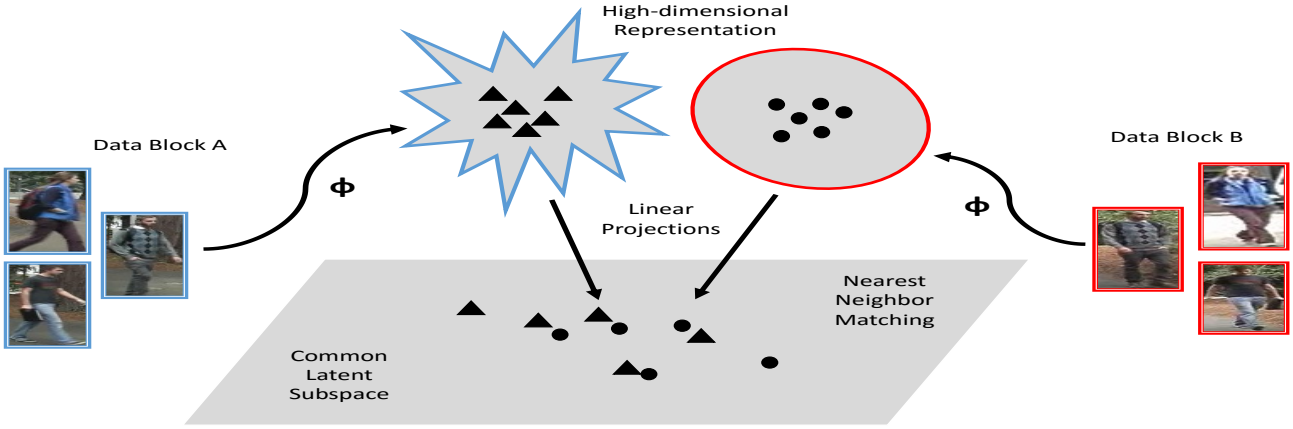

Fig. 1. Schematic representation of the novel Kernel HPCA method. Given two blocks of data (A and B), Kernel HPCA computes a nonlinear mapping function (Φ) to a high-dimensional space where linear projections to a low-dimensional common subspace are computed. These projections, which seek for a consensual direction between data blocks, are efficiently computed using a "kernel trick". Therefore, we can successfully match samples from different blocks projecting them onto the learned subspace.

A possible solution to tackles the person Re-ID problem is to use multiblock multivariate models [26] to learn a common subspace where the direct comparison between probe and gallery images results in a high matching performance. These models have been employed when additional information is available for grouping variables in a meaningful blocks (e.g, different camera views). Consensus PCA (CPCA) [27] and Hierarchical PCA (HPCA) [28] are two multiblock extensions of PCA that seek for a consensus direction among all the blocks and are useful to compare several descriptors measured on the same object. Despite being common in chemometrics [28], CPCA and HPCA methods are not well known in computer vision community. To the best of our knowledge, the method proposed by Qiu et *al.* [29] to detect changes in satellite images using principal components is the unique effort to tackle a computer vision problem using multiblock PCA.

We evaluate CPCA and HPCA as statistical methods to learn a common subspace where we match probe and gallery images using a simple nearest neighbor method. Experimental results demonstrate that both methods reach results comparable with state-of-the-art linear subspace learning methods at two widely used datasets (VIPeR and PRID450S). However, since CPCA and HPCA are linear models and images of the same person at two distinct cameras indicate a strong nonlinear relationship, we propose a nonlinear extension of Hierarchical PCA, which we name *Kernel Hierarchical PCA* (Kernel HPCA).

The method proposed in this work is a nonlinear extension of multiblock HPCA that pursues a consensus direction between block of variables (see Fig. 1). Differently from HPCA, this direction is computed in a high-dimensional feature space obtained by a nonlinear transformation of input variables. Therefore, Kernel HPCA captures high-order correlation between input variables when learning projections to a common subspace that tackles dimensionality reduction and camera transition problems. Furthermore, we propose an efficient algorithm to compute the principal components using the "kernel trick" and Nonlinear Iterative PArtial Least Squares (NIPALS) algorithm. Experimental results demonstrate that Kernel HPCA represents a great improvement when compared to Kernel PCA and reaches comparable results with a state-of-the-art kernel subspace learning methods.

To describe the multiblock PCA (Section II) and the novel Kernel HPCA (Section III) methods we use the following notation. Bold lower-case letters denote vectors and bold upper-case letters denote matrices (e. g., $\mathbf{z}$ and $\mathbf{Z}$, respectively). Furthermore, we represent the training data from data blocks A and B using matrices $\mathbf{X}_a,\ \mathbf{X}_b \in \mathbb{R}^{n \times m}$, respectively, where the *ith* row in $\mathbf{X}_a$ and $\mathbf{X}_b$ correspond to feature descriptors extracted from the *ith* object. In person Re-ID, that means we have labeled image pairs from both cameras, corresponding to the supervised Re-ID scenario. At testing stage, we represent a new sample from data block A and B (e.g, images captured by the same camera) as vectors $\hat{\mathbf{x}}_a, \hat{\mathbf{x}}_b \in \mathbb{R}^m$, respectively.

II. Consensus PCA and Hierarchical PCA

Despite being widely employed in chemometrics and biochemical process monitoring [26], Consensus PCA (CPCA) and Hierarchical PCA (HPCA) are not well known by the pattern recognition community. Therefore, this section provides an introduction to these multiblock PCA methods. A more deep discussion can be found in [26].

CPCA and HPCA are useful when there is a meaningful division of data into blocks. A typical example corresponds to multiple measures of the same object (e.g, images of the same subject captured by multiple cameras). To enable a better understanding, we present a two-block NIPALS algorithm and the description of these methods. However, it is important to highlight that these methods are able to handle multiple data blocks, not limited to two.

Let $\mathbf{X}_a$ and $\mathbf{X}_b$ be two blocks of data, CPCA starts with a super score $\mathbf{t} \in \mathbb{R}^n$, which is an initial guess of consensus and can be a column of these blocks. Then, this super score is regressed on blocks $\mathbf{X}_a$ and $\mathbf{X}_b$ to compute the loadings $\mathbf{w}_a,\ \mathbf{w}_b \in \mathbb{R}^m$. Loadings $\mathbf{w}_a$ and $\mathbf{w}_b$ are normalized to length one to compute the respective block scores $\mathbf{s}_a,\ \mathbf{s}_b \in \mathbb{R}^n$ that are combined in a super block $\mathbf{S} \in \mathbb{R}^{n \times 2}$. The super score $\mathbf{t}$

Algorithm 1: Consensus PCA (CPCA)

input : $\mathbf{X}_a$, $\mathbf{X}_b$ matrices and the number of factors (f)

randomly initialize $\mathbf{t}$
for *i=1* **to** f **do**
  while ***t*** *do not converge* **do**
    $\mathbf{w}_a = \mathbf{X}_a^\top \mathbf{t}/\mathbf{t}^\top \mathbf{t} \quad \mathbf{w}_a \leftarrow \frac{\mathbf{w}_a}{\|\mathbf{w}_a\|}$
    $\mathbf{w}_b = \mathbf{X}_b^\top \mathbf{t}/\mathbf{t}^\top \mathbf{t} \quad \mathbf{w}_b \leftarrow \frac{\mathbf{w}_b}{\|\mathbf{w}_b\|}$
    $\mathbf{s}_a = \mathbf{X}_a \mathbf{w}_a$
    $\mathbf{s}_b = \mathbf{X}_b \mathbf{w}_b$
    $\mathbf{S} = [s_a, s_b]$
    $\mathbf{w} = \mathbf{S}^\top \mathbf{t}/\mathbf{t}^\top \mathbf{t} \quad \mathbf{w} \leftarrow \frac{\mathbf{w}}{\|\mathbf{w}\|}$
    $\mathbf{t} = \mathbf{S}\mathbf{w}$
  end
  $\mathbf{w}_a = \mathbf{X}_a^\top \mathbf{t}/\mathbf{t}^\top \mathbf{t}$
  $\mathbf{w}_b = \mathbf{X}_b^\top \mathbf{t}/\mathbf{t}^\top \mathbf{t}$
  $\mathbf{X}_a = \mathbf{X}_a - \mathbf{t}\mathbf{w}_a^\top$
  $\mathbf{X}_b = \mathbf{X}_b - \mathbf{t}\mathbf{w}_b^\top$
end

is regressed on the super block $\mathbf{S}$ to obtain the super weight $\mathbf{w} \in \mathbb{R}^m$, which is normalized to length one. Then, we update the value of super score $\mathbf{t}$ by projecting the super block $\mathbf{S}$ onto super weight $\mathbf{w}$. This process repeats until the convergence of super score $\mathbf{t}$ to a predefined precision. Thus, while block scores ($\mathbf{s}_a$ and $\mathbf{s}_b$) are resultant of block variables, the super score $\mathbf{t}$ is derived using all variables. After the convergence, block variables $\mathbf{X}_a$ and $\mathbf{X}_b$ are deflated to obtain a new score vector at the next iterations.

Algorithm 2: Hierarchical PCA (HPCA)

input : $\mathbf{X}_a$, $\mathbf{X}_b$ matrices and the number of factors (f)

randomly initialize $\mathbf{t}$
for *i=1* **to** f **do**
  while ***t*** *do not converge* **do**
    $\mathbf{w}_a = \mathbf{X}_a^\top \mathbf{t}$
    $\mathbf{s}_a = \mathbf{X}_a \mathbf{w}_a \quad \mathbf{s}_a \leftarrow \frac{\mathbf{s}_a}{\|\mathbf{s}_a\|}$
    $\mathbf{w}_b = \mathbf{X}_b^\top \mathbf{t}$
    $\mathbf{s}_b = \mathbf{X}_b \mathbf{w}_b \quad \mathbf{s}_b \leftarrow \frac{\mathbf{s}_b}{\|\mathbf{s}_b\|}$
    $\mathbf{S} = [s_a, s_b]$
    $\mathbf{w} = \mathbf{S}^\top \mathbf{t}$
    $\mathbf{t} = \mathbf{S}\mathbf{w} \quad \mathbf{t} \leftarrow \frac{\mathbf{t}}{\|\mathbf{t}\|}$
  end
  $\mathbf{X}_a = \mathbf{X}_a - \mathbf{t}\mathbf{w}_a^\top$
  $\mathbf{X}_b = \mathbf{X}_b - \mathbf{t}\mathbf{w}_b^\top$
end

NIPALS algorithms 1 and 2 compute CPCA and HPCA, respectively. The main difference between these two approaches is in normalization step. While CPCA normalizes the loadings, HPCA applies a normalization in the scores. Therefore, loadings (CPCA) and scores (HPCA) from the same block are orthonormal.

Let $\mathbf{W}_a \in \mathbb{R}^{m \times f}$ be a projection matrix obtained by storing the loading vector $\mathbf{w}_a$ computed at each iteration of Algorithm 1. Similarly, $\mathbf{T} \in \mathbb{R}^{n \times f}$ is constructed storing the super scores $\mathbf{t}$. Due to the normalization and deflation steps of CPCA, $\mathbf{W}_a$ matrix is orthonormal ($\mathbf{W}_a^\top \mathbf{W}_a = \mathbf{I}$). Therefore, starting from the regression of super score on data block $\mathbf{X}_a$, we can obtain a closed-form for $\mathbf{T}$ as

$$\begin{aligned} \mathbf{X}_a &= \mathbf{T}\mathbf{W}_a^\top, \\ \mathbf{X}_a\mathbf{W}_a &= \mathbf{T}\mathbf{W}_a^\top\mathbf{W}_a, \\ \mathbf{T} &= \mathbf{X}_a\mathbf{W}_a \end{aligned} \tag{1}$$

and compute the super score $\hat{\mathbf{t}}_a$ for the testing sample $\hat{\mathbf{x}}_a$ as

$$\hat{\mathbf{t}}_a = \hat{\mathbf{x}}_a \mathbf{W}_a. \tag{2}$$

This vector corresponds to projections of $\hat{\mathbf{x}}_a$ to a common latent subspace in the CPCA method. Similarly, for a testing sample $\hat{\mathbf{x}}_b$, we can compute the super score $\hat{\mathbf{t}}_b$ using a projection matrix $\mathbf{W}_b$ as

$$\hat{\mathbf{t}}_b = \hat{\mathbf{x}}_b \mathbf{W}_b. \tag{3}$$

In HPCA method, $\mathbf{T}$ is a orthonormal matrix ($\mathbf{T}^\top\mathbf{T} = \mathbf{I}$). Therefore, the closed-form solution of $\mathbf{T}$ is obtained as

$$\begin{aligned} \mathbf{X}_a &= \mathbf{T}\mathbf{W}_a^\top, \\ \mathbf{X}_a\mathbf{W}_a &= \mathbf{T}\mathbf{W}_a^\top\mathbf{W}_a, \\ \mathbf{T} &= \mathbf{X}_a\mathbf{W}_a(\mathbf{W}_a^\top\mathbf{W}_a)^{-1}, \end{aligned} \tag{4}$$

where $\mathbf{W}_a$ can be computed as

$$\begin{aligned} \mathbf{X}_a &= \mathbf{T}\mathbf{W}_a^\top, \\ \mathbf{T}^\top\mathbf{X}_a &= \mathbf{T}^\top\mathbf{T}\mathbf{W}_a^\top, \\ \mathbf{W}_a &= \mathbf{X}_a^\top\mathbf{T}. \end{aligned} \tag{5}$$

Thus, combining Equations 4 and 5, we obtain the super scores $\hat{\mathbf{t}}_a$ for $\hat{\mathbf{x}}_a$ as

$$\hat{\mathbf{t}}_a = \hat{\mathbf{x}}_a\mathbf{X}_a^\top\mathbf{T}(\mathbf{T}^\top\mathbf{X}_a\mathbf{X}_a^\top\mathbf{T})^{-1}. \tag{6}$$

Similarly, we can compute the super scores $\hat{\mathbf{t}}_b$ for a testing sample $\hat{\mathbf{x}}_b$ using $\mathbf{X}_b$ in Equation 6.

In this work, we assume that the learned common subspace deals with appearance changes caused by the data capture executed by different cameras. Therefore, it is possible to compute the similarity between individuals $\hat{\mathbf{x}}_a$ and $\hat{\mathbf{x}}_b$ using the super scores $\hat{\mathbf{t}}_a$ and $\hat{\mathbf{t}}_b$ in a simple nearest neighbor method. Experimental results reinforce our hypothesis.

The complex appearance changes in images of the same individual captured by distinct cameras and the high matching performance of recent nonlinear person Re-ID models [6], [19] suggest that we can reach improved results using a nonlinear extension of multiblock PCA methods. Therefore, in Section III, we present a novel kernel extension to HPCA model, which we call *Kernel HPCA*. To the best of our knowledge, this is the first work to propose a kernel extension to a multiblock PCA method.

III. Proposed Method

In this section, we describe the proposed Kernel Hierarchical PCA (Kernel HPCA) method. Kernel HPCA relates data blocks nonlinearly with principal components to obtain block scores ($\mathbf{s}_a$ and $\mathbf{s}_b$). Therefore, our proposed method captures high-order correlation between input variables to learn the common latent subspace. Furthermore, we present an efficient derivation of NIPALS algorithm to iteratively compute the principal components (or factors) of Kernel HPCA.

Assuming a nonlinear transformation of the input variables to a feature space $\mathcal{F}$, i.e, $\mathbf{\Phi} : x_i \in \mathbb{R}^m \rightarrow \mathbf{\Phi}(x_i) \in \mathcal{F}$ and representing all mapped vectors by $\mathbf{\Phi}$, we can use the "kernel trick" to avoid explicitly mapping of data to a high-dimensional space substituting this cross-product by $K = \mathbf{\Phi}\mathbf{\Phi}^\top$, where $K \in \mathbb{R}^{n \times n}$ is the *kernel Gram matrix*. Particularly, we define the *kernel Gram matrices* $\mathbf{K}_a$ and $\mathbf{K}_b$ to represent the application of a kernel function using samples from blocks A and B, respectively. Thus, combining rows 4 and 5 from Algorithm 2 and applying the nonlinear transformation, we obtain that

$$\begin{aligned} \mathbf{s}_a &= \mathbf{X}_a\mathbf{X}_a^\top\mathbf{t}, \\ \mathbf{s}_a &= \mathbf{\Phi}_a\mathbf{\Phi}_a^\top\mathbf{t}, \\ \mathbf{s}_a &= \mathbf{K}_a\mathbf{t}. \end{aligned} \tag{7}$$

Likewise, we rewrite rows 6 and 7 from Algorithm 2 as

$$\mathbf{s}_b = \mathbf{K}_b\mathbf{t}. \tag{8}$$

Then, manipulating equations in 5, we derive a rank-one approximation of *kernel Gram matrix* $\mathbf{K}_a$ as

$$\begin{aligned} \mathbf{X}_a\mathbf{X}_a^\top &\approx \mathbf{t}\mathbf{t}^\top\mathbf{X}_a\mathbf{X}_a^\top, \\ \mathbf{\Phi}_a\mathbf{\Phi}_a^\top &\approx \mathbf{t}\mathbf{t}^\top\mathbf{\Phi}_a\mathbf{\Phi}_a^\top, \\ \mathbf{K}_a &\approx \mathbf{t}\mathbf{t}^\top\mathbf{K}_a. \end{aligned} \tag{9}$$

An analogous rank-one approximation to *kernel Gram matrix* $\mathbf{K}_b$ results in

$$\mathbf{K}_b \approx \mathbf{t}\mathbf{t}^\top\mathbf{K}_b. \tag{10}$$

Therefore, the rank-one deflation of kernel matrices $\mathbf{K}_a$ and $\mathbf{K}_b$ are, respectively,

$$\begin{aligned} \mathbf{K}_a &= \mathbf{K}_a - \mathbf{t}\mathbf{t}^\top\mathbf{K}_a \quad \text{and} \\ \mathbf{K}_b &= \mathbf{K}_b - \mathbf{t}\mathbf{t}^\top\mathbf{K}_b. \end{aligned} \tag{11}$$

Algorithm 3 presents our Kernel HPCA method. Kernel HPCA is based on the NIPALS algorithm for computation of Hierarchical PCA (see Algorithm 2) with the required modifications to efficiently handle with the nonlinear transformation of input variables. Similarly to HPCA, Kernel HPCA computes an orthonormal matrix $\mathbf{T} \in \mathbb{R}^{m \times f}$ whose columns store the super scores $\mathbf{t}$ obtained at each iteration. Thus, using the nonlinear mapping, we compute the super score $\hat{\mathbf{t}}_a$ for a new vector $\hat{\mathbf{x}}_a$ from block A as

$$\hat{\mathbf{t}}_a = \mathbf{\Phi}(\hat{\mathbf{x}}_a)\mathbf{\Phi}_a^\top\mathbf{T}(\mathbf{T}^\top\mathbf{\Phi}_a\mathbf{\Phi}_a^\top\mathbf{T})^{-1}. \tag{12}$$

Algorithm 3: Kernel Hierarchical PCA (Kernel HPCA).

input : $\mathbf{K}_a$, $\mathbf{K}_b$ and the number of factors (f)

1 randomly initialize $\mathbf{t}$
2 **for** *i=1* **to** f **do**
3 **while** *t do not converge* **do**
4 $\mathbf{s}_a = \mathbf{K}_a\mathbf{t}, \quad \mathbf{s}_a \leftarrow \frac{\mathbf{s}_a}{\|\mathbf{s}_a\|}$
5 $\mathbf{s}_b = \mathbf{K}_b\mathbf{t}, \quad \mathbf{s}_b \leftarrow \frac{\mathbf{s}_b}{\|\mathbf{s}_b\|}$
6 $\mathbf{S} = [s_a, s_b]$
7 $\mathbf{w} = \mathbf{S}^\top\mathbf{t}$
8 $\mathbf{t} = \mathbf{S}\mathbf{w} \quad \mathbf{t} \leftarrow \frac{\mathbf{t}}{\|\mathbf{t}\|}$
9 **end**
10 $\mathbf{K}_a \leftarrow \mathbf{K}_a - \mathbf{t}\mathbf{t}^\top\mathbf{K}_a$
11 $\mathbf{K}_b \leftarrow \mathbf{K}_b - \mathbf{t}\mathbf{t}^\top\mathbf{K}_b$
12 **end**

Furthermore, we define $\mathbf{K}_a(\hat{\mathbf{x}}_a, \mathbf{X}_a)$ as the computation of kernel function between $\hat{\mathbf{x}}_a$ and all vectors $\mathbf{x}_a \in \mathbf{X}_a$. Identically, $\mathbf{K}_b(\hat{\mathbf{x}}_b, \mathbf{X}_b)$ denotes the kernel function applied in $\hat{\mathbf{x}}_b$ and all vectors $\mathbf{x}_b \in \mathbf{X}_b$. Then, using the "kernel trick", we rewrite Equation 12 as

$$\hat{\mathbf{t}}_a = \mathbf{K}_a(\hat{\mathbf{x}}_a, \mathbf{X}_a)\mathbf{T}(\mathbf{T}^\top\mathbf{K}_a\mathbf{T})^{-1}. \tag{13}$$

Likewise, we compute $\hat{\mathbf{t}}_b$ for a new vector $\hat{\mathbf{x}}_b$ from block B as

$$\hat{\mathbf{t}}_b = \mathbf{K}_b(\hat{\mathbf{x}}_b, \mathbf{X}_b)\mathbf{T}(\mathbf{T}^\top\mathbf{K}_b\mathbf{T})^{-1}. \tag{14}$$

Similarly to HPCA and CPCA, we assume that the learned subspace handles the camera transition problem and the direct comparison between samples from different blocks results in a high performance when using their super scores. However, due the nonlinear transformation in the input variables, we believe that the common subspace learned with Kernel HPCA method is able to handle more complex feature transitions. Experimental results corroborate our hypothesis with great improvement when compared with the linear models HPCA and CPCA.

IV. Experimental Results

In this Section, we assess the matching performance of the proposed Kernel HPCA method on two widely used datasets for person Re-ID problem (VIPeR and PRID450S datasets). First, we describe these datasets and our evaluation protocol (Section IV-A). Then, we compare different variations of PCA-based methods for the specific task of subspace learning (Section IV-B). We end this section with a comparison between the proposed Kernel HPCA and state-of-the-art subspace learning and metric learning methods (Section IV-C). In the following tables, we use the symbol "*" to indicate a method learned without supervision (without labeled image pairs).

A. Datasets and Evaluation Protocol

VIPeR Dataset[1] **[30].** The Viewpoint Invariant Pedestrian Recognition (VIPeR) dataset contains 632 labeled image pairs

[1]Available at: https://vision.soe.ucsc.edu/projects

captured by two outdoor cameras. Each camera captures a single image of each subject (single-shot) and the images are normalized to 128×48 pixels. VIPeR is the most used dataset for supervised and unsupervised Re-ID. The main challenges are related to viewpoint changes, illumination and low-resolution images.

PRID 450S Dataset[2] **[31].** PRID 450S is a recently released dataset with more realistic images than VIPeR. It consists in 450 images pairs of pedestrians captured by two non-overlapping cameras. Each subject appears in single image at each camera (single-shot). The main challenges are related to changes in viewpoint, pose, camera characteristic as well as significant differences in background and illumination.

Experimental Setup. Similarly to other works in the literature, we achieve more stable results using ten random partitions of the dataset in training and testing subsets of equal sizes and reporting the mean values. In the testing subset, images from one camera are considered as gallery and images from the other camera are considered as probe. The results are reported using the matching performance at the *top-r* positions, which corresponds to the number of probe individuals correctly identified at the first r returned gallery individuals.

Feature Extraction. To represent images on VIPeR and PRID450S datasets, we use the combination of hand-crafted and Convolutional Neural Network (CNN) features in a 35024-dimensional feature descriptor proposed in [18].

Parameter Setting. To obtain a fair comparison, we use the same number of principal components (f) for all multiblock PCA-based methods. We set f to 100 for linear (CPCA, HPCA and PCA) and nonlinear models (Kernel HPCA and KPCA) due the best results obtained in a validation set (random splits between training and testing sets). Furthermore, we compute the *kernel Gram matrices* using exponential χ^2 kernel function, as described in [6], and we match individuals at the latent subspace using the cosine similarity.

TABLE I
VIPER: TOP RANKED CMC RESULTS.

Method	Viper (p=316)				
	r = 1	r = 5	r = 10	r = 20	r = 30
PCA*	13.7	26.4	36.5	47.7	54.2
PRDC [7]	15.7	38.4	53.9	70.0	-
EIML [4]	22.0	47.5	63.0	78.0	87.0
KISSME [3]	27.0	55.0	70.0	83.0	89.5
CPCA	28.9	63.5	78.9	**91.3**	**95.3**
RCCA+RD [8]	30.2	60.0	74.7	86.8	-
ROCCA [9]	30.4	-	75.6	86.6	-
HPCA	31.0	**65.2**	**79.6**	90.8	93.9
Prates and Schwartz [5]	**32.9**	62.3	78.7	87.8	91.6
KPCA*	14.6	29.8	39.7	51.5	58.4
KCCA [6]	37.0	-	85.0	93.0	-
KCCA + CNN Features	**39.6**	**73.8**	**85.5**	**93.7**	**96.6**
Kernel HPCA	39.4	73.0	85.1	93.5	96.1

[2]Available at: https://lrs.icg.tugraz.at/download.php

B. PCA Methods

Compared Methods. We perform a comparison between linear PCA-based methods: Hierarchical PCA (HPCA), Consensus PCA (CPCA) and the classical PCA. We focus on the specific task of common subspace learning. In addition, we compare our Kernel HPCA method with a baseline Kernel PCA that also learns a common subspace using a nonlinear transformation of input descriptors. Notice that we do not compare our method against the Kernel PCA method proposed in [24] since that method learns a specific model for each gallery image, while we are interested in learning a common subspace to the entire dataset.

Discussion. Tables I and II present the experimental results for VIPeR and PRID450S datasets, respectively. HPCA and CPCA capture cross-view discriminative information from labeled image pairs when learning the common subspace, which justifies the great improvement when compared with the unsupervised PCA. For instance, HPCA represents a gain at top-1 rank of 17.3 and 21.9 percentage points when compared to PCA. Therefore, experimental results corroborate our supposition that multiblock PCA methods are more indicated to person Re-ID problem than the classical PCA method. In addition, HPCA performs slightly superior than CPCA, mainly for the first ranking positions, which favor the choice of HPCA to propose our kernel extension.

From Tables I and II, we can conclude that KPCA and the proposed Kernel HPCA obtain higher matching performance than their linear counterpart PCA and HPCA, respectively. We attribute these results to the nonlinear transformations in images captured by two non-overlapping cameras. Furthermore, the Kernel HPCA presents a significant improvement when compared to the unsupervised KPCA, which is a consequence of using labeled image pairs from both cameras to learn the common subspace. Interestingly, the gain at top-1 rank of Kernel HPCA when compared to KPCA is 24.8 percentage points in VIPeR, while it reduces to 9.9 percentage points in PRID450S. These results are consistent with the difficulty of these datasets.

C. State-of-the-art Methods

Compared Methods. We compare the proposed Kernel HPCA method with a supervised PLS approach proposed by Prates and Schwartz [5], the well-known distance metric learning methods PRDC [7], KISSME [3] and EIML [4], and subspace learning approaches RCCA [8], ROCCA [9] and KCCA [6]. For KCCA, we provide the results reported by the authors (VIPeR) and the results obtained using the code provided by the authors and the features described in Section IV-A, which we denote as KCCA + CNN features in the tables. Notice that some approaches are missing in Table II because they neither provide their code nor results for the dataset PRID450S.

Discussion. Considering VIPeR (Table I) and PRID450S (Table II) datasets, Kernel HPCA reaches comparable results when compared to KCCA, which is the state-of-the-art on nonlinear subspace learning method in person Re-ID literature.

However, while KCCA requires more than 300 factors, Kernel HPCA requires only 100, which demonstrates its efficient combination of dimensionality reduction and subspace learning. Furthermore, when compared to other subspace learning and metric learning methods, we obtain higher matching performance. We attribute this improvement to the nonlinear mapping and the multiblock analysis that allow our model to learn more complex cross-view discriminant information.

TABLE II
PRID450S: TOP RANKED CMC RESULTS.

Method	PRID450S (p=225)				
	r = 1	r = 5	r = 10	r = 20	r = 30
PCA*	23.0	45.5	59.7	71.6	79.5
Prates and Schwartz [5]	29.3	52.5	63.1	75.0	82.1
KISSME [3]	33.0	59.8	71.0	79.0	84.5
EIML [4]	35.0	58.5	68.0	77.0	83.0
CPCA	39.6	69.3	82.2	90.2	93.8
HPCA	**44.9**	**73.8**	**83.3**	**91.5**	**95.0**
KPCA*	42.9	67.7	76.6	84.7	89.5
Kernel HPCA	52.2	**80.9**	**92.8**	94.4	96.8
KCCA + CNN Features	**52.8**	**80.9**	89.0	**95.1**	**97.2**

V. CONCLUSIONS AND FUTURE WORKS

This paper addressed person Re-ID problem using multiblock PCA methods. Experimental results showed that Hierarchical and Consensus PCA methods are sucessful for person Re-ID problem. For instance, they reach comparable results when compared to linear subspace learning methods. However, they cannot deal with the nonlinear feature transformations between distinct camera-views. Therefore, we proposed a kernel extension of HPCA method, called Kernel HPCA. According to experimental results, the Kernel HPCA is comparable with KCCA, which is the state-of-the-art nonlinear subspace learning method. However, Kernel HPCA requires significantly fewer factors (dimensions) than KCCA method.

As future works, we intend to evaluate the performance of Kernel HPCA method at other multimodal problems, such as sketches and mugshot photos matching [32].

ACKNOWLEDGMENTS

The authors would like to thank the Brazilian National Research Council – CNPq (Grant #477457/2013-4), the Minas Gerais Research Foundation – FAPEMIG (Grants APQ-00567-14 and PPM-00025-15) and the Coordination for the Improvement of Higher Education Personnel – CAPES (DeepEyes Project).

REFERENCES

[1] Bedagkar-Gala Apurva and Shishir K. Shah., "A survey of approaches and trends in person re-identification," *Image and Vision Computing*, vol. 32, no. 4, pp. 270 – 286, 2014.

[2] Riccardo Satta, "Appearance descriptors for person re-identification: a comprehensive review," *CoRR*, vol. abs/1307.5748, 2013.

[3] M. Koestinger, M. Hirzer, P. Wohlhart, P. Roth, and H. Bischof, "Large scale metric learning from equivalence constraints," in *IEEE CVPR*, 2012.

[4] H. Martin, R. Peter, and B. Horst, "Person re-identification by efficient impostor-based metric learning," in *IEEE AVSS*, 2012, pp. 203–208.

[5] R. Prates and W. Schwartz, "Appearance-based person re-identification by intra-camera discriminative models and rank aggregation," in *IAPR ICB*, 2015, pp. 65–72.

[6] G. Lisanti, Iacopo M., and A. Del Bimbo, "Matching people across camera views using kernel canonical correlation analysis," 2014, pp. 10:1–10:6, ACM.

[7] Z. Wei-Shi, G. Shaogang, and X. Tao, "Person re-identification by probabilistic relative distance comparison," in *IEEE CVPR*, 2011.

[8] L. An, M. Kafai, S. Yang, and B. Bhanu, "Reference-based person re-identification," in *IEEE AVSS*, 2013.

[9] L. An, S. Yang, and B. Bhanu, "Person re-identification by robust canonical correlation analysis," *Signal Processing Letters, IEEE*, vol. 22, no. 8, pp. 1103–1107, Aug 2015.

[10] W.R. Schwartz and L.S. Davis, "Learning discriminative appearance-based models using partial least squares," in *SIBGRAPI*, Oct 2009, pp. 322–329.

[11] Shengcai Liao, Yang Hu, Xiangyu Zhu, and Stan Z Li, "Person re-identification by local maximal occurrence representation and metric learning," in *IEEE CVPR*, 2015, pp. 2197–2206.

[12] Chunxiao Liu, Shaogang Gong, and Chen Change Loy, "On-the-fly feature importance mining for person re-identification," *Pattern Recognition*, vol. 47, no. 4, pp. 1602 – 1615, 2014.

[13] Bingpeng Ma, Yu Su, and Frdric Jurie, "Local descriptors encoded by fisher vectors for person re-identification," 2012, vol. 7583 of *Lecture Notes in Computer Science*, pp. 413–422, Springer Berlin Heidelberg.

[14] Bingpeng Ma, Yu Su, and Frederic Jurie, "Bicov: a novel image representation for person re-identification and face verification," in *BMVC*, 2012.

[15] D. S. Cheng, M. Cristani, M. Stoppa, L. Bazzani, and V. Murino, "Custom pictorial structures for re-identification," in *BMVC*, 2011.

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

[17] M. Farenzena, L. Bazzani, A. Perina, V. Murino, and M. Cristani, "Person re-identification by symmetry-driven accumulation of local features," in *IEEE CVPR*, 2010.

[18] X. Li J. You W. Zheng S. Wu, Y. Chen, "An enhanced deep feature representation for person re-identification," in *IEEE WACV*, March 2016.

[19] Fei Xiong, Mengran Gou, Octavia Camps, and Mario Sznaier, "Person re-identification using kernel-based metric learning methods," in *ECCV*, pp. 1–16. Springer, 2014.

[20] Bernhard Schölkopf, Alexander Smola, and Klaus-Robert Müller, "Nonlinear component analysis as a kernel eigenvalue problem," *Neural computation*, vol. 10, no. 5, pp. 1299–1319, 1998.

[21] Matthew A Turk and Alex P Pentland, "Face recognition using eigenfaces," in *IEEE CVPR*, 1991, pp. 586–591.

[22] Chiraz BenAbdelkader, Ross Cutler, Harsh Nanda, and Larry Davis, "Eigengait: Motion-based recognition of people using image self-similarity," in *Audio-and Video-Based Biometric Person Authentication*. Springer, 2001, pp. 284–294.

[23] Michael J Black and Allan D Jepson, "Eigentracking: Robust matching and tracking of articulated objects using a view-based representation," *IJCV*, vol. 26, no. 1, pp. 63–84, 1998.

[24] Jun Yang, Zhongke Shi, and P.A. Vela, "Person reidentification by kernel pca based appearance learning," in *CRV*, 2011.

[25] Niki Martinel and C Micheloni, "Person re-identification by modelling principal component analysis coefficients of image dissimilarities," *Electronics Letters*, vol. 50, no. 14, pp. 1000–1001, 2014.

[26] Johan A Westerhuis, Theodora Kourti, and John F MacGregor, "Analysis of multiblock and hierarchical pca and pls models," *Journal of chemometrics*, vol. 12, no. 5, pp. 301–321, 1998.

[27] Svante Wold, M. Hellberg, H. Sjostrom, and H. Wold, "Pls modeling with latent variables in two or more dimensions," *PLS Model Building: Theory and applications.*, 1987.

[28] Svante Wold, Nouna Kettaneh, and Kjell Tjessem, "Hierarchical multiblock pls and pc models for easier model interpretation and as an alternative to variable selection," *Journal of chemometrics*, vol. 10, no. 5-6, pp. 463–482, 1996.

[29] B Qiu, Véronique Prinet, Edith Perrier, and Olivier Monga, "Multi-block pca method for image change detection," in *ICIAP*, 2003.

[30] D. Gray, S. Brennan, and H. Tao, "Evaluating appearance models for recognition, reacquisition, and tracking," in *IEEE PETS*, 2007.

[31] Peter R., M. Hirzer, M. Kstinger, Csaba B., and Horst B., "Mahalanobis distance learning for person re-identification," in *Person Re-Identification*, Springer ACVPR, pp. 247–267. 2014.

[32] Brendan F Klare, Zhifeng Li, and Anil K Jain, "Matching forensic sketches to mug shot photos," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 33, no. 3, pp. 639–646, 2011.