

Person Re-Identification by Robust Canonical Correlation Analysis

Le An, Songfan Yang, *Member, IEEE*, and Bir Bhanu, *Fellow, IEEE*

Abstract—Person re-identification is the task to match people in surveillance cameras at different time and location. Due to significant view and pose change across non-overlapping cameras, directly matching data from different views is a challenging issue to solve. In this letter, we propose a robust canonical correlation analysis (ROCCA) to match people from different views in a coherent subspace. Given a small training set as in most re-identification problems, direct application of canonical correlation analysis (CCA) may lead to poor performance due to the inaccuracy in estimating the data covariance matrices. The proposed ROCCA with shrinkage estimation and smoothing technique is simple to implement and can robustly estimate the data covariance matrices with limited training samples. Experimental results on two publicly available datasets show that the proposed ROCCA outperforms regularized CCA (RCCA), and achieves state-of-the-art matching results for person re-identification as compared to the most recent methods.

Index Terms—Canonical correlation analysis (CCA), covariance estimation, person re-identification, subspace, surveillance.

I. INTRODUCTION

PERSON re-identification aims at matching people across non-overlapping cameras. It has a great potential in surveillance applications such as crowd monitoring [1] and people tracking [2]. Recently there has been a significant research interest in the design of re-identification algorithms [3], [4], [5], [6], [7], [8], [9], [10]. To associate individuals at different time and locations is a challenging task, since the captured images of the same subject may vary significantly due to pose variation, illumination change, occlusion, background clutter, and low image resolution.

In general, there are two classes of methods to improve the performance of person re-identification. Methods of the first

class extract robust appearance features such that discrimination is ensured for different subjects while invariance is maintained for the same subject in different camera views. Kviatkovsky *et al.* [11] discovered an intra-distribution color structure and applied it in conjunction with covariance descriptors to re-identify people. Kuo *et al.* [12] applied semantic color names to describe a person by the probabilities of the presence of predefined colors. This resulted in a better performance than what was achieved using color histograms only. Yang *et al.* [13] proposed a color descriptor based on salient color names which can guarantee that a higher probability will be assigned to the color name which is closest to the intrinsic color. Farenzena *et al.* [4] combined features including the overall chromatic content, the spatial arrangement and the presence of recurrent local motifs to handle individual matching with appearance variation. Liu *et al.* [14] studied the importance of different features and proposed a method for on-the-fly feature mining. Zhao *et al.* [15] learned discriminative mid-level filters from automatically discovered patch clusters to identify specific visual patterns.

Methods of the second class learn feature transformations or distance metrics such that the intra-class distance between the same person from different cameras is reduced and the inter-class distance between different persons is increased. Following this spirit, Zheng *et al.* [8] formulated re-identification as a relative distance comparison problem, and the likelihood is maximized such that the distance between the same-person image pair is smaller than the distance between a different-person image pair. Köstinger *et al.* [16] proposed a simple metric called “KISS” from a statistical inference perspective such that a decision of dissimilarity can be obtained by a likelihood ratio test. Tao *et al.* [9], [10] improved the “KISS” metric and achieved a better performance. Hirzer *et al.* [7] developed a relaxed pairwise learned metric based on Mahalanobis distance. Li *et al.* [6] partitioned the image spaces of two camera views into different configurations based on the similarity of cross-view transforms and image pairs with similar transforms were projected into a common feature space for matching. An *et al.* [17] proposed the use of reference descriptors instead of using image features directly. Loy *et al.* [18] proposed a cross canonical correlation analysis (xCCA) to model temporal and causal relationships between regional activities within and across camera views. Li *et al.* [19] proposed a filter pairing neural network in which misalignment, pose difference, occlusions and background clutter were jointly handled with the help of abundant data. Xiong *et al.* [20] applied multiple kernel-based metrics in conjunction with histogram-based features and showed improvement over state-of-the-art on several datasets.

Manuscript received October 08, 2014; revised December 09, 2014; accepted December 28, 2014. Date of publication January 12, 2015; date of current version January 14, 2015. This work supported in part by NSF Grant I330110 and by ONR Grant N00014-12-1-1026. L. An and S. Yang contributed equally to this work and are both considered first authors. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Mathew Magimai Doss.

L. An is with BRIC, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599 USA (e-mail: lan004@unc.edu).

S. Yang is with College of Electronics and Information Engineering, Sichuan University, Chengdu 610064, China (e-mail: syang@scu.edu.cn).

B. Bhanu is with Center for Research in Intelligent Systems, University of California-Riverside, Riverside, CA, 92521 USA (e-mail: bhanu@cris.ucr.edu). (Corresponding author: S. Yang.)

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/LSP.2015.2390222

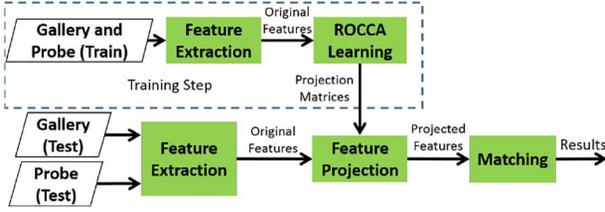


Fig. 1. The proposed robust canonical correlation analysis (ROCCA) method for person re-identification.

In this letter, we propose a robust canonical correlation analysis (ROCCA) based method to project the data from different views into a coherent subspace where matching can be reliably performed. Fig. 1 provides an overview of ROCCA based re-identification process. As the basis of the proposed ROCCA, canonical correlation analysis (CCA) is a multi-view data analysis technique and has been applied to various tasks such as face recognition [21] and image super-resolution [22]. However, for person re-identification, normally the number of training samples is significantly smaller than the feature dimension. As a result, direct application of CCA is prone to poor performance due to error in estimation of the data covariance matrices. The improvement of CCA has been studied in previous work using projection pursuit and alternating regressions [23]. An *et al.* [17] used regularized CCA (RCCA) in conjunction with a reference set for person re-identification. In this letter we propose an alternative method to apply CCA to the problem of person re-identification. Unlike RCCA used in [17], in this letter the proposed ROCCA is based on a shrinkage estimation and smoothing technique. The proposed method is less complicated and consistently outperforms the existing alternative (i.e., RCCA). As compared to the recent methods, state-of-the-art results are achieved on two public datasets. The implementation of ROCCA will be made available upon request.

II. CANONICAL CORRELATION ANALYSIS (CCA)

First introduced in [24], CCA is a multivariate statistical analysis tool. CCA aims at projecting two sets of multivariate data into a subspace such that the correlation between the projected data is maximized.

In the problem of person re-identification, given N image pairs from two cameras, appearance features with dimension p are first extracted from the images. These feature vectors are organized into two data matrices $X = \{X_i \in \mathbb{R}^p, i = 1, 2, \dots, N\}$ and $Y = \{Y_i \in \mathbb{R}^p, i = 1, 2, \dots, N\}$. The goal of CCA is to find a pair of projection vectors $w_X \in \mathbb{R}^p$ and $w_Y \in \mathbb{R}^p$ such that the correlation coefficient r of $w_X^T X$ and $w_Y^T Y$ is maximized. Specifically, the objective function to be maximized is

$$\begin{aligned} r &= \frac{\text{cov}(w_X^T X, w_Y^T Y)}{\sqrt{\text{var}(w_X^T X) \text{var}(w_Y^T Y)}} \\ &= \frac{w_X^T C_{XY} w_Y}{\sqrt{w_X^T C_{XX} w_X w_Y^T C_{YY} w_Y}}, \end{aligned} \quad (1)$$

where $\text{cov}(\cdot)$ is a covariance operator and $\text{var}(\cdot)$ computes the data variance. The covariance matrices are computed by

$C_{XX} = E[XX^T]$, $C_{YY} = E[YY^T]$, $C_{XY} = E[XY^T]$, and $C_{YX} = E[YX^T]$, where $E[\cdot]$ is the expectation operator.

Eqn. (1) can be reformulated as a constrained optimization problem as follows

$$\begin{aligned} &\text{maximize} && w_X^T C_{XY} w_Y \\ &\text{subject to} && w_X^T C_{XX} w_X = 1 \\ &&& w_Y^T C_{YY} w_Y = 1. \end{aligned} \quad (2)$$

The solution of w_X and w_Y in Eqn. (2) can be obtained through the following generalized eigenvalue problem

$$\begin{bmatrix} 0 & C_{XY} \\ C_{YX} & 0 \end{bmatrix} \begin{bmatrix} w_X \\ w_Y \end{bmatrix} = \lambda \begin{bmatrix} C_{XX} & 0 \\ 0 & C_{YY} \end{bmatrix} \begin{bmatrix} w_X \\ w_Y \end{bmatrix}, \quad (3)$$

and w_X is an eigenvector of $C_{XX}^{-1} C_{XY} C_{YY}^{-1} C_{YX}$, w_Y is an eigenvector of $C_{YY}^{-1} C_{YX} C_{XX}^{-1} C_{XY}$.

III. ROBUST CCA (ROCCA)

To obtain the solution of CCA, the inverse of the covariance matrices C_{XX} and C_{YY} needs to be computed. In practice, for problems like person re-identification, the number of available samples N is usually smaller than the feature dimension p per sample. In this case the covariance matrices may be singular and their inverse would be ill-conditioned. To tackle this problem, the goal is to find an estimated covariance matrix $\hat{\Sigma}$ which best approximates the underlying covariance matrix Σ in terms of the mean square error given by

$$E \left[\left\| \hat{\Sigma} - \Sigma \right\|_F^2 \right], \quad (4)$$

where $\|\cdot\|_F$ denotes matrix Frobenius norm. Directly solving Eqn. (4) is infeasible due to the unknown data covariance Σ . To obtain a feasible estimation, the shrinkage algorithm was proposed in [25]. Specifically, given sample covariance C , the shrinkage method generates the estimated covariance matrix $\hat{\Sigma}$ in the following form

$$\hat{\Sigma} = \rho \alpha I + (1 - \rho)C, \quad (5)$$

where ρ is the shrinkage coefficient, $\alpha = \text{tr}(C)/p$, where $\text{tr}(\cdot)$ computes the matrix trace, and I is an identity matrix.

To obtain the shrinkage coefficient ρ , the optimization, referred to as oracle estimator, is defined as

$$\begin{aligned} &\text{minimize}_{\rho} && E \left[\left\| \hat{\Sigma} - \Sigma \right\|_F^2 \right] \\ &\text{subject to} && \hat{\Sigma} = \rho \alpha I + (1 - \rho)C. \end{aligned} \quad (6)$$

Eqn. (6) is not approachable due to the unknown Σ . To overcome this issue, we apply Oracle Approximating Shrinkage (OAS) method [26] to approximate the solution of Eqn. (6). OAS is designed as an iterative algorithm for the oracle estimator. In the first step Σ is replaced with sample covariance C as an initial guess and $\hat{\Sigma}_0$ is obtained. In the second step, $\hat{\Sigma}_0$ replaces C and $\hat{\Sigma}_1$ is acquired. Subsequent steps iteratively refine the covariance estimation $\hat{\Sigma}$.

As proved in [26], the OAS iterative process converges to a closed form solution of ρ given by

$$\rho = \min \left(\frac{(1 - \frac{2}{p}) \text{tr}(C^2) + \text{tr}^2(C)}{(N + 1 - \frac{2}{p}) \left[\text{tr}(C^2) - \frac{\text{tr}^2(C)}{p} \right]}, 1 \right), \quad (7)$$

where N is the number of samples. This expression allows us to obtain ρ without iteration, and the estimated covariance $\hat{\Sigma}$ can be readily computed by Eqn. (5).

For more robust estimation of the covariance matrix, we apply the smoothing technique [9] to stabilize the small eigenvalues of $\hat{\Sigma}$, which are prone to estimation errors. To proceed, we first compute eigen-decomposition of $\hat{\Sigma}$ as

$$\hat{\Sigma} = Q\Lambda Q^T, \quad (8)$$

where $\Lambda = \text{diag}[\lambda_1, \lambda_2, \dots, \lambda_{p-1}, \lambda_p]$. To perform smoothing, the last k smallest eigenvalues are replaced by a constant γ which is computed as the average of the last k eigenvalues

$$\gamma = \frac{1}{k} \sum_{i=p-k+1}^p \lambda_i. \quad (9)$$

The smoothed version of eigenvalue matrix Λ becomes $\tilde{\Lambda} = \text{diag}[\lambda_1, \lambda_2, \dots, \gamma, \gamma]$, and the final covariance estimation is obtained by

$$\tilde{\Sigma} = Q\tilde{\Lambda}Q^T. \quad (10)$$

In our problem, instead of using the sample covariance matrices C_{XX} and C_{YY} in Eqn. (3) to calculate CCA projection matrices, we use the robustly estimated covariance matrices $\tilde{\Sigma}_{XX}$ and $\tilde{\Sigma}_{YY}$ to solve the CCA problem in Eqn. (3), which is referred to as ROCCA. The computation of $\tilde{\Sigma}_{XX}$ and $\tilde{\Sigma}_{YY}$ using shrinkage and smoothing is summarized in Algorithm 1. ROCCA projection matrices are computed during training. In the testing phase of re-identification, given a probe, image features are first extracted and then transformed into the subspace learned by ROCCA. The matching is based on the similarity between the projected features of a probe and a gallery using the Nearest Neighbor (NN) classifier. We have empirically found out that cosine similarity performs better than L2 distance and it is computationally more efficient. Therefore, we have chosen to use cosine similarity.

Algorithm 1 PRobust Covariance Estimation for CCA.

Input:

Sample covariance matrices C_{XX} and C_{YY}

- 1: Compute ρ using OAS estimator in Eqn. (7).
- 2: Perform shrinkage estimation using Eqn. (5) to obtain $\hat{\Sigma}_{XX}$ and $\hat{\Sigma}_{YY}$.
- 3: Compute eigen-decomposition of $\hat{\Sigma}_{XX}$ and $\hat{\Sigma}_{YY}$ using Eqn. (8).
- 4: Perform smoothing using Eqns. (9), (10).

Output:

$\tilde{\Sigma}_{XX}$ and $\tilde{\Sigma}_{YY}$

IV. EXPERIMENTS

A. Datasets

To evaluate our method, the VIPeR dataset [3] and the CUHK Campus dataset [27] are used. The VIPeR dataset contains im-



Fig. 2. Samples from VIPeR dataset (left) and CUHK Campus dataset (right).

ages of 632 persons in two cameras. For each person, one image is available in each view. For most of the subjects the view change is more than 90 degrees. The CUHK Campus dataset includes images of 971 subjects from two non-overlapping cameras. One camera captures the frontal or rear view of a person and the other camera captures the profile view of a person. Each person has two images in each camera view. Fig. 2 shows some samples from the two datasets.

B. Experimental Setup

All of the images in the experiments are normalized to 128×48 . Each image is divided into blocks of size 8×16 . The blocks are overlapped by 50% in both horizontal and vertical directions. For each block, the extracted color features include the quantized mean values of different channels in the HSV and Lab color space, and the semantic color names [12]. The texture features are extracted using the 8-bit Local Binary Patterns (LBP) [28]. The final appearance features for one block is the concatenation of both color and texture features with dimension $3 + 3 + 12 + 256 = 274$.

To determine the parameter k for smoothing in Eqn. (9), we adopt the elbow method [29]. Elbow method has been widely used to determine the number of clusters for clustering algorithms such as K-means. In our case, we choose k such that at the “elbow” position, k most significant eigenvalues are kept. Specifically, k is set to 314 for the VIPeR dataset and 480 for the CUHK Campus dataset in the experiments.

To fairly compare with other methods, we follow the standard experimental protocols used in the previous work (e.g., [4], [7], [16], [30]) for both datasets. Each dataset is randomly split into two subsets of equal size. Half of the data are used for training and the other half are used for testing. Gallery consists of images from one camera and images from the other camera are used as probes. For each dataset, the experiments are conducted 10 times and the averaged matching results are reported. The results from the competing methods are directly obtained from the corresponding papers.

C. Results

Experimental results are reported in terms of matching rates at different ranks. To evaluate and compare the performance of the proposed ROCCA with RCCA [17], the CMC curves for the top 10 ranks on the VIPeR dataset and the CUHK Campus dataset are plotted in Fig. 3 and Fig. 4, respectively. Note that the same image features (see Section IV-B) are used for both methods. Fig. 3 and Fig. 4 provide a major justification of the proposed approach. For the VIPeR dataset, the rank-1 matching rate is 26.19% using RCCA. When ROCCA is used, the rank-1 matching rate rises to 30.44% due to the improved covariance estimation, yielding a relative increase of 16%. At higher ranks,

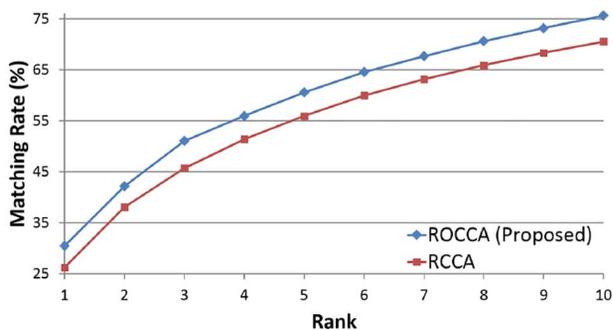


Fig. 3. CMC curves of RCCA and ROCCA on the VIPeR dataset.

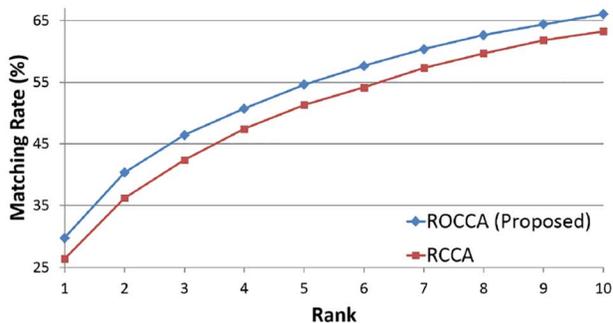


Fig. 4. CMC curves of RCCA and ROCCA on the CUHK Campus dataset.

the ROCCA based method consistently outperforms RCCA. For the CUHK Campus dataset, the rank-1 matching rate goes up from 26.36% to 29.77% when RCCA is replaced by ROCCA, which indicates a relative improvement of 13%.

To compare our method with the state-of-the-art, 16 recent methods are included in the comparison on the VIPeR dataset. The matching rates at various ranks for different methods are reported in Table I. Compared to the rest of the methods, the ROCCA based approach achieved the highest matching rate of 30.44% at rank-1. Note that in the method of [17] referred to as RCCA + RD, RCCA has been used to first project the image features into a coherent subspace to learn the reference descriptors. Using ROCCA in conjunction with semantic color names in the image features, very competitive results are achieved by the proposed method with a much simpler framework. Note that when the same image features are used, ROCCA outperformed RCCA, as illustrated in Fig. 3 and Fig. 4. Compared to the rest of the methods, ROCCA based method also has superior performance.

For the CUHK Campus dataset, we compare the proposed method with eight different approaches. Table II lists the matching rates at different ranks. The results of ROCCA and RCCA + RD[17] are very close at different ranks, while outperforming the rest of the methods. This is similar to the observation made in Table I. To evaluate if ROCCA can be incorporated with an existing method [17] for improved performance, we compare ROCCA in conjunction with RD (ROCCA + RD) to RCCA + RD[17] using the same image features and the same matching method (i.e., cosine similarity + NN) as in [17]. Table III shows that improved matching results (in terms of rank-1 matching rate) are achieved when ROCCA is used in an existing framework [17].

Regarding the computational cost, the one-time offline covariance estimation (see Algorithm 1) takes less than 10 seconds

TABLE I
MATCHING RATES (IN%) ON THE VIPeR DATASET

Rank→	$r = 1$	10	20	50	100
ROCCA (Proposed)	30.44	75.63	86.61	95.98	98.80
RCCA+RD [17]	30.25	74.68	86.82	95.70	99.24
SalMatch [30]	30.16	65.54	79.15	91.49	98.10
LAF [6]	29.60	69.30	84.50	96.80	99.00
RPLM [7]	27.34	69.02	82.69	94.56	98.54
RS-KISS [9]	24.50	66.60	81.70	93.50	98.00
CPS [31]	21.84	57.21	71.00	87.00	91.77
BiCov [32]	20.66	56.18	68.00	81.56	88.66
KISSME [16]	20.03	62.39	77.46	92.81	98.19
LMNN-R [33]	20.00	66.00	79.00	92.50	95.18
SDALF [4]	19.87	49.37	65.73	84.84	90.43
MRank [34]	19.34	55.51	70.44	87.69	96.90
PCCA [35]	19.27	64.91	80.28	95.00	97.01
DDC [36]	19.00	52.00	65.00	80.00	91.00
LMNN [37]	17.41	53.86	67.88	88.13	96.23
PRDC [8]	15.66	53.86	70.09	87.79	92.84
ITML [38]	15.54	53.13	69.05	88.54	96.93

TABLE II
MATCHING RATES (IN%) ON THE CUHK CAMPUS DATASET

Rank→	$r = 1$	10	20	50	100
ROCCA (Proposed)	29.77	66.02	76.76	88.47	94.82
RCCA+RD [17]	29.97	67.78	77.04	87.24	94.21
SalMatch [30]	28.45	55.68	67.95	83.53	92.10
LAF [6]	25.80	64.50	78.00	93.00	98.50
ITML [38]	15.98	45.60	59.81	76.61	88.32
LMNN [37]	13.45	42.25	54.11	73.29	86.65
SDALF [4]	9.90	30.33	41.03	55.99	67.39
L2 distance [30]	9.84	26.42	33.13	46.98	63.48
L1 distance [30]	10.33	26.34	33.52	45.62	61.95

TABLE III
RANK-1 MATCHING RATE (IN%) OF ROCCA + RD AND RCCA + RD[17]
USING THE SAME IMAGE FEATURES AS IN [17]

Dataset→	VIPeR	CUHK Campus
ROCCA+RD	32.59	32.71
RCCA+RD [17]	30.25	29.97

on both datasets using the unoptimized Matlab implementation on a laptop with Intel i7 2.4 GHz CPU and 8 GB RAM. This can be further improved to ensure efficiency. In addition, the proposed method can be generalized with different feature descriptors or classifiers.

V. CONCLUSIONS

In this letter, we have presented a robust canonical correlation analysis (ROCCA) based method for person re-identification. Given the nature of the re-identification problem in which a small sample size of data encounters a high feature dimension, the proposed ROCCA can robustly estimate the data covariance matrices in order to learn feature transformations. The learned transformation matrices project the original data from different cameras into a coherent subspace with maximized correlation between the same person. Despite the simplicity in its implementation and efficiency in computation, the proposed ROCCA consistently outperformed regularized CCA (RCCA), and achieved competitive or better results as compared to the most recent methods on two widely used person re-identification datasets. In addition, the proposed ROCCA can be incorporated into existing re-identification methods to obtain improved matching results.

REFERENCES

- [1] W. Li, V. Mahadevan, and N. Vasconcelos, "Anomaly detection and localization in crowded scenes," *IEEE Trans. Patt. Anal. Mach. Intell.*, vol. 36, no. 1, pp. 18–32, Jan. 2014.
- [2] W. Hu, M. Hu, X. Zhou, T. Tan, J. Lou, and S. Maybank, "Principal axis-based correspondence between multiple cameras for people tracking," *IEEE Trans. Patt. Anal. Mach. Intell.*, vol. 28, no. 4, pp. 663–671, Apr. 2006.
- [3] D. Gray, S. Brennan, and H. Tao, "Evaluating appearance models for recognition, reacquisition, and tracking," in *Proc. PETS*, Sep. 2007.
- [4] M. Farenzena, L. Bazzani, A. Perina, V. Murino, and M. Cristani, "Person re-identification by symmetry-driven accumulation of local features," in *Proc. CVPR Workshop*, Jun. 2010, pp. 2360–2367.
- [5] X. Wang, G. Doretto, T. Sebastian, J. Rittscher, and P. Tu, "Shape and appearance context modeling," in *Proc. ICCV*, Oct. 2007, pp. 1–8.
- [6] W. Li and X. Wang, "Locally aligned feature transforms across views," in *Proc. CVPR*, Jun. 2013, pp. 3594–3601.
- [7] M. Hirzer, P. M. Roth, M. Köstinger, and H. Bischof, "Relaxed pairwise learned metric for person re-identification," in *Proc. ECCV*, 2012, pp. 780–793.
- [8] W.-S. Zheng, S. Gong, and T. Xiang, "Reidentification by relative distance comparison," *IEEE Trans. Patt. Anal. Mach. Intell.*, vol. 35, no. 3, pp. 653–668, 2013.
- [9] D. Tao, L. Jin, Y. Wang, Y. Yuan, and X. Li, "Person re-identification by regularized smoothing kiss metric learning," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 23, no. 10, pp. 1675–1685, 2013.
- [10] D. Tao, L. Jin, Y. Wang, and X. Li, "Person reidentification by minimum classification error-based kiss metric learning," *IEEE Trans. Cybern.*, no. 99, pp. 1–1, 2014.
- [11] I. Kviatkovsky, A. Adam, and E. Rivlin, "Color invariants for person reidentification," *IEEE Trans. Patt. Anal. Mach. Intell.*, vol. 35, no. 7, pp. 1622–1634, Jul. 2013.
- [12] C.-H. Kuo, S. Khamis, and V. Shet, "Person re-identification using semantic color names and rankboost," in *Proc. WACV*, 2013, pp. 281–287.
- [13] Y. Yang, J. Yang, J. Yan, S. Liao, D. Yi, and S. Li, "Salient color names for person re-identification," in *Proc. ECCV*, 2014, pp. 536–551.
- [14] C. Liu, S. Gong, and C. C. Loy, "On-the-fly feature importance mining for person re-identification," *Patt. Recognition*, vol. 47, no. 4, pp. 1602–1615, 2014.
- [15] R. Zhao, W. Ouyang, and X. Wang, "Learning mid-level filters for person re-identification," in *Proc. CVPR*, Jun. 2014, pp. 144–151.
- [16] M. Köstinger, M. Hirzer, P. Wohlhart, P. Roth, and H. Bischof, "Large scale metric learning from equivalence constraints," in *Proc. CVPR*, Jun. 2012, pp. 2288–2295.
- [17] L. An, M. Kafai, S. Yang, and B. Bhanu, "Reference-based person re-identification," in *Proc. AVSS*, Aug. 2013, pp. 244–249.
- [18] C. C. Loy, T. Xiang, and S. Gong, "Multi-camera activity correlation analysis," in *Proc. CVPR*, Jun. 2009, pp. 1988–1995.
- [19] W. Li, R. Zhao, T. Xiao, and X. Wang, "DeepReID: Deep filter pairing neural network for person re-identification," in *Proc. CVPR*, Jun. 2014, pp. 152–159.
- [20] F. Xiong, M. Gou, O. Camps, and M. Szaiaier, "Person re-identification using kernel-based metric learning methods," in *Proc. ECCV*, 2014, pp. 1–16.
- [21] A. Sharma, M. A. Haj, J. Choi, L. S. Davis, and D. W. Jacobs, "Robust pose invariant face recognition using coupled latent space discriminant analysis," *Comput. Vis. Image Understand.*, vol. 116, no. 11, pp. 1095–1110, 2012.
- [22] H. Huang, H. He, X. Fan, and J. Zhang, "Super-resolution of human face image using canonical correlation analysis," *Patt. Recognit.*, vol. 43, no. 7, pp. 2532–2543, 2010.
- [23] J. Branco, C. Croux, P. Filzmoser, and M. Oliveira, "Robust canonical correlations: A comparative study," *Comput. Statist.*, vol. 20, no. 2, pp. 203–229, 2005.
- [24] H. Hotelling, "Relations between two sets of variates," *Biometrika*, vol. 28, no. 3/4, pp. 321–377, 1936.
- [25] O. Ledoit and M. Wolf, "Nonlinear shrinkage estimation of large-dimensional covariance matrices," *Ann. Statist.*, vol. 40, no. 2, pp. 1024–1060, 04, 2012.
- [26] Y. Chen, A. Wiesel, Y. Eldar, and A. Hero, "Shrinkage algorithms for MMSE covariance estimation," *IEEE Trans. Signal Process.*, vol. 58, no. 10, pp. 5016–5029, Oct. 2010.
- [27] W. Li, R. Zhao, and X. Wang, "Human reidentification with transferred metric learning," in *Proc. ACCV*, 2012, pp. 31–44.
- [28] T. Ojala, M. Pietikainen, and T. Maenpaa, "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns," *IEEE Trans. Patt. Anal. Mach. Intell.*, vol. 24, no. 7, pp. 971–987, Jul. 2002.
- [29] R. L. Thorndike, "Who belong in the family?," in *Psychometrika*, 1953.
- [30] R. Zhao, W. Ouyang, and X. Wang, "Person re-identification by saliency matching," in *Proc. ICCV*, 2013.
- [31] D. S. Cheng, M. Cristan, M. Stoppa, L. Bazzani, and V. Murino, "Custom pictorial structures for re-identification," in *Proc. BMVC*, 2011, pp. 68.1–68.11.
- [32] B. Ma, Y. Su, and F. Jurie, "BiCov: A novel image representation for person re-identification and face verification," in *Proc. BMVC*, 2012, pp. 57.1–57.11.
- [33] M. Dikmen, E. Akbas, T. S. Huang, and N. Ahuja, "Pedestrian recognition with a learned metric," in *Proc. ACCV*, 2011, pp. 501–512.
- [34] C. C. Loy, C. Liu, and S. Gong, "Person re-identification by manifold ranking," in *Proc. ICFP*, Sep. 2013, pp. 3567–3571.
- [35] A. Mignon and F. Jurie, "PCCA: A new approach for distance learning from sparse pairwise constraints," in *Proc. CVPR*, Jun. 2012, pp. 2666–2672.
- [36] M. Hirzer, C. Beleznaï, P. M. Roth, and H. Bischof, "Person re-identification by descriptive and discriminative classification," in *Proc. SCIA*, 2011, pp. 91–102.
- [37] K. Q. Weinberger and L. K. Saul, "Distance metric learning for large margin nearest neighbor classification," *J. Mach. Learn. Res.*, vol. 10, pp. 207–244, Jun. 2009.
- [38] J. V. Davis, B. Kulis, P. Jain, S. Sra, and I. S. Dhillon, "Information-Theoretic metric learning," in *Proc. ICML*, 2007, pp. 209–216.