TAICHI DISTANCE FOR PERSON RE-IDENTIFICATION

Zheng Wang[1], *Ruimin Hu*[1,*], *Yi Yu*[2], *Chao Liang*[1] *and Chen Chen*[3]

[1]State Key Laboratory of Software Engineering, School of Computer, Wuhan University, China
[2]Digital Content and Media Sciences Research Division, National Institute of Informatics, Japan
[3]Center for Research in Computer Vision, University of Central Florida, USA

ABSTRACT

Metric learning is an important issue in person re-identification, and Mahalanobis-distance based metric learning methods prevail in this field. All of these approaches can be considered as equivalently projecting all samples to a new metric space and calculating the Euclidean distance there. However, the performance of distinguishing similar samples from dissimilar ones via absolute distance is limited. In this paper, we suggest using relative distance instead. We adopt a bi-target perspective. The core idea is to construct a virtual opposite target for each original target. Then, the similarity between a sample and the others is judged by using both the original and opposite targets of the sample. In this way, we propose a bi-target metric method, named TAICHI distance. Considering simplicity and efficiency, we follow the KISSME metric in this paper. Extensive evaluations on challenging datasets confirm the effectiveness of the proposed method.

Index Terms— person re-identification, metric learning, TAICHI distance

1. INTRODUCTION

Person re-identification, aiming to identify images of the same person from various cameras configured in different places, has attracted much attention in the signal processing community [1–8]. Due to low resolution, motion blur, view change, and illumination variation in the individual's appearance, constructing a discriminative representation to adapt to different camera conditions is extremely challenging [9–16]. Distance metric based methods, which aim at seeking a proper distance metric, have gradually become a main stream procedure in solving the person re-identification problem. In addition, if we could find excellent visual descriptions for person appearance, the metric could still promote the results [17].

* R. Hu is the corresponding author, and is also with the National Engineering Research Center for Multimedia Software, the Collaborative Innovation Center of Geospatial Technology, and the Hubei Provincial Key Laboratory of Multimedia and Network Communication Engineering, Wuhan University.

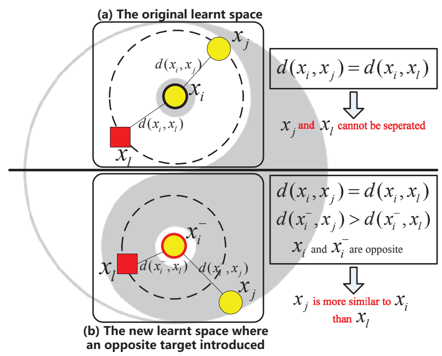


Fig. 1. **An example illustrating the relationships merely with one target (upper) versus after the opposite target is introduced (lower)**. In the figure, x_i denotes the target, x_j denotes the sample similar to the target, x_l denotes the dissimilar sample, and x_i^- denotes the introduced opposite target. (a) The distance $d(x_i, x_j)$ is equal to the distance $d(x_i, x_l)$. (b) By introducing an opposite target x_i^-, the distance $d(x_i^-, x_j)$ is greater than the distance $d(x_i^-, x_l)$.

Traditional distance metrics, such as Euclidean Distance, Minkowski Distance, Manhattan Distance, Chebyshev Distance, are non-learning and fixed metrics, and the distance is directly calculated by the untransformed feature difference. The effectiveness of these handcrafted metric models is proved to be limited because of the appearance's large variation [18]. In comparison, the well-studied Mahalanobis-based distance metric learning helps to find global transformations of the feature space such that relevant feature dimensions are emphasized while irrelevant ones are suppressed, and plays an important role. It has been extended in a number of follow-up methods, including LMNN [19], ITML [20], IDML [21], PRDC [22], LADF [23] and KISSME [24].

The difference among the above methods mainly lies in their different objective functions, which are designed for different specific tasks with different constraints. After the metric matrix $\mathbf{M}$ is learnt, all of the above metric learning methods utilize a uniform form as $d_{\mathbf{M}}(x_i, x_j) = (x_i - x_j)^\top \mathbf{M}(x_i - x_j)$ to obtain the distance between a pair of samples (x_i, x_j). Meanwhile, the distance can be also rewritten as $d_{\mathbf{M}}(x_i, x_j) = \|\mathbf{L}x_i - \mathbf{L}x_j\|^2$, when performing eigenvalue decomposition on $\mathbf{M}$ with $\mathbf{M} = \mathbf{L}^\top \mathbf{L}$. With this definition, it is easy to see that the essence of the metric learning is to seek a suitable projection matrix $\mathbf{L}$, transforming the original feature space to a new one [25]. Then, the distance is computed as an Euclidean distance. Here, we name this kind of distance as **absolute distance**.

As is known, a good distance metric should compute a small distance for a pair of similar samples and a large distance for a pair of dissimilar samples. Let us focus on a special case (Fig.1(a)): after being projected to the new feature space, the distance from the target x_i to a similar sample x_j is the same as that to a dissimilar sample x_l. In this condition, it is still impossible to distinguish the similar sample from the dissimilar one. To solve this kind of problem by breaking the preservation property of the equality relationship in the classic metric learning algorithm, this paper introduces a virtual target, to re-define the transformed sample distance. For instance, according to the target x_i, if we can construct a virtual target x_i^- which is a dissimilar sample of x_i, project this target and samples to a new metric space, and make the distances from the virtual target to x_j and x_l different, the samples can be separated indirectly.

Inspired by the famous Chinese ancient Yin-Yang philosophy: everything in the universe can be viewed as a product of a constant conflict between opposites - Yin and Yang [26], we introduce a bi-target concept. The original target acts as the Yang target. A Yin (invisible) target is introduced, acting as the constructed virtual target described above. However, in the Yin-Yang philosophy, the Yin and the Yang are opposite to each other, and the constructed Yin target should be the dissimilar sample of the Yang target. When the nearness to the Yang target x_i cannot distinguish two samples x_j and x_l, an alternative way is to find the farness of the two samples from the Yin target x_i^-. Here, x_i^- is a duality of x_i in terms of a certain criterion. Fig.1(b) shows that the distance from the Yin target x_i^- to the sample x_l is smaller than that to x_j. Therefore, we consider that x_j is relatively more similar to the target than x_l. We argue that absolute distances between images are not necessarily the best for person re-identification tasks that put more emphasis on the relative order or ranking positions. The essence of the ranking requirement is to identify whether a sample is farther away from or closer to the target than other samples [27, 28]. Therefore, **relative distance** is more important to satisfy these conditions.

Based on the above idea, we propose the **TAICHI distance** to demonstrate the relative distance. Learned from [24], we consider two independent learning process in a statistical perspective. The relative distance is defined by the likelihood ratio of the probability of the distance between the Yang target sample pair to the probability of the distance between the Yin target sample pair. We evaluate our method on the VIPeR dataset [29] and CUHK Campus dataset [30], which outperforms the state-of-the-art metric learning methods. In addition, with two threads running in parallel, it runs as fast as the KISSME method in the training process, and is more efficient than the other approaches.

2. TAICHI DISTANCE

Learning a distance metric based on the class of Mahalanobis distance functions has gained considerable interest in person re-identification. In general, a Mahalanobis distance metric measures the squared distance between two data points by a uniform metric $\mathbf{M}$. It is proved that LMNN, ITML, LDML, PRDC and LDAF rely on an iterative optimization scheme, which is computationally expensive for large scale datasets, while the KISSME introduces a non-iterative formulation, which builds on a statistical inference perspective, and is very effective, simple and fast. Given this, we propose the TAICHI distance also in a statistical inference perspective, although it applies to other metrics as well.

2.1. Yang Metric Learning

We follow the KISSME method [24] to learn the Yang metric. To facilitate the discussion, we make the following definitions. Pairs of samples from similar set $S = \{(x_i, x_j) | y(x_i) = y(x_j)\}$ or dissimilar set $D = \{(x_i, x_j) | y(x_i) \neq y(x_j)\}$ are utilized to train the Mahalanobis-like metric $\mathbf{M}$. Here, $y(\cdot)$ indicates the class label of a sample. Each sample $x_i \in \mathbb{R}^{N_x \times 1}$ is N_x dimension feature vector. $C_{ij} = (x_i - x_j)(x_i - x_j)^\top$ is used to denote the outer product of pairwise differences. And Σ_S and Σ_D are the covariance matrices of S and D, which can be estimated as $\Sigma_S = \frac{1}{|S|} \sum_{(x_i, x_j) \in S} C_{ij}, \Sigma_D = \frac{1}{|D|} \sum_{(x_i, x_j) \in D} C_{ij}$.

Considering two independent generation processes for observed commonalities of similar and dissimilar pairs, the proposed method defines the distance of a sample pair as the probability that it belongs to a dissimilar pair or a similar pair. From a statistical inference point of view the optimal statistical decision on whether a pair is dissimilar or not can be obtained by a likelihood ratio test. Therefore, we test the hypothesis H_0 that a pair (x_i, x_j) is dissimilar against H_1 that it is similar: $\delta(x_i, x_j) = \log(\frac{p(x_i, x_j | H_0)}{p(x_i, x_j | H_1)}) = \log(\frac{f(x_i, x_j, \theta_0)}{f(x_i, x_j, \theta_1)})$, where δ is the log-likelihood ratio, and $f(x_i, x_j, \theta)$ is a PDF (probability density function) with the parameter set θ. A high value of δ means that H_0 is validated. In contrast, a low value means that H_0 is rejected and the pair is considered as similar. To be independent of the actual locality in the feature space, we cast the problem in the space of pairwise differences with zero mean. Assuming a Gaussian structure of the difference space, the equation can be re-written as

$$\delta(x_i, x_j) = \log(\frac{\frac{1}{\sqrt{2\pi\Sigma_D}} \exp(-1/2(x_i - x_j)^\top \Sigma_D^{-1}(x_i - x_j))}{\frac{1}{\sqrt{2\pi\Sigma_S}} \exp(-1/2(x_i - x_j)^\top \Sigma_S^{-1}(x_i - x_j))}). \quad (1)$$

By taking the log and discarding the constant terms, Eq.1 can be simplified to $\delta(x_i, x_j) = (x_i - x_j)^\top(\Sigma_S^{-1} - \Sigma_D^{-1})(x_i - x_j)$. This is the expression of the absolute distance of the Yang Metric, where the metric $\mathbf{M}$ is expressed as $\mathbf{M} = \Sigma_S^{-1} - \Sigma_D^{-1}$. We can see that the metric is computed from C_{ij} of similar pairs and dissimilar pairs, which is depending on the **vector difference** $x_i - x_j$. In the following, we introduce a relative distance, which not only exploits two different metrics, but also brings another form of features relationship.

2.2. Constructing the Pre-train Yin target

All the samples are located in a limited space, and they have a center, which can be estimated from the mean value of the samples. Let u stand for the sample center. u is calculated as $u = \frac{1}{N}\sum_{k=1}^{N} x_k$, where N denotes the number of all the training samples. In the Yin-Yang philosophy, every thing will trend toward a balance of Yin and Yang. We argue that the Yin target and the Yang target are opposite, and in this paper, we choose the symmetric point of the Yang target with respect to the center as the Yin target, so $u = \frac{1}{2}(x_i^+ + x_i^-)$. Then, from $x_i^+ + x_i^- = 2u = v$ and $x_i^+ = x_i$, we directly compute the Yin target as $x_i^- = v - x_i$, where v denotes a constant vector related to the samples center u. Following this rule, for each target x_i, its corresponding pre-train Yin target is constructed as $x_i^- = \frac{2}{N}\sum_{k=1}^{N} x_k - x_i$. For each sample pair (x_i, x_j), a new sample pair (x_i^-, x_j) is constructed. Then pairs of samples form the similar set $S' = \{(x_i^-, x_j) | y(x_i) \neq y(x_j)\}$ and dissimilar set $D' = \{(x_i^-, x_j) | y(x_i) = y(x_j)\}$ for Yin Metric learning.

2.3. TAICHI Metric Learning

Whether a Yin pair is dissimilar or not is also obtained by a likelihood ratio test as the Yang Metric Learning does. By duplicating and expanding the progress in Yang Metric Learning, we test not only the hypothesis H_0 that a pair (x_i, x_j) is dissimilar against H_1 that the pair is similar, but also the hypothesis H_1' that a pair (x_i^-, x_j) is similar against H_0' that the pair is dissimilar: $\delta'(x_i, x_j) = \log(\frac{p(x_i,x_j|H_0)p(x_i^-,x_j|H_1')}{p(x_i,x_j|H_1)p(x_i^-,x_j|H_0')}) = \log(\frac{f(x_i,x_j,\theta_0)f(x_i^-,x_j,\theta_1')}{f(x_i,x_j,\theta_1)f(x_i^-,x_j,\theta_0')})$, where δ' is the new log-likelihood ratio or relative distance, and $f(x_i^-, x_j, \theta')$ is a PDF with the parameter set θ'. Assuming zero-mean Gaussian distributions, and taking the log and discarding the constant terms, the equation can be simplified to Eq.2. Here, $\Sigma_{S'}$ and $\Sigma_{D'}$ are the covariance matrices of S' and D'. To increase readability, we introduce the notation $B_{ij} = (x_i^- - x_j)(x_i^- - x_j)^\top$. S' and D' can be respectively computed by $\Sigma_{S'} = \frac{1}{|S'|}\sum_{(x_i^-,x_j)\in S'} B_{ij}, \Sigma_{D'} = \frac{1}{|D'|}\sum_{(x_i^-,x_j)\in D'} B_{ij}$.

$$\begin{aligned}\delta'(x_i, x_j) &= (x_i - x_j)^\top(\Sigma_S^{-1} - \Sigma_D^{-1})(x_i - x_j) \\ &- (x_i^- - x_j)^\top(\Sigma_{S'}^{-1} - \Sigma_{D'}^{-1})(x_i^- - x_j).\end{aligned} \quad (2)$$

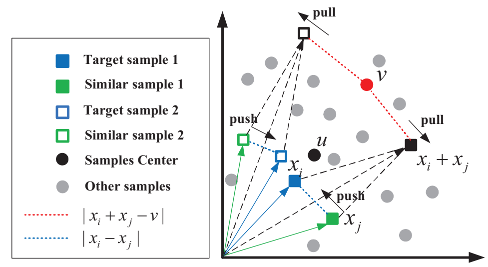


Fig. 2. **Two pairs of samples in a 2D sample space**. The blue line denotes the distance of a sample pair (vector difference), which traditional metric learning method tries to optimize. The red line denotes the distance between the sample v and the vector commonness, which the new metric tries to enlarge.

From the Eq.2, we can see that the learning process actually obtains two metrics in constructing relative distances. The **Yang metric** is the same as the original one $\mathbf{M}_{yang} = \Sigma_S^{-1} - \Sigma_D^{-1}$, and the **Yin metric** can be expressed as $\mathbf{M}_{yin} = \Sigma_{S'}^{-1} - \Sigma_{D'}^{-1}$. The metrics are respectively trained from C_{ij} and B_{ij}, which are depending on not only the vector difference, but also the other feature relationship introduced by the Yin target.

3. ANALYSIS OF TAICHI DISTANCE

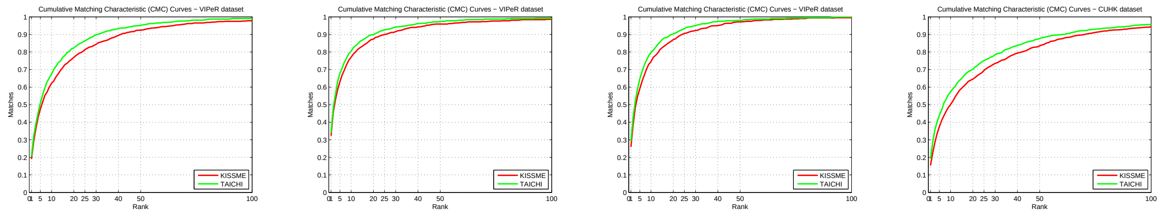
Considering Yin target $x_i^- = v - x_i$, Eq.2 changes to

$$\begin{aligned}\delta'(x_i, x_j) &= (x_i - x_j)^\top \mathbf{M}_{yang}(x_i - x_j) \\ &- (x_i + x_j - v)^\top \mathbf{M}_{yin}(x_i + x_j - v).\end{aligned} \quad (3)$$

The above distance function has the following desirable properties:

(1) Fully informed distance decision. The distance function depends not only on $x_i - x_j$, the **vector difference** usually considered by conventional metric learning, but also on the **vector commonness** $x_i + x_j$, which contains orthogonal information of (x_i, x_j) that would otherwise be neglected when using $x_i - x_j$ alone.

(2) Distances more discriminative. Eq.3 utilizes both the metric $\mathbf{M}_{yang}$ and $\mathbf{M}_{yin}$, where the new metric $\mathbf{M}_{yin}$ is designed by $B_{ij} = (x_i + x_j - v)(x_i + x_j - v)^\top$. This has been neglected before where only $\mathbf{M}_{yang}$ is used. Fig.2 demonstrates the effectiveness of $\mathbf{M}_{yang}$ and $\mathbf{M}_{yin}$ in a 2D sample space. It shows that traditional metric $\mathbf{M}_{yang}$ considers to pull the similar samples near to the targets, while the new metric $\mathbf{M}_{yin}$ simultaneously give more power to push the vector commonness $x_i + x_j$ away from the sample v. This will make the sample pairs more decentralized, at the same time, decrease the confusing possibility.



(a) VIPeR/Basic features (b) VIPeR/SCNCD (c) VIPeR/LOMO (d) CUHK/LOMO

Fig. 3. **Person re-identification results.** In (a) (b) (c), methods are evaluated on the VIPeR dataset. In (a), basic features are used. In (b) the SCNCD feature is exploited. In (c) the LOMO feature is exploited. In (d) the LOMO feature is exploited on the CUHK Campus dataset.

4. EXPERIMENTS

Datasets: To show the effectiveness of the TAICHI distance, we conduct experiments on two different datasets with different features. The widely used VIPeR dataset [29] contains 632 person image pairs of two different camera views. All images of individuals are normalized to a size of 128×48 pixels. Most of the example pairs contain a viewpoint change, making the dataset one of the most challenging datasets currently available for person re-identification. As the compared metric learning approaches did, we divide 632 image pairs randomly into two sets of 316 image pairs each, one for training and the other for testing. The CUHK Campus dataset [30] was captured with two camera views in a campus environment. It contains 971 persons, and each person has two images in each camera view. All images were scaled to 160×60 pixels. The persons were split to two groups, 485 for training and 486 for test.

Features: In order to demonstrate the independence of the proposed method to different features. We conducted experiments on the basic features described in [24], the salient color names (SCNCD [11]) feature, and the Local Maximal Occurrence Representation (LOMO [10]) feature.

This paper reports Cumulative Matching Characteristic (CMC) [31] curves of various algorithms, which represent the expectation of the true match found within the first n ranks. To obtain a reasonable statistical significance, the experiment is repeated 20 times, and the average results are reported in Fig.3(a), Fig.3(b), Fig.3(c) and Fig.3(d). From the figures, we can conclude that the proposed TAICHI distance significantly outperforms the KISSME method. Moreover, in Table 1, this paper compares the performance of our approach in the range of the first 50 ranks to state-of-the-art methods on the VIPeR dataset. As can be seen, we obtain competitive results when the TAICHI distance is used.

Generally speaking, training time cost is one of the main evaluating indicator of metric learning approaches. We can observe a common fact that traditional methods all rely on an iterative optimization scheme which is computationally expensive. In comparison, a non-iterative metric learning method, which builds on a statistical inference perspective, is very fast, especially when the amount of data constantly grows. Following the non-iterative framework KISSME, from the Table 1, we can also see that the TAICHI distance is computationally much more efficient than LADF and PRDC.

Table 1. Person re-identification matching rates (%) on the VIPeR dataset.

Method	1	10	25	50	time
LADF [23]	30	79	93	97	781s
PRDC [22]	19.9	49.4	70.5	84.8	904s
SCNCD [11]	20.7	60.6	79.1	90.4	–
KISSME [24]	19.6	62.2	80.7	91.8	**0.007**s
TAICHI	20.54	67.51	85.83	95.02	**0.007**s
KISSME+LOMO	25.95	74.53	90.51	97.15	0.019s
TAICHI+LOMO	29.11	79.75	**93.35**	**97.94**	0.019s
KISSME+SCNCD	32.52	76.67	89.97	94.56	0.010s
TAICHI+SCNCD	**33.68**	**80.38**	93.04	97.31	0.010s

5. CONCLUSION

We convert the traditional absolute distance task to a more proper relative distance task, and adopt a bi-target perspective to reform existing metric learning methods and proposed the TAICHI distance. Our method improves the original KISSME method significantly, and also achieves the best results compared to state-of-the-art metric learning approaches.

6. ACKNOWLEDGMENT

The research was supported by National High Technology Research and Development Program of China (2015AA016306), Natural Science Foundation of China (61231015, 61671336, 61671332, 61562048), Natural Science Foundation of Jiang-Su Province (BK20160386), Natural Science Foundation of Hubei Province (2016CFB573), Technology Research Program of Ministry of Public Security (2016JSYJA12), EU FP7 QUICK project (PIRSES-GA-2013-612652), and Jiangxi Youth Science Foundation of China (20151BAB217013).

References

[1] Cong Ma, Zhenjiang Miao, and Min Li, "Saliency preprocessing for person re-identification images," in *ICASSP*, 2016.

[2] Liu Hong, Ma Liqian, and Wang Can, "Body-structure based feature representation for person re-identification," in *ICASSP*, 2015.

[3] J. Oliver, Antonio Albiol, Antonio Albiol, and Jose Manuel Mossi, "Re-identifying people in the wild," in *ICASSP*, 2013.

[4] Kai Liu, Xin Guo, Zhicheng Zhao, and Anni Cai, "Person re-identification using matrix completion," in *ICIP*, 2013.

[5] Yanna Zhao, Xu Zhao, and Yuncai Liu, "Person re-identification by free energy score space encoding," in *ICIP*, 2014.

[6] Junjun Jiang, Ruimin Hu, Zhongyuan Wang, and Zhihua Cai, "Cdmma: Coupled discriminant multi-manifold analysis for matching low-resolution face images," *Signal Process.*, vol. 124, pp. 162–172, 2016.

[7] Chen Chen, Mengyuan Liu, Baochang Zhang, Jungong Han, Junjun Jiang, and Hong Liu, "3d action recognition using multi-temporal depth motion maps and fisher vector," in *IJCAI*, 2016.

[8] Mengyuan Liu, Hong Liu, Chen Chen, and Maryam Najafian, "Energy-based global ternary image for action recognition using sole depth sequences," in *3DV*, 2016.

[9] Michela Farenzena, Loris Bazzani, Alessandro Perina, Vittorio Murino, and Marco Cristani, "Person re-identification by symmetry-driven accumulation of local features," in *CVPR*, 2010.

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

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

[12] Rui Zhao, Wanli Ouyang, and Xiaogang Wang, "Unsupervised salience learning for person re-identification," in *CVPR*, 2013.

[13] Le An, Xiaojing Chen, Songfan Yang, and Xuelong Li, "Person re-identification by multi-hypergraph fusion," *IEEE Trans. Neural Netw. Learn. Syst.*, 2016.

[14] Le An, Mehran Kafai, Songfan Yang, and Bir Bhanu, "Person reidentification with reference descriptor," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 26, no. 4, pp. 776–787, 2016.

[15] Le An, Xiaojing Chen, Songfan Yang, and Bir Bhanu, "Sparse representation matching for person re-identification," *Inform. Sciences*, vol. 355, pp. 74–89, 2016.

[16] Jin Wang, Zheng Wang, Changxin Gao, Nong Sang, and Rui Huang, "Deeplist: Learning deep features with adaptive listwise constraint for person re-identification," *IEEE Trans. Circuits Syst. Video Technol.*, 2016.

[17] Zheng Wang, Ruimin Hu, Chao Liang, and Yi Yu, "Zero-shot person re-identification via cross-view consistency," *IEEE Trans. Multimedia*, vol. 18, no. 2, pp. 260–272, 2016.

[18] Mang Ye, Chao Liang, Yi Yu, Zheng Wang, Qingming Leng, Chunxia Xiao, Jun Chen, and Ruimin Hu, "Person re-identification via ranking aggregation of similarity pulling and dissimilarity pushing," *IEEE Trans. Multimedia*, vol. 18, no. 12, pp. 2553, 2016.

[19] Kilian Q. Weinberger and Lawrence K. Saul, "Distance metric learning for large margin nearest neighbor classification," *J. Mach. Learn. Res.*, vol. 10, no. 1, pp. 207–244, 2009.

[20] Jason V Davis, Brian Kulis, Prateek Jain, Suvrit Sra, and Inderjit S Dhillon, "Information-theoretic metric learning," in *NIPS*, 2007.

[21] Matthieu Guillaumin, Jakob Verbeek, and Cordelia Schmid, "Is that you? metric learning approaches for face identification," in *ICCV*, 2009.

[22] Wei-Shi Zheng, Shaogang Gong, and Tao Xiang, "Person re-identification by probabilistic relative distance comparison," in *CVPR*, 2011.

[23] Zhen Li, Shiyu Chang, Feng Liang, Thomas S Huang, Liangliang Cao, and John R Smith, "Learning locally-adaptive decision functions for person verification," in *CVPR*, 2013.

[24] Martin Kostinger, Martin Hirzer, Paul Wohlhart, Peter M Roth, and Horst Bischof, "Large scale metric learning from equivalence constraints," in *CVPR*, 2012.

[25] Yimin Wang, Ruimin Hu, Chao Liang, and Chunjie Zhang, "Camera compensation using a feature projection matrix for person reidentification," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 24, no. 8, pp. 1–6, 2013.

[26] Minh Nhut Nguyen, Daming Shi, and C Quek, "Fcmac-byy: Fuzzy cmac using bayesian ying-yang learning," *IEEE Trans. Syst. Man, Cybern. B, Cybern.*, vol. 36, no. 5, pp. 1180–90, 2006.

[27] David G Lowe, "Distinctive image features from scale-invariant keypoints," *Int. J. Comput. Vision*, vol. 60, no. 2, pp. 91–110, 2004.

[28] Junjun Jiang, Ruimin Hu, Zhongyuan Wang, Zhen Han, and Jiayi Ma, "Facial image hallucination through coupled-layer neighbor embedding," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 26, no. 9, pp. 1–1, 2015.

[29] Douglas Gray, Shane Brennan, and Hai Tao, "Evaluating appearance models for recognition, reacquisition, and tracking," in *PETS*, 2007.

[30] Wei Li, Rui Zhao, and Xiaogang Wang, "Human reidentification with transferred metric learning," in *ACCV*, 2012.

[31] Xiaogang Wang, Gianfranco Doretto, Thomas Sebastian, Jens Rittscher, and Peter Tu, "Shape and appearance context modeling," in *ICCV*, 2007.