

Human Identity and Gender Recognition From Gait Sequences With Arbitrary Walking Directions

Jiwen Lu, *Member, IEEE*, Gang Wang, *Member, IEEE*, and Pierre Moulin, *Fellow, IEEE*

Abstract—We investigate the problem of human identity and gender recognition from gait sequences with arbitrary walking directions. Most current approaches make the unrealistic assumption that persons walk along a fixed direction or a pre-defined path. Given a gait sequence collected from arbitrary walking directions, we first obtain human silhouettes by background subtraction and cluster them into several clusters. For each cluster, we compute the cluster-based averaged gait image as features. Then, we propose a sparse reconstruction based metric learning method to learn a distance metric to minimize the intra-class sparse reconstruction errors and maximize the inter-class sparse reconstruction errors simultaneously, so that discriminative information can be exploited for recognition. The experimental results show the efficacy of our approach.

Index Terms—Human gait analysis, identity recognition, gender recognition, metric learning, sparse reconstruction.

I. INTRODUCTION

HERE has been much work on human gait analysis over the past decades for human identification [2]–[10], gender classification [11]–[13], and age estimation [14]–[16]. Several benchmark gait datasets [5], [17]–[21] are also available. However, most datasets and gait recognition approaches assume that people walk along a fixed direction or a pre-defined path. This is generally unrealistic because people usually walk freely in the scene and the walking directions may be time-varying.

In this paper, we investigate the problem of human identity and gender recognition from gait sequences with arbitrary walking directions, so that a more practical human gait analysis system can be developed. Since human gait analysis is known to be sensitive to the varying views [19], [22]–[24], human identification and gender recognition from gait

Manuscript received May 16, 2013; revised August 5, 2013 and November 11, 2013; accepted November 15, 2013. Date of publication November 20, 2013; date of current version December 16, 2013. This work was supported by the research grant for the Human Sixth Sense Program with the Advanced Digital Sciences Center from the Agency for Science, Technology and Research of Singapore. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Jaihie Kim.

J. Lu is with the Advanced Digital Sciences Center, Singapore 138632 (e-mail: jiwen.lu@adsc.com.sg).

G. Wang is with the School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore 639798, and also with the Advanced Digital Sciences Center, Singapore 138632 (e-mail: wanggang@ntu.edu.sg).

P. Moulin is with the Department of Electrical and Computer Engineering, University of Illinois at Urbana-Champaign, Urbana, IL 61801 USA, and also with the Advanced Digital Sciences Center, Singapore 138632 (e-mail: moulin@ifp.uiuc.edu).

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/TIFS.2013.2291969



Fig. 1. Five frames in a single gait sequence of our dataset which was captured from arbitrary walking directions. From top to bottom are the original gait images, binary segmented silhouette images, and normalized and aligned silhouette images of one gait sequence, respectively.

sequences with arbitrary walking directions are difficult. To study this new problem, we first collect a new gait dataset, where people walk freely in the scene, and the walking directions are arbitrary and time-varying throughout the sequence. Fig. 1 shows some frames of a gait sequence from our dataset, as well as the segmented and aligned human silhouettes.

We obtain human silhouettes by background subtraction in each gait sequence and cluster them into several clusters. For each cluster, we compute the cluster-based averaged gait image (C-AGI) as the feature. Subsequently, we propose a metric learning approach to learn a distance metric to extract more discriminative information for recognition. Experimental results show the efficacy of our proposed approach.

The main contributions of this work are:

- We have constructed a new gait database named ADSC-AWD (Advanced Digital Sciences Center-Arbitrary Walking Directions) using the Microsoft Kinect depth sensor. We have made this dataset and labels publicly available online.¹
- We have proposed a sparse reconstruction based metric learning (SRML) method for discriminative gait feature extraction. The proposed SRML method can minimize the intra-class sparse reconstruction errors and maximize the inter-class sparse reconstruction errors simultaneously, so that more discriminative information can be exploited for human identity and gender recognition.
- We have applied our approach to the widely used USF and CASIA-B gait databases and achieved competitive recognition performance with state-of-the-art methods.
- We have conducted a number of human identity and gender recognition experiments from gait sequences with arbitrary walking directions. Our experimental results

¹Available at: <https://sites.google.com/site/adscawdgait/>.

have demonstrated the effectiveness of our approach to this challenging problem.

The rest of the paper is organized as follows. Section II briefly reviews related work. Section III presents our dataset. Section IV details our proposed approach. Section V provides the experimental results, and Section VI concludes the paper.

II. RELATED WORK

In this section, we briefly review three related topics: 1) view-invariant gait recognition, 2) gait-based gender classification, and 3) metric learning.

A. View-Invariant Gait Recognition

The most relevant work is view-invariant human gait recognition [22]–[29]. Existing view-invariant gait recognition methods can be classified into two categories: model-based [26] and motion-based [28], [29]. Generally, model-based methods aim to extract view-invariant part features in human body, which are matched for gait sequences captured across views. Motion-based methods describe human gait sequences with view-invariance feature representations. However, most existing view-invariant gait recognition methods assume that the view is consistent or the walking path is predefined in a single video clip except [30]. As we mentioned before, this assumption is unrealistic in many real applications because people usually walk freely and the walking direction may be time-varying. While Huang and Boulgouris [30] proposed a robust gait recognition system where the walking direction changes during the walking period, the direction changes within each gait sequence are still not very significant. Unlike these studies, we attempt to identify people and recognize their genders from their gait sequences which are captured from arbitrary walking directions.

B. Gait-Based Gender Classification

Recent work on gait-based gender classification includes [11]–[13], [22]. For example, Yu *et al.* [12] and Li *et al.* [11] partitioned each gait energy image (GEI) [31] into several different local parts such as head, chest, and legs, and performed classification with support vector machine (SVM) on each local part. The outputs of different local parts were then combined and fused. Lu and Tan [22] presented a view-invariant gait-based gender classification method with a subspace learning approach. Hu *et al.* [13] proposed a mixed conditional random field method for gait-based gender classification. Unlike these existing methods, we aim to recognize human genders from their gait sequences which are captured from arbitrary walking directions.

C. Metric Learning

A number of metric learning algorithms have been proposed in recent years [32]–[34]. Representative methods include neighborhood component analysis (NCA) [32], large margin nearest neighbor (LMNN) [33], and information theoretic metric learning (ITML) [34]. While these methods achieve

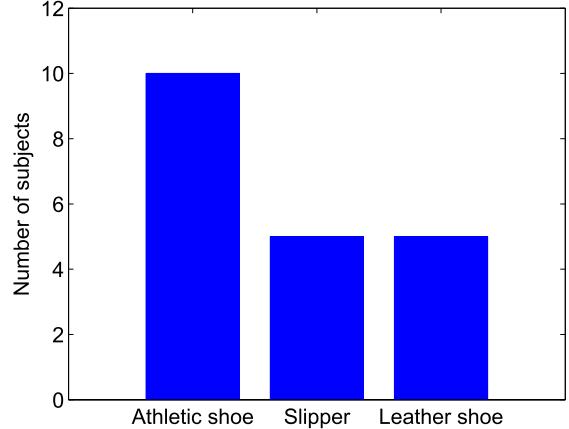


Fig. 2. Statistics in terms of the distribution of the number of subjects with respect to the type of shoes.

encouraging performance in many computer vision applications such as face recognition [35], human activity recognition [36], person re-identification [37], image retrieval [38], object recognition [39] and visual tracking [40], most of them learn the distance metric to force data points from the same category to be close. This approach cannot effectively handle large intra-class variations of the samples. For our application, there are large inter-class variations for gait sequences captured from arbitrary walking direction. To address this, we propose to learn a reconstruction-based distance metric by using a set of samples to estimate these variations. This is more suitable for the problem of gait-based human identity and gender recognition from arbitrary walking directions.

III. DATASETS

Several benchmark databases for gait recognition are available [5], [17]–[21]. Each gait sequence in these databases was collected in controlled environments where humans walk along a fixed direction except the USF [18] and ACTIBIO [30] gait databases. The walking direction of the USF gait dataset was restricted to an elliptical path and hence people cannot walk freely. For the ACTIBIO gait database, the changes of the walking directions in each gait sequences are not very significant.

To study the problem of gait-based human identity and gender recognition from arbitrary walking directions, we collect a new ADSC-AWD dataset using the Microsoft Kinect depth sensor. Compared with the RGB cameras used by the aforementioned datasets, the depth sensor preserves human privacy information, which could make gait recognition easier to be accepted by the public. Still our proposed approach can also be applied on RGB videos. The Kinect depth camera is fixed on a tripod and captured gait sequences on two different days in two rooms. The size of each room is about $8 \times 6 \times 4$ meters. The distance from our camera to each subject is between 1.5 to 5 meters. There are 20 subjects (13 males and 7 females) in our dataset, each of which walks freely in the room and wears one of three types of shoes: athletic shoe, slipper, and leather shoe. The subjects do not wear coats. The distribution of the number of subjects with respect to the type of shoes is shown in Fig. 2.



Fig. 3. Some sample images from two gait sequences of two subjects in our ADSC-AWD dataset. The top two rows are image frames of one male subject captured in the first room and the bottom two rows are image frames of one female subject captured in the second room, respectively.

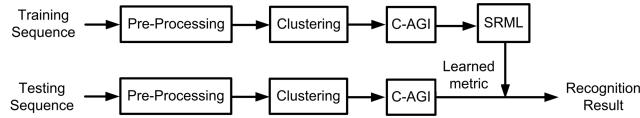


Fig. 4. Flow chart of the proposed approach.

The gait sequences for each subject is collected four times, hence there is a total of 80 gait sequences in our ADSC-AWD database. These depth sequence images are captured at a rate of 30 frames per second, and the original resolution is 320×240 . The length of each gait sequence varies from 300 to 800 frames. Some sample images of two subjects are shown in Fig. 3. Hence, our ADSC-AWD dataset is complement to the existing gait databases and provides a new benchmark dataset.

IV. PROPOSED APPROACH

Fig. 4 shows the flow-chart of our approach. For each gait sequence, we obtain human silhouettes by background subtraction and adaptively cluster them into several clusters. For each cluster, we compute the cluster-based averaged gait image (C-AGI) as the gait feature. Then, we learn a distance metric under which the intra-class sparse reconstruction errors are minimized and the inter-class sparse reconstruction errors are maximized, simultaneously. The approach is detailed in the following subsections.

A. Pre-Processing

Given a gait sequence, the human silhouettes are extracted by background subtraction using the approach in [41]. To make gait feature representation insensitive to the distance between the camera and the subject, we align each gait silhouette into 64×44 following the method in [18]. Fig. 1 shows preprocessing results for several gait frames. We see there are large variations within the same gait sequence because

the person walks from arbitrary directions and the viewpoint between the person and the camera in each frame is arbitrary. The subject's pose in each frame cannot be easily predicted.

B. Clustering and Feature Extraction

The gait energy image (GEI) feature is powerful in representing human gaits owing to its robustness against preprocessing noises [31]. Since persons walk freely, it is very challenging to estimate the gait period in the gait sequence [10], [18]. Moreover, previous studies have shown that the GEI feature is sensitive to the varying views [19], [22]. Hence we cannot compute the GEI feature for the whole gait sequence directly. To address this issue, we cluster each gait sequence into many clusters. Each cluster is expected to gather human silhouettes of similar views or poses. Our previous work [1] used the K -means to obtain the clusters for each gait sequence. However, the number of clusters is not known in advance because the view and pose variations in a testing gait sequence are unknown. To address this, we apply the affinity propagation (AP) clustering method [42] to obtain the clusters without pre-specifying the cluster number for each gait sequence. In our dataset, each sequence is clustered into 8-15 clusters.

Given a gait sequence, assume there are N_k frames in the k^{th} cluster. We then compute the cluster-based averaged gait image (C-AGI) as the gait feature:

$$G_k(x, y) = \frac{1}{N_k} \sum_{p=1}^{N_k} I_{pk}(x, y) \quad (1)$$

where $I_{pk}(x, y)$ is the p^{th} human silhouette in the k^{th} cluster, and x and y are 2-D image coordinates. Our C-AGI feature differs from the conventionally used GEI feature what was proposed in [31]. Our C-AGI feature is the average silhouette of all frames within the same cluster where the GEI feature is the average silhouette of a complete gait cycle. Previous studies have shown that there is not a significant performance difference on these two features [19]. However, due to the fact that gait period estimation is very challenging for sequences captured from arbitrary walking directions, we use the C-AGI feature in our experiments.

Fig. 5 shows forty C-AGIs for two subjects' gait sequences. A pixel with higher intensity value in C-AGI represents more variations of static poses in gait feature, and a pixel with lower intensity means that human walking occurs more frequently at this position and more dynamic information is represented.

C. SRML

For each gait sequence, we obtain multiple C-AGI features using the above pre-processing and feature extraction procedures, and these C-AGIs describe human gaits from different views and poses. Let $X = [X^1, X^2, \dots, X^c]$ be the training set of c different subjects, where $X^c = [x_{11}^c, x_{12}^c, \dots, x_{1K_{1c}}^c, \dots, x_{LK_{ic}}^c]$ denotes the C-AGI feature set of the c^{th} subject, $x_{ij}^c \in R^d$ is the C-AGI feature of the j^{th} cluster in the i^{th} gait sequence of the c^{th} subject, L is the number of gait sequences per subject, and K_{ic} is the number of C-AGI features of the i^{th} gait sequence of the c^{th} subject.



Fig. 5. Forty C-AGIs of two subjects' gait sequences. The top two rows show C-AGIs of the first person and the bottom two rows show C-AGIs of the second person.

We aim to learn a discriminative distance metric for recognition. Conventional metric learning methods [32], [33], [34] usually force two data samples from the same category to be close, and two data samples from different categories to be far. However, our C-AGI features exhibit too much intra-class variations, because they represent human gaits from many views and poses. We argue that these C-AGIs have strong group correlations since the views and poses they represent lie in a continuous space. Hence one C-AGI feature is expected to be more similar to a linear combination of several C-AGIs rather than to a single one. Based on this intuition, we develop a new metric learning method which we call sparse reconstruction based metric learning (SRML) to minimize the intra-class reconstruction errors and maximize the inter-class reconstruction errors, simultaneously. We enforce the sparse reconstruction constraint due to its great success in face and object recognition [43].

We first show how to calculate the reconstruction coefficients given a distance metric M . For a data sample x_{ij}^c , the intra-class and inter-class reconstruction errors are calculated as below:

$$\min_{a_{ij}^c} (x_{ij}^c - H_{ij}^c a_{ij}^c)^T M (x_{ij}^c - H_{ij}^c a_{ij}^c) + \lambda \|a_{ij}^c\|_1 \quad (2)$$

$$\min_{b_{ij}^c} (x_{ij}^c - G_{ij}^c b_{ij}^c)^T M (x_{ij}^c - G_{ij}^c b_{ij}^c) + \lambda \|b_{ij}^c\|_1 \quad (3)$$

where a_{ij}^c and b_{ij}^c are the corresponding intra-class and inter-class reconstruction coefficients, M is a $d \times d$ square, symmetric, and positive semi-definite matrix to satisfy the distance metric requirement, H_{ij}^c and G_{ij}^c are two dictionaries constructed from all the intra-class and inter-class samples of

x_{ij}^c , $1 \leq i \leq L$, $1 \leq j \leq K_{ic}$, $1 \leq c \leq C$. Finally, the parameter λ balances the reconstruction error and the sparsity of the reconstruction coefficients, and is empirically specified as 0.01 in our experiments.

In the training procedure, we make the intra-class reconstruction variations for the whole training set as small as possible and the inter-class reconstruction variations for the whole training set as large as possible simultaneously, by learning the distance metric M , so that more discriminative information can be exploited for recognition. We formulate the objective function as shown in Eq. (4) (at the bottom of the page), where A and B are the corresponding reconstruction coefficient matrices. The objective function J_1 is to ensure that each x_{ij}^c is sparsely reconstructed by its intra-class samples and the reconstruction error is as small as possible. Similarly, the objective function J_2 is to ensure each x_{ij}^c is sparsely reconstructed by its inter-class samples and the reconstruction error is as large as possible under the learned distance metric M . To the best of our knowledge, there is no closed-form solution to such a max-min optimization problem [44] because we aim to learn M but have to infer A and B simultaneously. Hence, we solve this problem iteratively similar to previous EM-like learning algorithms [45]. The basic idea is to fix M , update A and B , and then update M , etc, iteratively.

We first initialize M as the Euclidean metric, and obtain A and B by solving the following l^1 -minimization problems [43], [46]

$$a_{ij}^c = \arg \min \|a_{ij}^c\|, \text{ s.t. } \|x_{ij}^c - H_{ij}^c a_{ij}^c\|_2^2 \leq \varepsilon. \quad (5)$$

$$b_{ij}^c = \arg \min \|b_{ij}^c\|, \text{ s.t. } \|x_{ij}^c - G_{ij}^c b_{ij}^c\|_2^2 \leq \varepsilon. \quad (6)$$

for a given error tolerance $\varepsilon > 0$.

Having obtained A and B , (2) and (3) can be combined and rewritten as the following optimization problem:

$$\begin{aligned} \max_M J = & \sum_{c=1}^C \sum_{i=1}^L \sum_{j=1}^{K_{ci}} (x_{ij}^c - G_{ij}^c b_{ij}^c)^T M (x_{ij}^c - G_{ij}^c b_{ij}^c) \\ & - \sum_{c=1}^C \sum_{i=1}^L \sum_{j=1}^{K_{ci}} (x_{ij}^c - H_{ij}^c a_{ij}^c)^T M (x_{ij}^c - H_{ij}^c a_{ij}^c). \end{aligned} \quad (7)$$

Since M is positive semi-definite and symmetric, there exists a non-square matrix W of size $d \times l$, where $l \leq d$, such that

$$M = WW^T \quad (8)$$

$$\begin{aligned} \max_{M,A,B} J = & \max_M \left(\min_B J_2 - \min_A J_1 \right) = \max_M \left(\min_B \sum_{c=1}^C \sum_{i=1}^L \sum_{j=1}^{K_{ci}} (x_{ij}^c - G_{ij}^c b_{ij}^c)^T M (x_{ij}^c - G_{ij}^c b_{ij}^c) + \lambda \|b_{ij}^c\|_1 \right) \\ & - \left(\min_A \sum_{c=1}^C \sum_{i=1}^L \sum_{j=1}^{K_{ci}} (x_{ij}^c - H_{ij}^c a_{ij}^c)^T M (x_{ij}^c - H_{ij}^c a_{ij}^c) + \lambda \|a_{ij}^c\|_1 \right) \end{aligned} \quad (4)$$

Combining (7) and (8), we simplify J as follows:

$$\begin{aligned} J &= \sum_{c=1}^C \sum_{i=1}^L \sum_{j=1}^{K_{ci}} (x_{ij}^c - G_{ij}^c b_{ij}^c)^T W W^T (x_{ij}^c - G_{ij}^c b_{ij}^c) \\ &\quad - \sum_{c=1}^C \sum_{i=1}^L \sum_{j=1}^{K_{ci}} (x_{ij}^c - H_{ij}^c a_{ij}^c)^T W W^T (x_{ij}^c - H_{ij}^c a_{ij}^c) \\ &= \text{tr}[W^T \sum_{c=1}^C \sum_{i=1}^L \sum_{j=1}^{K_{ci}} (x_{ij}^c - G_{ij}^c b_{ij}^c)(x_{ij}^c - G_{ij}^c b_{ij}^c)^T W] \\ &\quad - \text{tr}[W^T \sum_{c=1}^C \sum_{i=1}^L \sum_{j=1}^{K_{ci}} (x_{ij}^c - H_{ij}^c a_{ij}^c)(x_{ij}^c - H_{ij}^c a_{ij}^c)^T W] \\ &= \text{tr}[W^T (Z_1 - Z_2) W] \end{aligned} \quad (9)$$

where

$$Z_1 \triangleq \sum_{c=1}^C \sum_{i=1}^L \sum_{j=1}^{K_{ci}} (x_{ij}^c - G_{ij}^c b_{ij}^c)(x_{ij}^c - G_{ij}^c b_{ij}^c)^T, \quad (10)$$

$$Z_2 \triangleq \sum_{c=1}^C \sum_{i=1}^L \sum_{j=1}^{K_{ci}} (x_{ij}^c - H_{ij}^c a_{ij}^c)(x_{ij}^c - H_{ij}^c a_{ij}^c)^T. \quad (11)$$

Now, we formulate SRML as the following constrained optimization problem:

$$\begin{aligned} \max_W J(W) &= \text{tr}[W^T (Z_1 - Z_2) W] \\ \text{subject to } W^T W &= I \end{aligned} \quad (12)$$

where $W^T W = I$ is a constraint to restrict the scale of W so that the optimization problem in (12) with respect to W is well-posed. Then, W can be obtained by solving the following eigenvalue problem

$$(Z_1 - Z_2)w = \lambda w. \quad (13)$$

Let w_1, w_2, \dots, w_l be the eigenvectors of (13) corresponding to the l largest eigenvalues ordered according to $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_l$. The $d \times l$ transformation matrix $W = [w_1, w_2, \dots, w_l]$ is used to project each C-AGI feature sample x_{ij}^c into low-dimensional feature vectors y_{ij}^c , as follows:

$$y_{ij}^c = W^T x_{ij}^c. \quad (14)$$

where $1 \leq i \leq L$, $1 \leq j \leq K_{ci}$ and $1 \leq c \leq C$.

Having obtained W and y_{ij}^c , we re-calculate the sparse reconstruction coefficients A and B in (5) and (6) by using y_{ij}^c instead of x_{ij}^c . Then, we update W by re-solving the eigenvalue equation in Eq. (13). The proposed SRML algorithm is summarized in **Algorithm 1**.

D. Classification

Given a testing gait sequence T , we first extract human silhouettes using background subtraction and then cluster them into K groups. For each group, we calculate the C-AGI as the gait feature: T_1, T_2, \dots, T_K . For the k^{th} C-AGI T_k , we

Algorithm 1: SRML

Input: Training set: $X = [X^1, X^2, \dots, X^C]$, iteration number T , convergence error ϵ , where $X^c = [x_{11}^c, x_{12}^c, \dots, x_{1K}^c, x_{21}^c, \dots, x_{LK_{ci}}^c]$.

Output: Distance metric W .

Step 1 (Initialization):

Initialize $W: W = I^{d \times d}$.

Step 2 (Local optimization):

For $r = 1, 2, \dots, T$, repeat

- 2.1. Compute A and B by using (5) and (6).
- 2.2. Compute Z_1 and Z_2 of (10) and (11).
- 2.3. Solve the eigenvalue problem in (13).
- 2.4. Obtain $W^r = [w_1, w_2, \dots, w_l]$.
- 2.5. Update $x_{ij}^c: x_{ij}^c = W^T x_{ij}^c$.
- 2.6. If $r > 2$ and $|W^r - W^{r-1}| < \epsilon$, go to Step 3.

Step 3 (Output distance metric):

Output distance metric $W = W^r$.

calculate the residuals $r_c(T_k)$ using Sparse Representation-based Classification (SRC) [43] under the learned metric M . Then, we assign a label z to T as follows:

$$z = \arg \min_c \sum_{k=1}^K r_c(T_k), \quad c = 1, 2, \dots, C \quad (15)$$

where

$$r_c(T_k) = \|W^T T_k - W^T X \delta_c(T_k)\|_2, \quad (16)$$

$\delta_c(T_k)$ is a coefficient vector whose only nonzero entries in $\delta(T_k)$ that are associated with the c th class, and $\delta(T_k)$ is the coefficient vector for the testing sample which is sparsely reconstructed from X .

V. EXPERIMENTAL RESULTS

In this section, we perform gait-based human identity and gender recognition experiments to verify the efficacy of our proposed approach.

A. Gait-Based Human Identity Recognition

In our gait-based human identity recognition experiments, we randomly selected n gait sequences per subject from our dataset to construct the training set and use the remaining to form the testing set. We run the experiment 10 times by randomly splitting the training and test sets. The feature dimension l of our SRML method is empirically set as 50.

1) Comparison With Existing Metric Learning Algorithms: We compare our proposed SRML method with three existing metric learning algorithms: neighborhood component analysis (NCA) [32], large margin nearest neighbor (LMNN) [33], and information-theoretic metric learning (ITML) [34]. For these three methods, we empirically set the number of the nearest neighbors to 5. We apply principal component analysis (PCA) to reduce each C-AGI feature into 100 dimensions for all metric learning algorithms to improve the speed. For our SRML method, the feature dimension is determined by the parameter l . For the other three metric learning methods, the feature dimension is the selected as the the PCA subspace

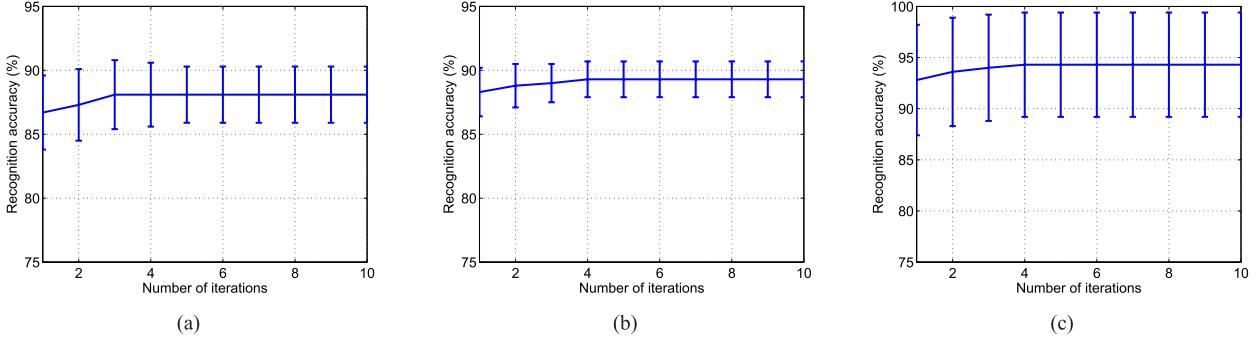


Fig. 6. Rank-1 recognition rate and the standard deviation of our approach versus different number of iterations when (a) $n = 1$, (b) $n = 2$, and (c) $n = 3$, respectively.

TABLE I

RANK-1 RECOGNITION RATE (%) OF DIFFERENT METRIC LEARNING METHODS ON THE ADSC-AWD DATASET. HERE n IS THE NUMBER OF GAIT SEQUENCES PER SUBJECT USED IN THE TRAINING SET

| Method | $n = 1$ | $n = 2$ | $n = 3$ |
|-----------|----------------|----------------|----------------|
| NCA [13] | 77.5 ± 2.3 | 82.5 ± 1.5 | 90.8 ± 5.2 |
| LMNN [46] | 69.2 ± 2.2 | 77.4 ± 2.3 | 87.0 ± 5.4 |
| ITML [8] | 79.5 ± 2.3 | 86.3 ± 1.7 | 87.5 ± 5.3 |
| SRML | 88.1 ± 2.2 | 89.2 ± 1.4 | 94.3 ± 5.1 |

TABLE II

RANK-1 RECOGNITION RATE (%) OF OUR APPROACH AND OTHER EXISTING GAIT RECOGNITION METHODS ON THE ADSC-AWD DATASET

| Method | $n = 1$ | $n = 2$ | $n = 3$ |
|----------------|----------------|----------------|----------------|
| Method in [16] | 77.0 ± 4.4 | 85.1 ± 3.5 | 87.5 ± 5.3 |
| Method in [49] | 79.5 ± 3.5 | 85.5 ± 3.3 | 90.5 ± 5.3 |
| Method in [18] | 86.3 ± 2.5 | 87.5 ± 1.6 | 93.3 ± 5.3 |
| Our approach | 88.1 ± 2.2 | 89.2 ± 1.4 | 94.3 ± 5.1 |

dimension. Table I shows the rank-1 recognition rate and the standard deviation of different methods on our database. We observe that our proposed SRML method consistently outperforms its competitors in terms of recognition rate. That is because our method applies the point-to-set distance to learn the distance metric while others use the point-to-point distance metric which may not effectively model large intra-class view variation, especially when the number of training samples is limited.

2) *Comparison With Existing Gait Recognition Methods:* We apply three existing gait recognition methods in [31], [47] and [30], implement them, and compare them with our approach. Table II shows the recognition accuracy. From this table, we see that our approach performs better than the other three methods. That is because the methods in [31] and [47] have not effectively exploited the characteristics of gait sequences captured from arbitrary walking directions for recognition. Our metric learning is also more powerful than the LDA method used in [30].

3) *Comparison With No Metric Learning:* In SRML, if M is not updated with metric learning and set as the Euclidean distance, the approach is equivalent to the SRC method

TABLE III

RANK-1 RECOGNITION RATE (%) OF OUR APPROACH WITH/WITHOUT METRIC LEARNING ON THE ADSC-AWD DATASET

| Method | $n = 1$ | $n = 2$ | $n = 3$ |
|----------------------|----------------|----------------|----------------|
| No metric learning | 80.3 ± 3.5 | 85.3 ± 2.7 | 88.6 ± 5.8 |
| With metric learning | 88.1 ± 2.2 | 89.2 ± 1.4 | 94.3 ± 5.1 |

TABLE IV

RANK-1 RECOGNITION RATE (%) OF DIFFERENT CLASSIFIERS ON THE ADSC-AWD DATASET

| Method | $n = 1$ | $n = 2$ | $n = 3$ |
|----------------|----------------|----------------|----------------|
| NN | 83.4 ± 3.2 | 86.6 ± 2.5 | 92.0 ± 5.4 |
| NFS | 86.8 ± 2.5 | 87.8 ± 1.6 | 93.3 ± 5.3 |
| Our classifier | 88.1 ± 2.2 | 89.2 ± 1.4 | 94.3 ± 5.1 |

in [43]. Table III shows the recognition accuracy. We see that our approach with metric learning performs better than that without metric learning because SRML can produce more discriminative features.

4) *Comparison With Different Classifiers:* We adopt the nearest neighbor (NN) classifier and the nearest feature space (NFS) classifier [48] to perform classification and compare it with our sparse reconstruction-based classification method. For the NFS classifier, the number of nearest neighbors was set as 3 in our experiments. For all these methods, the SRML metric learning is adopted for a fair comparison. Table IV shows the recognition accuracy. We observe that our classification approach achieves higher recognition rate than the other two, which is consistent with previous findings in [43] that SRC achieves better performance than the NN and NFS classifiers in face recognition.

5) *Influence of Number of Iterations:* Since SRML is an iterative algorithm, we evaluate its performance with different numbers of iterations. Fig. 6 shows the recognition accuracy versus number of iterations. We see that our proposed SRML method achieves stable performance in just a few iterations.

6) *Successful and Fail Recognition Examples:* Fig. 7 shows two examples of the recognition results of our approach, one successful and the other one not. In Fig. 7(a), two gait sequences are from the same person, one is from the training set, and the other one is from the testing set. Our approach

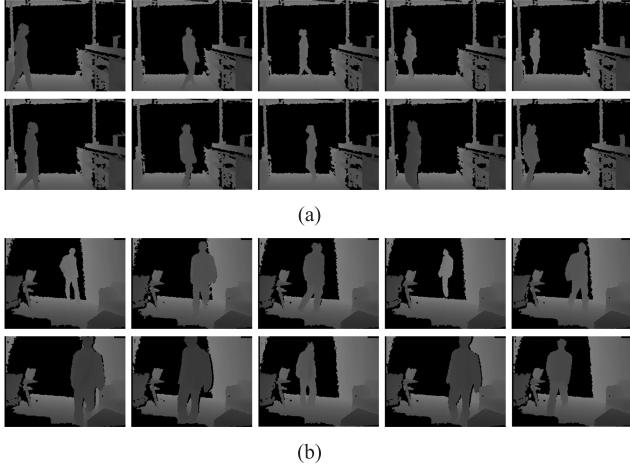


Fig. 7. Two recognition results of our approach, where image frames in each row are from the same gait sequence. (a) One successful example, and (b) one unsuccessful example.

performs the recognition successfully. In Fig. 7(b), there are also two gait sequences from the same person. However, our approach cannot correctly recognize the testing sequence. That is because in Fig. 7(a), the person walks in the scene with similar walking direction changes and similar speed. In this case, our approach yields very similar clusters for both the training and testing sequences, and it correctly recognizes the testing sample. In Fig. 7(b), even though the subjects in the two sequences are the same person, there is a large difference between the walking styles, namely walking speeds and directions. In this case, the clusters obtained by our approach are different for two sequences and our approach fails to recognize the testing sample.

B. Gait-Based Gender Recognition

Besides gait recognition, we also investigate gait-based gender recognition from arbitrary walking directions. In our dataset, there are 20 subjects (13 males and 7 females), and each subject is captured in 4 video clips. We adopt the leave-one-person-out strategy to conduct gender recognition experiments. Specifically, we take four gait videos for each person as the testing set and the remaining as the training set. We repeat this 20 times and record the average classification rate as the final classification accuracy.

1) Comparison With Existing Metric Learning Algorithms: We compare our SRML method with NCA [32], LMNN [33], ITML [34] and our previous method in [1]. For these four methods, we empirically set the number of the nearest neighbors to 5. We apply PCA to reduce each C-AGI feature into 100 dimensions for all metric learning algorithms to improve the speed. Table V shows the classification rate of different methods. We observe that our proposed SRML method consistently outperforms the other compared methods in terms of recognition accuracy.

2) Comparison With Existing Gait-Based Gender Recognition Methods: We apply the gait-based gender recognition methods of [11] and [12], implement them, and compare them with our approach. Table VI shows the recognition accuracy.

TABLE V
CORRECT CLASSIFICATION RATE (%) OF DIFFERENT METRIC LEARNING METHODS FOR GAIT-BASED GENDER RECOGNITION ON THE ADSC-AWD DATASET

| Method | Correct classification rate |
|----------------|-----------------------------|
| NCA | 89.3 |
| LMNN | 88.3 |
| ITML | 86.7 |
| Method in [32] | 91.7 |
| SRML | 93.7 |

TABLE VI
CORRECT CLASSIFICATION RATE (%) OF DIFFERENT GAIT-BASED GENDER RECOGNITION METHODS ON THE ADSC-AWD DATASET

| Method | Correct classification rate |
|----------------|-----------------------------|
| Method in [28] | 87.7 |
| Method in [53] | 89.3 |
| Our approach | 93.7 |

TABLE VII
CORRECT CLASSIFICATION RATE (%) OF OUR APPROACH WITH/WITHOUT METRIC LEARNING FOR GAIT-BASED GENDER RECOGNITION ON THE ADSC-AWD DATASET

| Method | Correct classification rate |
|----------------------|-----------------------------|
| No metric learning | 87.7 |
| With metric learning | 93.7 |

TABLE VIII
CORRECT CLASSIFICATION RATE (%) OF DIFFERENT CLASSIFIERS FOR GAIT-BASED GENDER RECOGNITION ON THE ADSC-AWD DATASET

| Method | Correct classification rate |
|----------------|-----------------------------|
| NN | 84.3 |
| NFS | 90.3 |
| Our classifier | 93.7 |

Our approach outperforms the other two. That is because these methods have not effectively exploited the characteristics of gait sequences captured from arbitrary walking directions for gender recognition.

3) Comparison With No Metric Learning: Table VII shows that our approach with metric learning outperforms that without metric learning because SRML can produce more discriminative features.

4) Comparison With Different Classifiers: Similar to human identity recognition, we also adopt the nearest neighbor (NN) classifier and the nearest feature space (NFS) classifier [48] to perform gender recognition. For the NFS classifier, the number of nearest neighbors was set to 3 in our experiments. For all these methods, the SRML metric learning was adopted to yield a fair comparison. Table VIII shows the recognition accuracy. We observe that our approach achieves higher recognition rate than the other two.

5) Influence of the Number of Iterations: Fig. 8 shows the recognition accuracy of SRML versus number of iterations. We see that our proposed SRML method achieves stable performance in just a few iterations.

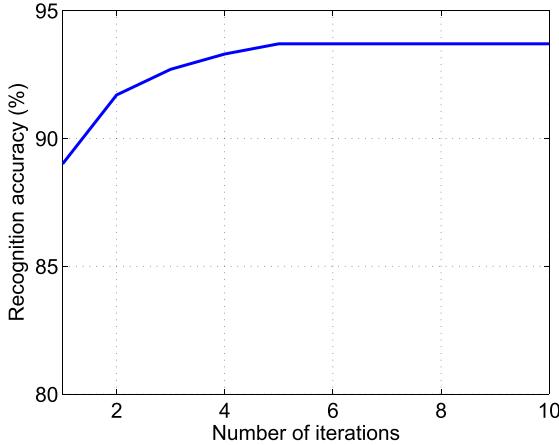


Fig. 8. Gender recognition rate of our approach versus different number of iterations.

TABLE IX

CPU TIME (IN SECONDS) USED BY DIFFERENT METRIC LEARNING METHODS ON OUR ADSC-AWD DATABASE

| Method | Training | Testing |
|--------|----------|---------|
| NCA | 15.32 | 2.54 |
| LMNN | 200.46 | 2.54 |
| ITML | 88.65 | 2.54 |
| SRML | 4502.65 | 2.54 |

C. Computational Time

We analyze the computational complexity of the proposed SRML method. There are three steps in each iteration: compute the coefficient matrices A and B , calculate two matrices Z_1 and Z_2 , and solve an eigenvalue equation. The computational complexity of computing these three steps in one iteration is $O(CLd^2(K-1))$, $O(CLKd^2)$, and $O(d^3)$. Hence, the computational complexity of our SRML is $O(FT)$, where $F = \max\{CLK, d\}$.

We also compare the computational time of our SRML with those of the NCA [32], LMNN [33], and ITML [34] metric learning methods. Our hardware contains a 2.4-GHz CPU and a 6GB RAM. Table IX shows the computational time of different metric learning methods, using the Matlab software and our dataset for gender recognition experiments.

From Table IX, we see that the computational time of our SRML is higher than that of other metric learning methods for training. However, training is usually performed offline and only the recognition is performed online. As shown in Table IX, the recognition time of the proposed method is the same to other metric learning methods when the same feature dimension is used. Hence, the computational complexity will not affect the practical applications of our proposed SRML.

D. Experiments on the USF and CASIA-B Datasets

Since SRML is a discriminative learning method in general, it should be also suitable to gait-based identity and gender recognition under controlled environments. Hence, we apply our proposed SRML method to conduct gait-based identity and gender recognition experiments on the widely used USF

HumanID [18] and CASIA-B [19] gait databases to further demonstrate its effectiveness.

1) *Gait-Based Human Identity Recognition*: We conduct gait recognition experiments on the USF HumanID and CASIA-B datasets. For the USF database, we follow the experimental setting in [18], where the gait energy image (GEI) feature extraction method [31] was used to represent each gait sequence. Indeed, previous studies [31], [47], [49] have demonstrated experimentally that GEI is more effective than the other binary silhouette-based features [5], [22], [41] for gait-based identity recognition. We compare our SRML method with four state-of-the-art gait recognition methods including baseline [18], 2DLDA [49], TLPP [50], CGI [10], RTRDA [51], Gabor-PDF-LGSR [52], and RSM [53]. Table X shows the rank-1 recognition rates of different gait recognition methods on the USF database. We observe that our proposed SRML achieves comparable recognition rate with most existing state-of-the-art gait recognition methods, which further demonstrates its effectiveness for gait recognition.

For the CASIA-B database, we follow the experimental settings in [24], where 6 normal walking sequences of each subject collected from the 54° , 72° , 90° , 108° and 126° are selected for our cross-view gait recognition experiments. We compare our approach with eight existing cross-view gait recognition methods including CSC+PSA [5], view rectification [26], PSC+PSA [23], FT-SVD [25], GEI-SVD [54], GEI+SVR [29], GEI+CCA [27], and TILT+PMS [24]. For a fair comparison, the same number of subjects (i.e. 65 subjects) are used by all methods in this experiment. Table XI shows the rank-1 recognition rates of different gait recognition methods when gait sequences collected from the 90° view are selected as the probe set on the CASIA-B database. We see that our proposed SRML achieves comparable recognition rate with state-of-the-art cross-view gait recognition methods.

2) *Gait-Based Gender Recognition*: For the USF database, we follow the experimental settings in [11]. Specifically, we use the gallery set for training and the remaining probe sets for testing. Table XII shows the recognition rate of our method and that in [11] on the USF database. We see that our approach achieves better recognition rate than that the method in [11].

For the CASIA-B database, we follow the experimental settings in [13]. Since there are only 31 female subjects in this dataset, 31 male subjects were randomly selected to construct a balance dataset for gender recognition. Similar to [13], the side-view gait sequences with normal clothes and without any bag, were used in the experiments. We employ a 31-fold cross-validation strategy to compute the recognition rate. Specifically, we constructed 31 non-overlapping subsets, and each subset consists of gait sequences of one male and one female. Each time, one subset is used for testing and the remaining 30 subsets for training. Table XIII shows the recognition rate of our approach and several existing gait-based gender recognition methods. We see that our approach achieves comparable recognition rate with the other methods.

E. Discussion

We make the following three observations from the experimental results listed in Tables I-XIII:

TABLE X
RANK-1 RECOGNITION RATE (%) OF DIFFERENT GAIT RECOGNITION METHODS ON THE USF DATASET

| Method | Baseline | 2DLDA | TLPP | CGI | RTRDA | Gabor-PDF-LGSR | RSM | SRML |
|---------|----------|-------|------|------|-------|----------------|------|------|
| Probe A | 87 | 89 | 87 | 91 | 95 | 95 | 92 | 93 |
| Probe B | 93 | 93 | 93 | 93 | 91 | 93 | 93 | 94 |
| Probe C | 72 | 80 | 72 | 78 | 85 | 89 | 84 | 85 |
| Probe D | 42 | 28 | 25 | 51 | 59 | 62 | 39 | 52 |
| Probe E | 43 | 33 | 35 | 53 | 62 | 62 | 40 | 52 |
| Probe F | 26 | 17 | 17 | 35 | 36 | 39 | 22 | 37 |
| Probe G | 28 | 19 | 18 | 38 | 36 | 38 | 26 | 40 |
| Probe H | 62 | 74 | 62 | 84 | 84 | 94 | 90 | 86 |
| Probe I | 58 | 71 | 62 | 78 | 76 | 91 | 82 | 85 |
| Probe J | 50 | 49 | 43 | 64 | 73 | 78 | 72 | 68 |
| Probe K | 6 | 16 | 12 | 3 | 3 | 21 | 18 | 18 |
| Probe L | 6 | 16 | 15 | 9 | 21 | 21 | 14 | 15 |
| Mean | 51.6 | 51.0 | 47.0 | 61.7 | 65.4 | 70.1 | 65.6 | 66.5 |

TABLE XI
RANK-1 RECOGNITION RATE (%) OF DIFFERENT CROSS-VIEW GAIT RECOGNITION METHODS ON THE CASIA-B DATASET

| Method | 54° | 72° | 90° | 108° | 126° |
|-------------------------|-----|-----|------|------|------|
| CSC+PSA [44] | 27 | 71 | 85 | 67 | 21 |
| view rectification [12] | 62 | 64 | 87 | 64 | 70 |
| PSC+PSA [24] | 52 | 93 | 97 | 77 | 28 |
| FT-SVD [35] | 27 | 36 | 87 | 58 | 28 |
| GEI-SVD [22] | 52 | 75 | 93 | 79 | 45 |
| GEI+SVR [25] | 63 | 92 | 93 | 91 | 65 |
| GEI+CCA [1] | 60 | 96 | N.A. | 94 | 95 |
| TILT+PMS [26] | 70 | 97 | 98 | 93 | 55 |
| Our approach | 68 | 94 | 99 | 96 | 70 |

TABLE XII
CORRECT CLASSIFICATION RATE (%) OF DIFFERENT GAIT-BASED GENDER RECOGNITION METHODS ON THE USF DATASET

| Method | Method in [28] | SRML |
|---------|----------------|------|
| Probe A | 98 | 98 |
| Probe B | 98 | 98 |
| Probe C | 96 | 97 |
| Probe D | 82 | 84 |
| Probe E | 84 | 86 |
| Probe F | 76 | 78 |
| Probe G | 74 | 76 |
| Probe H | 92 | 93 |
| Probe I | 84 | 86 |
| Probe J | 93 | 94 |
| Probe K | 94 | 94 |
| Probe L | 82 | 84 |

TABLE XIII
CORRECT CLASSIFICATION RATE (%) OF DIFFERENT GAIT-BASED GENDER RECOGNITION METHODS ON THE CASIA-B DATASET

| Method | Correct classification rate |
|----------------|-----------------------------|
| Method in [28] | 93.3 |
| Method in [52] | 96.0 |
| Method in [17] | 98.4 |
| Our approach | 98.0 |

- While the problem of gait-based human identity and gender recognition from arbitrary walking directions is challenging, our approach achieves 87.6% rank-1 identity recognition rate and 93.1% gender recognition rate on our ADSC-AWD dataset containing 20 subjects, respectively.

- Our proposed SRML method consistently outperforms the other compared metric learning methods on our human identity and gender recognition tasks. That is because our method applies the point-to-set distance to learn the distance metric while others use the point-to-point distance metric which may not effectively model the large variations in intra-class pose and view, especially when the number of training samples is limited.
- Our proposed SRML method achieves comparable recognition rate with existing state-of-the-art gait-based identity and gender recognition methods on the widely used USF and CASIA-B gait databases. This demonstrates its effectiveness in the gait-based identity and gender recognition under controlled environments.

VI. CONCLUSION

We have proposed a sparse reconstruction based metric learning approach for gait-based human identity and gender recognition from arbitrary walking directions. Experimental results on our dataset have shown the efficacy of our approach. We have applied our proposed SRML method on the existing USF and CASIA-B gait datasets and achieved comparable recognition rate with most existing state-of-the-art gait-based human identity and gender recognition methods.

Our future work will explore several interesting directions:

- Our approach will fail if the walking style in the testing sequence is significantly different from that in the training sequences. That is because there are large differences for the C-AGI features of the same subject in such scenarios. To address this, we will consider extracting more discriminative modal-based features to improve the performance of our proposed approach.
- Our ADSC-AMD gait dataset was collected in indoor environments, however, our approach can also be used for outdoor human gait analysis because only human silhouettes are required for gait feature representation in our approach.
- Our proposed SRML method could be applied to other computer vision applications such as face and object recognition.

REFERENCES

- [1] J. Lu, G. Wang, and T. S. Huang, "Gait-based gender classification in unconstrained environments," in *Proc. 21st IEEE Int. Conf. Pattern Recognit.*, Nov. 2012, pp. 3284–3287.
- [2] S. A. Niyogi and E. H. Adelson, "Analyzing and recognizing walking figures in XYT," in *Proc. IEEE Int. Conf. CVPR*, Jun. 1994, pp. 469–474.
- [3] A. F. Bobick and A. Y. Johnson, "Gait recognition using static, activity-specific parameters," in *Proc. IEEE Int. Conf. CVPR*, vol. 1, Dec. 2001, pp. 423–430.
- [4] L. Lee and W. E. L. Grimson, "Gait analysis for recognition and classification," in *Proc. 5th IEEE Int. Conf. Autom. Face Gesture Recognit.*, May 2002, pp. 148–155.
- [5] L. Wang, T. Tan, H. Ning, and W. Hu, "Silhouette analysis-based gait recognition for human identification," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 25, no. 12, pp. 1505–1518, Dec. 2003.
- [6] L. Wang, T. Tan, W. Hu, and H. Ning, "Automatic gait recognition based on statistical shape analysis," *IEEE Trans. Image Process.*, vol. 12, no. 9, pp. 1120–1131, Sep. 2003.
- [7] A. Kale, A. Sundaresan, A. Rajagopalan, N. Cuntoor, A. Roy-Chowdhury, V. Kruger, et al., "Identification of humans using gait," *IEEE Trans. Image Process.*, vol. 13, no. 9, pp. 1163–1173, Sep. 2004.
- [8] N. Boulgouris, D. Hatzinakos, and K. Plataniotis, "Gait recognition: A challenging signal processing technology for biometric identification," *IEEE Signal Process. Mag.*, vol. 22, no. 6, pp. 78–90, Nov. 2005.
- [9] N. Boulgouris and Z. Chi, "Gait recognition using radon transform and linear discriminant analysis," *IEEE Trans. Image Process.*, vol. 16, no. 3, pp. 731–740, Mar. 2007.
- [10] C. Wang, J. Zhang, J. Pu, X. Yuan, and L. Wang, "Chrono-gait image: A novel temporal template for gait recognition," in *Proc. Eur. Conf. Comput. Vis.*, 2010, pp. 257–270.
- [11] X. Li, S. Maybank, S. Yan, D. Tao, and D. Xu, "Gait components and their application to gender recognition," *IEEE Trans. Syst., Man, Cybern. C. Appl. Rev.*, vol. 38, no. 2, pp. 145–155, Mar. 2008.
- [12] S. Yu, T. Tan, K. Huang, K. Jia, and X. Wu, "A study on gait-based gender classification," *IEEE Trans. Image Process.*, vol. 18, no. 8, pp. 1905–1910, Aug. 2009.
- [13] M. Hu, Y. Wang, Z. Zhang, and D. Zhang, "Gait-based gender classification using mixed conditional random field," *IEEE Trans. Syst., Man, Cybern. B, Cybern.*, vol. 41, no. 5, pp. 1429–1439, Oct. 2011.
- [14] J. Lu and Y.-P. Tan, "Gait-based human age estimation," *IEEE Trans. Inf. Forensics Security*, vol. 5, no. 4, pp. 761–770, Dec. 2010.
- [15] Y. Makihara, M. Okumura, H. Iwama, and Y. Yagi, "Gait-based age estimation using a whole-generation gait database," in *Proc. IEEE Int. Joint Conf. Biometr.*, Oct. 2011, pp. 1–6.
- [16] J. Lu and Y.-P. Tan, "Ordinary preserving manifold analysis for human age and head pose estimation," *IEEE Trans. Human-Mach. Syst.*, vol. 43, no. 2, pp. 249–258, Mar. 2013.
- [17] D. S. Matovski, M. S. Nixon, S. Mahmoodi, and J. N. Carter, "The effect of time on gait recognition performance," *IEEE Trans. Inf. Forensics Security*, vol. 7, no. 2, pp. 543–552, Apr. 2012.
- [18] S. Sarkar, P. Phillips, Z. Liu, I. Vega, P. Grother, and K. Bowyer, "The humanID gait challenge problem: Data sets, performance, and analysis," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 27, no. 2, pp. 162–177, Feb. 2005.
- [19] S. Yu, D. Tan, and T. Tan, "Modelling the effect of view angle variation on appearance-based gait recognition," in *Proc. Asian Conf. Comput. Vis.*, 2006, pp. 807–816.
- [20] D. Tan, K. Huang, S. Yu, and T. Tan, "Efficient night gait recognition based on template matching," in *Proc. 18th IEEE Int. Conf. Pattern Recognit.*, Aug. 2006, pp. 1000–1003.
- [21] H. Iwama, M. Okumura, Y. Makihara, and Y. Yagi, "The OU-ISIR gait database comprising the large population dataset and performance evaluation of gait recognition," *IEEE Trans. Inf. Forensics Security*, vol. 7, no. 5, pp. 1511–1521, Oct. 2012.
- [22] J. Lu and Y.-P. Tan, "Uncorrelated discriminant simplex analysis for view-invariant gait signal computing," *Pattern Recognit. Lett.*, vol. 31, no. 5, pp. 382–393, 2010.
- [23] W. Kusakunniran, Q. Wu, J. Zhang, and H. Li, "Pairwise shape configuration-based PSA for gait recognition under small viewing angle change," in *Proc. 8th IEEE Int. Conf. Adv. Video Signal-Based Survell.*, Aug./Sep. 2011, pp. 17–22.
- [24] W. Kusakunniran, Q. Wu, J. Zhang, Y. Ma, and H. Li, "A new view-invariant feature for cross-view gait recognition," *IEEE Trans. Inf. Forensics Security*, vol. 8, no. 5, pp. 1642–1653, Oct. 2013.
- [25] Y. Makihara, R. Sagawa, Y. Mukaigawa, T. Echigo, and Y. Yagi, "Gait recognition using a view transformation model in the frequency domain," in *Proc. Eur. Conf. Comput. Vis.*, 2006, pp. 151–163.
- [26] M. Goffredo, I. Bouchrika, J. Carter, and M. Nixon, "Self-calibrating view-invariant gait biometrics," *IEEE Trans. Syst., Man, Cybern. B, Cybern.*, vol. 40, no. 4, pp. 997–1008, Aug. 2010.
- [27] K. Bashir, T. Xiang, and S. Gong, "Cross view gait recognition using correlation strength," in *Proc. Brit. Mach. Vis. Conf.*, 2010, pp. 1–11.
- [28] W. Kusakunniran, Q. Wu, J. Zhang, and H. Li, "Support vector regression for multi-view gait recognition based on local motion feature selection," in *Proc. IEEE Int. Conf. Comput. Vis. Pattern Recognit.*, Jun. 2010, pp. 974–981.
- [29] W. Kusakunniran, Q. Wu, J. Zhang, and H. Li, "Gait recognition under various viewing angles based on correlated motion regression," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 22, no. 6, pp. 966–980, Jun. 2012.
- [30] X. Huang and N. V. Boulgouris, "Gait recognition with shifted energy image and structural feature extraction," *IEEE Trans. Image Process.*, vol. 21, no. 4, pp. 2256–2268, Apr. 2012.
- [31] J. Han and B. Bhanu, "Individual recognition using gait energy image," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 28, no. 2, pp. 316–322, Feb. 2006.
- [32] J. Goldberger, S. Roweis, G. Hinton, and R. Salakhutdinov, "Neighborhood component analysis," in *Advances in Neural Information Processing Systems*. Cambridge, MA, USA: MIT Press, 2004, pp. 2539–2544.
- [33] K. Weinberger, J. Blitzer, and L. Saul, "Distance metric learning for large margin nearest neighbor classification," in *Advances in Neural Information Processing Systems*. Cambridge, MA, USA: MIT Press, 2005, pp. 1473–1480.
- [34] J. Davis, B. Kulis, P. Jain, S. Sra, and I. Dhillon, "Information-theoretic metric learning," in *Proc. 24th Int. Conf. Mach. Learn.*, 2007, pp. 209–216.
- [35] M. Guillaumin, J. Verbeek, and C. Schmid, "Is that you? Metric learning approaches for face identification," in *Proc. IEEE 12th Int. Conf. Comput. Vis.*, Sep./Oct. 2009, pp. 498–505.
- [36] D. Tran and A. Sorokin, "Human activity recognition with metric learning," in *Proc. Eur. Conf. Comput. Vis.*, 2008, pp. 548–561.
- [37] W.-S. Zheng, S. Gong, and T. Xiang, "Person re-identification by probabilistic relative distance comparison," in *Proc. IEEE Int. Conf. Comput. Vis. Pattern Recognit.*, Jun. 2011, pp. 649–656.
- [38] L. Yang, R. Jin, L. Mummert, R. Sukthankar, A. Goode, B. Zheng, et al., "A boosting framework for visibility-preserving distance metric learning and its application to medical image retrieval," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 32, no. 1, pp. 30–44, Jan. 2010.
- [39] A. Frome, Y. Singer, F. Sha, and J. Malik, "Learning globally-consistent local distance functions for shape-based image retrieval and classification," in *Proc. IEEE 11th Int. Conf. Comput. Vis.*, Oct. 2007, pp. 1–8.
- [40] X. Wang, G. Hua, and T. X. Han, "Discriminative tracking by metric learning," in *Proc. Eur. Conf. Comput. Vis.*, 2010, pp. 200–214.
- [41] J. Lu and E. Zhang, "Gait recognition for human identification based on ICA and fuzzy SVM through multiple views fusion," *Pattern Recognit. Lett.*, vol. 28, no. 16, pp. 2401–2411, 2007.
- [42] B. J. Frey and D. Dueck, "Clustering by passing messages between data points," *Science*, vol. 315, no. 5814, pp. 972–976, 2007.
- [43] J. Wright, A. Yang, A. Ganesh, S. Sastry, and Y. Ma, "Robust face recognition via sparse representation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 31, no. 2, pp. 210–227, Feb. 2009.
- [44] W. Bian and D. Tao, "Max-min distance analysis by using sequential SDP relaxation for dimension reduction," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 33, no. 5, pp. 1037–1050, May 2011.
- [45] S. Yan, D. Xu, Q. Yang, L. Zhang, X. Tang, and H. Zhang, "Multilinear discriminant analysis for face recognition," *IEEE Trans. Image Process.*, vol. 16, no. 1, pp. 212–220, Jan. 2007.
- [46] D. Donoho and Y. Tsai, "Fast solution of ℓ_1 -norm minimization problems when the solution may be sparse," *IEEE Trans. Inf. Theory*, vol. 54, no. 11, pp. 4789–4812, Nov. 2008.
- [47] D. Xu, S. Yan, D. Tao, S. Lin, and H. Zhang, "Marginal Fisher analysis and its variants for human gait recognition and content-based image retrieval," *IEEE Trans. Image Process.*, vol. 16, no. 11, pp. 2811–2821, Nov. 2007.
- [48] J. Chien and C. Wu, "Discriminant waveletfaces and nearest feature classifiers for face recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 24, no. 12, pp. 1644–1649, Dec. 2002.
- [49] D. Tao, X. Li, X. Wu, and S. Maybank, "General tensor discriminant analysis and gabor features for gait recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 29, no. 10, pp. 1700–1715, Oct. 2007.

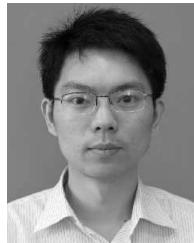
- [50] C. Chen, J. Zhang, and R. Fleischer, "Distance approximating dimension reduction of Riemannian manifolds," *IEEE Trans. Syst., Man, Cybern. B, Cybern.*, vol. 40, no. 1, pp. 208–217, Feb. 2010.
- [51] Y. Huang, D. Xu, and F. Nie, "Regularized trace ratio discriminant analysis with patch distribution feature for human gait recognition," in *Proc. 17th IEEE Int. Conf. Image Process.*, Sep. 2010, pp. 2449–2452.
- [52] D. Xu, Y. Huang, Z. Zeng, and X. Xu, "Human gait recognition using patch distribution feature and locality-constrained group sparse representation," *IEEE Trans. Image Process.*, vol. 21, no. 1, pp. 316–326, Jan. 2012.
- [53] Y. Guan, C.-T. Li, and Y. Hu, "Random subspace method for gait recognition," in *Proc. IEEE Int. Conf. Multimedia Expo Workshops*, Jul. 2012, pp. 284–289.
- [54] W. Kusakunniran, Q. Wu, H. Li, and J. Zhang, "Multiple views gait recognition using view transformation model based on optimized gait energy image," in *Proc. IEEE 12th Int. Conf. Comput. Vis. Workshops*, Sep./Oct. 2009, pp. 1058–1064.



Jiwen Lu (S'10–M'11) received the B.Eng. degree in mechanical engineering and the M.Eng. degree in electrical engineering from the Xi'an University of Technology, Xi'an, China, and the Ph.D. degree in electrical engineering from Nanyang Technological University, Singapore.

He is currently a Research Scientist with the Advanced Digital Sciences Center, Singapore. His research interests include computer vision, pattern recognition, machine learning, and biometrics. He has authored or co-authored more than 70 scientific papers in peer-reviewed journals and conferences including some top venues such as the IEEE TPAMI, TIP, TCSVT, TSMC-B, TIFS, ICCV, CVPR, and ACM MM.

Dr. Lu was a recipient of the First-Prize National Scholarship and the National Outstanding Student Award from the Ministry of Education of China in 2002 and 2003, respectively, and the Best Student Paper Award from PREMIA of Singapore in 2012.



Gang Wang (M'11) received the B.S. degree in electrical engineering from the Harbin Institute of Technology in 2005 and the Ph.D. degree from the Department of Electrical and Computer Engineering, University of Illinois at Urbana-Champaign, in 2010.

He is currently an Assistant Professor with the School of Electrical and Electronic Engineering, Nanyang Technological University, and a Research Scientist at the Advanced Digital Sciences Center, Singapore. His current research interests include computer vision and machine learning. He is focusing on object recognition, scene analysis, large-scale machine learning, and deep learning.



Pierre Moulin (F'03) received the Doctoral degree from Washington University, St. Louis, MO, USA, in 1990, after which he joined Bell Communications Research, Morristown, NJ, USA, as a Research Scientist. In 1996, he joined the University of Illinois at Urbana-Champaign, where he is currently Professor with the Department of Electrical and Computer Engineering, Research Professor at the Beckman Institute and the Coordinated Science Laboratory, and an Affiliate Professor with the Department of Statistics. His fields of professional interest include

image and video processing, compression, statistical signal processing and modeling, media security, decision theory, and information theory.

He has served on the editorial boards of the IEEE TRANSACTIONS ON INFORMATION THEORY, the IEEE TRANSACTIONS ON IMAGE PROCESSING, and the PROCEEDINGS OF IEEE. He currently serves on the editorial board of *Foundations and Trends in Signal Processing*. He was co-founding Editor-in-Chief of the IEEE TRANSACTIONS ON INFORMATION FORENSICS AND SECURITY from 2005 to 2008, a member of the IEEE Signal Processing Society Board of Governors from 2005 to 2007, and has served the IEEE in various other capacities.

Dr. Moulin received the 1997 Career Award from the National Science Foundation and the IEEE Signal Processing Society 1997 Senior Best Paper Award. He received the IEEE Signal Processing Society 2002 Young Author Best Paper Award. In 2003, he became a Beckman Associate of the UIUC's Center for Advanced Study. From 2007 to 2009, he was a Sony Faculty Scholar at UIUC. He was a plenary speaker for ICASSP 2006, ICIP 2011, and several other conferences. He is a Distinguished Lecturer of the IEEE Signal Processing Society from 2012 to 2013.