On Reducing the Effect of Covariate Factors in Gait Recognition: a Classifier Ensemble Method

Yu Guan, Chang-Tsun Li, *Senior Member, IEEE*, and Fabio Roli, *Fellow, IEEE*

***Abstract*—Robust human gait recognition is challenging because of the presence of covariate factors such as carrying condition, clothing, walking surface, etc. In this paper, we model the effect of covariates as an unknown partial feature corruption problem. Since the locations of corruptions may differ for different query gaits, relevant features may become irrelevant when walking condition changes. In this case, it is difficult to train one fixed classifier that is robust to a large number of different covariates. To tackle this problem, we propose a classifier ensemble method based on the Random Subspace Method (RSM) and Majority Voting (MV). Its theoretical basis suggests it is insensitive to locations of corrupted features, and thus can generalize well to a large number of covariates. We also extend this method by proposing two strategies, i.e., Local Enhancing (LE) and Hybrid Decision-level Fusion (HDF) to suppress the ratio of false votes to true votes (before MV). The performance of our approach is competitive against the most challenging covariates like clothing, walking surface, and elapsed time. We evaluate our method on the USF dataset and OU-ISIR-B dataset, and it has much higher performance than other state-of-the-art algorithms.**

***Index Terms*—Classifier ensemble, random subspace method, local enhancing, hybrid decision-level fusion, gait recognition, covariate factors, biometrics**

I. Introduction

Compared with other biometric traits like fingerprint or iris, the most significant advantage of gait is that it can be used for remote human identification without subject cooperation. Gait analysis has contributed to convictions in criminal cases in some countries like Denmark [1] and UK [2]. However, the performance of automatic gait recognition can be affected by covariate factors such as carrying condition, camera viewpoint, walking surface, clothing, elapsed time, etc. Designing a robust system to address these problems is an acute challenge. Existing gait recognition methods can be roughly divided into two categories: model-based and appearance-based approaches. Model-based methods (e.g., [3]) employ the parameters of the body structure, while appearance-based approaches extract features directly from gait sequences regardless of the underlying structure. This work falls in the category of appearance-based methods, which can also work well on low-quality gait videos, when the parameters of the body structure are difficult to estimate precisely.

The average silhouette over one gait cycle, known as Gait Energy Image (GEI), is a popular appearance-based representation [4]. The averaging operation encodes the information of the binary frames into a single grayscale image, which makes GEI less sensitive to segmentation errors [4]. Several GEI samples from the USF dataset [5] are shown in Fig. 1. Recently, Iwama et al. evaluated the effectiveness of several feature templates on a gait dataset consisting of more than 3000 subjects. They found that good performance can be achieved by directly matching the GEI templates, when there are no covariates [6]. However, it is error-prone when covariates exist. Therefore, many researchers have been

Y. Guan, and C.-T. Li are with the Department of Computer Science, University of Warwick, Coventry, CV4 7AL, U.K.
E-mail: yg.yuguan@gmail.com, c-t.li@warwick.ac.uk
F. Roli is with the Department of Electrical and Electronic Engineering, University of Cagliari, 09123, Cagliari, Italy
E-mail: roli@diee.unica.it

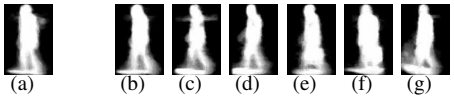


Fig. 1. GEIs of the same subject from the USF dataset [5]. The leftmost (a) is the gallery GEI in normal condition, while the rest (b)-(g) are probe GEIs with covariates (b) viewpoint, (c) walking surface, (d) viewpoint and walking surface, (e) carrying condition, (f) carrying condition and viewpoint, (g) elapsed time, shoe type, clothing, and walking surface

formulating various feature descriptors to capture discriminant information from GEIs to deal with different covariates. In [4], Han and Bhanu utilized Principle Component Analysis (PCA) and Linear Discriminant Analysis (LDA) to extract features from the concatenated GEIs. To extract gait descriptors, Li et al. proposed a Discriminant Locally Linear Embedding (DLLE) framework, which can preserve the local manifold structure [7]. By using two subspace learning methods, Coupled Subspaces Analysis (CSA) and Discriminant Analysis with Tensor Representation (DATER), Xu et al. extracted features directly from the GEIs [8]. They demonstrated that the matrix representation can yield much higher performance than the vector representation reported in [4]. In [9], after convolving a number of Gabor functions with the GEI representation, Tao et al. used the Gabor-filtered GEI as a new gait feature template. They also proposed the General Tensor Discriminant Analysis (GTDA) for feature extraction on the high-dimensional Gabor features [9]. To preserve the local manifold structure of the high-dimensional Gabor features, Chen et al. proposed a Tensor-based Riemannian Manifold distance-approximating Projection (TRIMAP) framework [10]. Since spatial misalignment may degrade the performance, based on Gabor representation, Image-to-Class distance was utilized in [11] to allow feature matching to be carried out within a spatial neighborhood. By using the techniques of Universal Background Model (UBM) learning and Maximum A Posteriori (MAP) adaptation, Xu et al. proposed the Gabor-based Patch Distribution Feature (Gabor-PDF) in [12], and the classification is performed based on Locality-constrained Group Sparse Representation (LGSR). Compared with GEI features (e.g., [4], [8]), Gabor features (e.g., [9], [11], [12]) tend to be more discriminant and can yield higher accuracies. However, due to the high dimension of Gabor features, these methods normally incur high computational costs.

Through the "cutting and fitting" scheme, synthetic GEIs can be generated to simulate the walking surface effect [4]. In [13], the population hidden Markov model was used for Dynamic-Normalized Gait Recognition (DNGR). Both methods can yield encouraging performance against the walking surface covariate. In [14], Wang et al. proposed Chrono-Gait Image (CGI), with the temporal information of the gait sequence encoded. Although CGI has similar performance to GEI in most cases [6] [14], it outperforms GEI in tackling the carrying condition covariate [14]. In [15], through fusing gait features from different cameras, it was found that short elapsed time does not affect the performance significantly in a controlled environment. Clothing was instead deemed as the most challenging covariate [15]. Based on a gait dataset with 32 different clothes combinations, Hossain et al. proposed an adaptive scheme for weighting different body parts to reduce the effect of clothing [16]. Yet it requires an additional training set that covers "all" possible clothes types, which is less practical in real-world applications.

Most of the previous works have satisfactory performance against some covariates like carrying condition, shoe type, (small changes in) viewpoint, etc. However, so far it still remains an

open question to tackle covariates like walking surface, clothing, and elapsed time in *less controlled environments*. As such, our aim is not only to build a general framework that can generalize to a large number of covariates in unseen walking conditions, but also to address these open issues. Based on GEIs, we model the effect caused by various covariates as an unknown partial feature corruption problem and propose a weak classifier ensemble method to reduce such effect. Each weak classifier is generated by random sampling in the full feature space, so they may generalize to the unselected features in different directions [17]. This concept, named Random Subspace Method (RSM), was initially proposed by Ho [17] for constructing decision forests. It was successfully applied to face recognition (e.g., [18]). In our previous works [19], [20], we employed this concept in gait recognition. Empirical results suggested RSM-based classifier ensemble can generalize well to several covariates. However, more theoretical findings and experimental results are needed to support this gait recognition method, especially on addressing the aforementioned open issues. Our contributions in this work can be summarized as follows:

1) **Modelling of Gait Recognition Challenges:** Based on GEIs, we model the effect of different covariates as a partial feature corruption problem with unknown locations.
2) **Classifier Ensemble Solution and its Extensions:** Through ideal cases analysis, we provide the theoretical basis of the proposed classifier ensemble solution. To tackle the hard problems in real cases, we further propose two strategies, namely, local enhancing and hybrid decision-level fusion.
3) **Performance and Open Issues Solving:** On the USF dataset, our method has more than 80% accuracy, more than 10% higher than the second best. It has very competitive performance against covariates such as walking surface and elapsed time. On the OU-ISIR-B dataset, the 90% plus accuracy suggests our method is robust to clothing.
4) **Parameters and Time Complexity:** Our method only has a few (i.e., 3) parameters, and the performance is not sensitive to them. Our method also has low time complexity. On the USF dataset, with the high-dimensional Gabor features, the training time can be within several minutes while the query time per sequence can be less than 1 second.

The rest of this paper is organized as follows, Section II discusses the gait recognition challenges, and provides classifier ensemble solution and its extensions. Details of the random subspace construction, local enhancing, and hybrid decision-level fusion are provided in Section III-V. Our method is experimentally evaluated and compared with other algorithms in Section VI. Section VII concludes this paper.

II. Problem Modelling and Solution

Metric-learning methods are popular for gait recognition. The learned metrics (e.g., in [4], [7], [8], [9], [10], [21], [22], [23]), used as feature extractors, may reduce the effect of covariates to some extent. However, an effective metric can only be learned based on representative training set, while this assumption may not hold in real-world scenarios. Since most of the walking conditions of query gaits are unknown, the training set collected in normal condition cannot represent the whole population. In this case, overfitting may occur. To build a model that generalizes to unknown covariates, in this section we first model the effect of covariates as a partial feature corruption problem with unknown locations, and then we provide the theoretical analysis of the proposed RSM-based classifier ensemble solution and its extensions.

Fig. 2. An example of modelling the covariate effect by image difference between the gallery GEI (i.e., Fig.1(a)) and a probe GEI (i.e., Fig.1(e))

Fig. 3. From left to right: difference images between the gallery GEI Fig.1(a) and the probe GEIs Fig.1(b)-Fig.1(g)

A. Gait Recognition Challenges

We model the effect of covariates by difference images between the probe and gallery. Given a subject with the gallery GEI $I_{gallery}$, and the probe GEIs $\{I_i^{probe}\}_{i=1}^{F}$ in F different walking conditions, we define the corresponding difference images as:

$$\hat{I}_i = I_i^{probe} - I^{gallery}, \quad i \in [1, F], \tag{1}$$

as shown in Fig. 2. Several examples of $\hat{I}_i$ corresponding to different walking conditions are also illustrated in Fig. 3. For an unknown walking condition i, $\hat{I}_i$ indicates the corrupted gait features with unknown locations. Before matching, we need to find a feature extractor T^* that can suppress $\hat{I}_i$, i.e.,

$$T^* = \text{argmin} \, \|\hat{I}_i T\|^2, \quad i \in [1, F], \tag{2}$$

where T can extract features in the column direction of $\hat{I}_i$. However, the locations of corruptions may differ for different walking conditions, as shown in Fig. 3. Given such *non-deterministic* nature of $\hat{I}_i$, from (2) we can see that it is difficult to find a *fixed* T^* that can extract effective features that can generalize to a large number of different covariates. In light of this, an effective T^* should only extract the relevant features *dynamically* for different walking conditions.

B. Problem Formulation

For high dimensional gait feature templates, PCA is usually used for dimension reduction (e.g., [4], [14]). In this work, we instead use the matrix-based two-dimensional PCA (2DPCA) [24] because of its lower computational costs and higher performance.

Let T be the 2DPCA transformation matrix consisting of the leading d non-zero principle components such that $T = [t_1, t_2, ..., t_d]$. Since $[t_1, t_2, ..., t_d]$ are pairwise orthogonal vectors, (2) can be written as:

$$T^* = \text{argmin} \sum_{j=1}^{d} \|\hat{I}_i t_j\|^2, \quad i \in [1, F]. \tag{3}$$

It is difficult for traditional 2DPCA with a fixed $T^* = [t_1, t_2, ..., t_d]$ to reduce the effect of F different walking conditions. In (3), since the terms $\|\hat{I}_i t_j\|^2, j \in [1, d]$ are non-negative, it is possible to reduce the effect of covariates by selecting a subset of $[t_1, t_2, ..., t_d]$ to form a new feature extractor.

To see the reduced effect of covariates, we form different feature extractors by randomly selecting 2 projection directions from $[t_1, t_2, ..., t_d]$. In Fig. 4, we report the corresponding normalized Euclidean distances from the gallery sample in Fig.1(a) to the probe samples in Fig.1(b)-Fig.1(g), respectively. We can see the effect of covariates in Fig.1(b)/Fig.1(e) can be reduced by using $[t_2, t_{35}]$, while the effect in Fig.1(g) can be reduced based on $[t_{15}, t_{61}]$. These observations empirically validate that it is

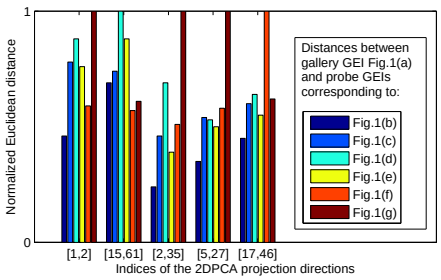


Fig. 4. Normalized Euclidean distances from gallery sample Fig. 1(a) to probe samples Fig.1(b)-Fig.1(g) based on different projection directions

possible to select subsets of $[t_1, t_2, ..., t_d]$ as feature extractors to reduce the effect of covariates. However, the optimal subset may vary for different walking conditions due to the *non-deterministic* nature of $\hat{I}_i$.

Therefore, we formulate the problem of reducing the effect of unknown covariates as a *dynamic* feature selection problem. For an unknown walking condition, we assume the irrelevant features (mainly corresponding to the corrupted pixels, as indicated by $\hat{I}_i$) exist in $m(0 < m < d)$ unknown non-negative terms of (3). We aim to *dynamically* select $N \in [1, d]$ relevant features (mainly corresponding to the uncorrupted pixels) for classification. It is obvious that the probability of selecting N relevant features is 0 when $N > d - m$. When $N \in [1, d-m]$, by randomly selecting N features out of d features, the probability of *not* choosing the m irrelevant features $P(N)$ follows the hypergeometric distribution [25]. This can be expressed as the ratio of binomial coefficients:

$$P(N) = \frac{\binom{d-m}{N}\binom{d-(d-m)}{N-N}}{\binom{d}{N}} = \frac{(d-m)!(d-N)!}{d!(d-m-N)!}, N \in [1, d-m]. \tag{4}$$

$P(N)$ is a generalization measure that is related in someway to the performance of the matching algorithm. Classifiers with greater $P(N)$ tend to generalize well to unseen covariates, and vice versa.

Lemma 1. *Let d, m, N be the numbers of the total features, irrelevant features (with $0 < m < d$), and randomly selected features, respectively, then $P(N)$ in (4) is inversely proportional to N in the range $N \in [1, d-m]$.*

Proof: It can be proved using mathematical induction. From (4), it is easy to see that $P(N+1)/P(N) = 1 - m/(d-N) < 1$ for $N \in [1, d-m]$. This completes the proof of the lemma. ■

It is preferable to set N to smaller values, which can yield greater $P(N)$, according to Lemma 1. Specifically, when $N \ll d$, the hypergeometric distribution can be deemed as a binomial distribution [25], we have

$$P(N) = \binom{N}{N} p^N (1-p)^{(N-N)} = p^N, \quad N \ll d, \tag{5}$$

where p is the probability of randomly selecting one relevant feature. Since $p = 1 - m/d$, (5) can be written as:

$$P(N) = (1 - \frac{m}{d})^N, \quad N \ll d, \tag{6}$$

which clearly reflects the simple relationship between $P(N)$ and the other parameters when we set $N \ll d$.

C. Classifier Ensemble Solution and its Extensions

We use a classifier ensemble strategy for covariate-invariant gait recognition. After repeating the random feature selection process L times (L is a *large number*), L base classifiers with generalization ability $P(N)$ (see (6)) are generated. Classification can then be performed by Majority Voting (MV) [26] among the labeling results of the base classifiers.

Given a query gait and the gallery consisting of c classes, we define the true votes V_{true} as the number of classifiers with correct prediction, while the false votes $\{V^i_{false}\}_{i=1}^{c-1}$ as the incorrectly predicted classifier number distribution over the other $c-1$ classes. Given that, through MV, the final correct classification can be achieved only when $V_{true} > \max\{V^i_{false}\}_{i=1}^{c-1}$.

Ideal Cases Analysis: We make two assumptions in ideal cases:

1) When irrelevant features are not assigned to a classifier (referred to as *unaffected classifier*), correct classification should be achieved.
2) When irrelevant features are assigned to a classifier (referred to as *affected classifier*), the output label should be a random guess from the c classes, with a probability of $1/c$.

Given L classifiers, in the ideal cases the true votes $\bar{V}_{true}$ can be approximated as the sum of the number of *unaffected classifiers* $P(N)L$ (based on *the law of large numbers* [27]) and the number of *affected classifiers*:

$$\bar{V}_{true} \approx \text{round}\left(P(N)L + \frac{(1-P(N))L}{c}\right). \tag{7}$$

Similarly, in the ideal cases the false votes should correspond to the number of *affected classifiers* for the other $c-1$ classes. Based on assumption 2), its maximum value, $\max\{\bar{V}^i_{false}\}_{i=1}^{c-1}$, can be roughly expressed as the corresponding mean value:

$$\max\{\bar{V}^i_{false}\}_{i=1}^{c-1} \approx \text{mean}\{\bar{V}^i_{false}\}_{i=1}^{c-1} \approx \text{round}\left(\frac{(1-P(N))L}{c}\right). \tag{8}$$

Correct classification can be achieved when $\bar{V}_{true} > \max\{\bar{V}^i_{false}\}_{i=1}^{c-1}$. From (7) and (8), it is obvious that this condition can be met when $P(N) > \varepsilon$ (ε is a small positive number).

*From ideal cases analysis, it can be seen that we **only care** about the number of unaffected classifiers, instead of which ones are. According to (6), by setting N to a small number to make $P(N) > \varepsilon$, the corresponding classifier ensemble method **is insensitive to the locations of corrupted features** (since $\bar{V}_{true} > max\{\bar{V}^i_{false}\}_{i=1}^{c-1}$, as supported by (7) and (8)). Therefore, it is robust to a large number of covariates as long as the irrelevant feature number is not extremely large (i.e., $m < d$ in (6)).*

Real Cases Analysis: Here the two assumptions made in the ideal cases analysis do not hold precisely: 1) *unaffected classifiers* do not always make the correct classification; 2) although the *affected classifiers* would result in label assigning in a relatively random manner, it is difficult to estimate $\max\{V^i_{false}\}_{i=1}^{c-1}$ precisely. However, in real cases we can deem $\bar{V}_{true}$ and $\max\{\bar{V}^i_{false}\}_{i=1}^{c-1}$ from ideal cases analysis as the upper bound and lower bound, respectively, i.e.,

$$\begin{aligned} V_{true} &\le \bar{V}_{true}, \\ \max\{V^i_{false}\}_{i=1}^{c-1} &\ge \max\{\bar{V}^i_{false}\}_{i=1}^{c-1}. \end{aligned} \tag{9}$$

Since correct classification can be achieved only when $V_{true} > \max\{V^i_{false}\}_{i=1}^{c-1}$, our objective is to increase V_{true} and decrease $\max\{V^i_{false}\}_{i=1}^{c-1}$. Since it is difficult to estimate $\max\{V^i_{false}\}_{i=1}^{c-1}$ precisely, we relax the objective by decreasing $\sum_{i=1}^{c-1} V^i_{false}$. Then we can reformulate our objective as a problem of suppressing Γ defined as:

$$\Gamma = \frac{\sum_{i=1}^{c-1} V^i_{false}}{V_{true} + \epsilon}, \tag{10}$$

where ϵ is a small positive number to avoid trivial results. To suppress Γ in (10), we extend the classifier ensemble method by proposing two strategies:

1) Local Enhancing (LE): If we use the output labels of all the L classifiers as valid votes such that $L = V_{true} + \sum_{i=1}^{c-1} V_{false}^{i}$, then according to (10) we can get $\Gamma \propto L/(V_{true} + \epsilon)$. Since L is a constant, the first strategy is to increase V_{true}. To realize it, based on supervised learning, for each subspace we extract more discriminant features in order to enhance the classification performance. We name the second-stage supervised feature extraction (after the first-stage random feature selection) as Local Enhancing (LE). Details of LE used in this work are presented in Section IV.

2) Hybrid Decision-level Fusion (HDF): The second strategy is to decrease $\sum_{i=1}^{c-1} V_{false}^{i}$ by *dynamically* eliminating classifiers corresponding to the irrelevant features, before MV. Irrelevant features would lead to label assignment in a relatively random manner. Based on the *same* irrelevant features, a classifier pair corresponding to two *different* LE methods would produce two random labels, which are unlikely to be the same. Based on the "AND" rule, classifier pairs with different output labels are deemed as invalid votes and simply discarded. Although this scheme may also decrease V_{true} to some extent, it significantly reduces the value of $\sum_{i=1}^{c-1} V_{false}^{i}$ to suppress Γ in (10). Details of HDF are presented in Section V.

III. Random Subspace Construction

We construct the feature space using 2DPCA, from which we can generate L feature extractors. Given n gait templates (e.g., GEIs) $\{I_i \in \mathbb{R}^{N_1 \times N_2}\}_{i=1}^{n}$ in the training set (i.e., the gallery), we can compute the scatter matrix $S^* = \frac{1}{n}\sum_{i=1}^{n}(I_i - \mu)^T(I_i - \mu)$, where $\mu = \frac{1}{n}\sum_{i=1}^{n} I_i$. The eigenvectors of S^* can be calculated, and the leading d eigenvectors associated with non-zero eigenvalues are retained as candidates $T = [t_1, t_2, ..., t_d] \in \mathbb{R}^{N_2 \times d}$ to construct the random subspaces. By repeating L times the process of randomly selecting subsets of T (with size $N \ll d$), the random subspaces $\{R^l \in \mathbb{R}^{N_2 \times N}\}_{l=1}^{L}$ are generated and can be used as random feature extractors. Then a gait template $I \in \mathbb{R}^{N_1 \times N_2}$ can be represented as a set of random features $\{X^l \in \mathbb{R}^{N_1 \times N}\}_{l=1}^{L}$ such that

$$X^l = IR^l, \qquad l = 1, 2, ..., L. \tag{11}$$

Note the random feature extraction is only performed in the column direction of I, with the dimension reduced to $\mathbb{R}^{N_1 \times N}$.

IV. Local Enhancing (LE)

The random features can be used for classification directly. However, these features may be redundant and less discriminant since 1) the random feature extraction process based on (11) is performed only in the column direction and 2) the random feature extractors are trained in an unsupervised manner, without using the label information. As a result, these features may lead to high computational costs and low recognition accuracies. Local enhancing is used to address this issue, as discussed in Section II-C. In each subspace, we further extract more discriminant features based on two different supervised learning methods, i.e., two-dimensional LDA (2DLDA) [28], and IDR/QR [29]. In this paper, these two types of feature extractors are referred to as local enhancers (i.e., LE1 and LE2), and the process of training 2DLDA-based LE1 and IDR/QR-based LE2 are summarized in Algorithm 1 and Algorithm 2, respectively. More details about 2DLDA and IDR/QR can be found in [28] [29].

In the l^{th} subspace, a gait template I with extracted random feature matrix X^l (or the corresponding concatenated vector $\hat{X}^l$) can be enhanced by W^l (the output of Algorithm 1) or V^l (the output of Algorithm 2) through

$$x^l = (W^l)^T X^l, \quad l \in [1, L], \tag{12}$$

or

$$\hat{x}^l = (V^l)^T \hat{X}^l, \quad l \in [1, L]. \tag{13}$$

x^l (resp. $\hat{x}^l$) are the enhanced random features by LE1 (resp. LE2), and they can be used for classification in the l^{th} subspace.

Based on the (enhanced) random features, we also reconstruct the GEIs corresponding to Fig. 1(a)-1(g), which can be found in this paper's supplemental material.

Algorithm 1 2DLDA-based LE1

Input: Training set (i.e., the gallery) $\{I_i \in \mathbb{R}^{N_1 \times N_2}\}_{i=1}^{n}$ in c classes, random feature extractors $\{R^l \in \mathbb{R}^{N_2 \times N}\}_{l=1}^{L}$, and the number of LE1 projection directions M;
Output: LE1-based feature extractors $\{W^l \in \mathbb{R}^{N_1 \times M}\}_{l=1}^{L}$;
Step 1: Random feature extraction on training set $X_i^l = I_i R^l, \quad i = 1, 2, ...n, \quad l \in [1, L]$;
for $l = 1$ to L **do**
Step 2: For $\{X_i^l\}_{i=1}^{n}$, letting μ^l be the global centroid, D_j^l be the j^{th} class (out of c classes) with sample number n_j and centroid m_j^l;
Step 3: Calculating $S_b^l = \sum_{j=1}^{c} n_j (m_j^l - \mu^l)(m_j^l - \mu^l)^T$;
Step 4: Calculating $S_w^l = \sum_{j=1}^{c} \sum_{X_i^l \in D_j^l} (X_i^l - m_j^l)(X_i^l - m_j^l)^T$;
Step 5: Setting $W^l = \{\phi_i\}_{i=1}^{M}$, which are the M leading eigenvectors of $(S_w^l)^{-1} S_b^l$.
end for

Algorithm 2 IDR/QR-based LE2

Input: Training set (i.e., the gallery) $\{I_i \in \mathbb{R}^{N_1 \times N_2}\}_{i=1}^{n}$ in c classes, random feature extractors: $\{R^l \in \mathbb{R}^{N_2 \times N}\}_{l=1}^{L}$, and the number of LE2 projection directions M;
Output: LE2-based feature extractors $\{V^l \in \mathbb{R}^{S \times M}\}_{l=1}^{L}$, where $S = N_1 N$;
Step 1: Random feature extraction on training set $X_i^l = I_i R^l, \quad i = 1, 2, ...n, \quad l \in [1, L]$;
Step 2: Concatenating $X_i^l \in \mathbb{R}^{N_1 \times N}$ to $\hat{X}_i^l \in \mathbb{R}^{S}, i = 1, 2, ..., n, \quad l \in [1, L]$;
for $l = 1$ to L **do**
Step 3: For $\{\hat{X}_i^l\}_{i=1}^{n}$, letting $\hat{\mu}^l$ be the global centroid, $\hat{D}_j^l$ be the j^{th} class (out of c classes) with sample number n_j and centroid $\hat{m}_j^l$;
Step 4: Constructing the set of within-class centroids: $C = [\hat{m}_1^l, \hat{m}_2^l, ..., \hat{m}_c^l]$, and performing QR decomposition [30] of C as $C = QR$, where $Q \in \mathbb{R}^{S \times c}$;
Step 5: After setting $e_j = (1, 1, ..., 1)^T \in \mathbb{R}^{n_j}$, computing $H_b^l = [\sqrt{n_1}(\hat{m}_1^l - \hat{\mu}^l), \sqrt{n_2}(\hat{m}_2^l - \hat{\mu}^l), ..., \sqrt{n_c}(\hat{m}_c^l - \hat{\mu}^l)]$, $H_w^l = [\hat{D}_1^l - \hat{m}_1^l {e_1}^T, \hat{D}_2^l - \hat{m}_2^l {e_2}^T, ..., \hat{D}_c^l - \hat{m}_c^l {e_c}^T]$;
Step 7: Calculating $S_B^l = Y^T Y$, where $Y = (H_b^l)^T Q$;
Step 8: Calculating $S_W^l = Z^T Z$, where $Z = (H_w^l)^T Q$;
Step 9: For $(S_W^l)^{-1} S_B^l$, calculating its M leading eigenvectors, $U = \{\phi_i\}_{i=1}^{M}$;
Step 10: Setting $V^l = QU$;
end for

V. Hybrid Decision-level Fusion (HDF)

Given the random feature extractors $\{R^l\}_{l=1}^{L}$, the corresponding LE1-based $\{W^l\}_{l=1}^{L}$, and LE2-based $\{V^l\}_{l=1}^{L}$, we can extract

(a) (b)

Fig. 5. Images from (a) the USF dataset [5], and (b) the OU-ISIR-B dataset [16]

TABLE I
12 PRE-DESIGNED EXPERIMENTS ON THE USF DATASET

Exp.	**A**	**B**	**C**	**D**	**E**	**F**
# Seq.	122	54	54	121	60	121
Covariates	V	H	VH	S	SH	SV
Exp.	**G**	**H**	**I**	**J**	**K**	**L**
# Seq.	60	120	60	120	33	33
Covariates	SHV	B	BH	BV	THC	STHC

Abbreviation note: V-View, H-Shoe, S-Surface, B-Briefcase, T-Time, C-Clothing

TABLE II
DIFFERENT CLOTHING COMBINATIONS IN THE OU-ISIR-B DATASET

Type	s_1	s_2	s_3	Type	s_1	s_2	Type	s_1	s_2
3	RP	HS	Ht	**0**	CP	CW	**F**	CP	FS
4	RP	HS	Cs	**2**	RP	HS	**G**	CP	Pk
6	RP	LC	Mf	**5**	RP	LC	**H**	CP	DJ
7	RP	LC	Ht	**9**	RP	FS	**I**	BP	HS
8	RP	LC	Cs	**A**	RP	Pk	**J**	BP	LC
C	RP	DJ	Mf	**B**	RP	Dj	**K**	BP	FS
X	RP	FS	Ht	**D**	CP	HS	**L**	BP	Pk
Y	RP	FS	Cs	**E**	CP	LC	**M**	BP	DJ
N	SP	HS	-	**P**	SP	Pk	**R**	RC	-
S	Sk	HS	-	**T**	Sk	FS	**U**	Sk	PK
V	Sk	DJ	-	**Z**	SP	FS	-	-	-

Abbreviation note: RP - Regular pants, BP - Baggy pants, SP - Short pants, HS - Half shirt, FS - Full shirt, LC - Long coat, CW - Casual wear, RC - Rain coat, Ht - Hat, CP-Casual pants, Sk - Skirt, Pk - Parker, DJ - Down jacket, Cs - Casquette cap, Mf - Muffler, s_i - i^{th} clothes slot

the sets of LE1-based (resp. LE2-based) features from a gait template using (11), and (12) (resp. (13)). Based on the new gait descriptors, for the l^{th} subspace let $[\phi_1^l, \phi_2^l, ..., \phi_c^l]$ be the centroids of the gallery. For a query sequence P^l (including n_p gait cycles) with the corresponding descriptors $[p_1^l, p_2^l, ..., p_{n_p}^l]$, the distance between P^l and the j^{th} class centroid ϕ_j^l is defined as: $\delta(P^l, \phi_j^l) = \frac{1}{n_p}\sum_{i=1}^{n_p} \|p_i^l - \phi_j^l\|, j \in [1, c]$. Nearest Neighbour (NN) rule is employed for classification. Let $\{\omega_j\}_{j=1}^c$ be class labels of c subjects in the gallery, then the output label $\Omega^l(P^l) \in \{\omega_j\}_{j=1}^c$ of the l^{th} base classifier can be expressed as:

$$\Omega^l(P^l) = \underset{\omega_j}{\operatorname{argmin}}\, \delta(P^l, \phi_j^l), \quad j \in [1, c]. \tag{14}$$

Then we use HDF to identify the query gait $P = \{P^l\}_{l=1}^L$. For simplicity reasons, let $\{\Omega_{LE1}^l(P^l), \Omega_{LE2}^l(P^l)\}_{l=1}^L$ be the output labels of the classifier pairs corresponding to LE1 and LE2. HDF can be performed by assigning the class label to $\bar{\Omega}(P)$ with:

$$\bar{\Omega}(P) = \underset{\omega_j}{\operatorname{argmax}} \sum_{l=1}^{L} \Theta_{\omega_j}^l, \quad j \in [1, c], \tag{15}$$

where

$$\Theta_{\omega_j}^l = \begin{cases} 1, & \text{if } \Omega_{LE1}^l(P^l) = \Omega_{LE2}^l(P^l) = \omega_j, \\ 0, & \text{otherwise,} \end{cases} \quad j \in [1, c]. \tag{16}$$

It can be seen that $\sum_{j=1}^c \sum_{l=1}^L \Theta_{\omega_j}^l \leq L$. When there are a large number of irrelevant features for the query gait P, most of the corresponding votes can be eliminated such that $\sum_{j=1}^c \sum_{l=1}^L \Theta_{\omega_j}^l \ll L$. In this case, $\sum_{i=1}^{c-1} V_{false}^i$ of (10) can be significantly reduced while V_{true} of (10) are less affected. As such, Γ of (10) can be effectively suppressed.

VI. EXPERIMENTS

In this section, first we describe the two benchmark databases, namely USF [5] and OU-ISIR-B [16] datasets. Then we discuss the parameter settings and the time complexity of our method. Performance gain analysis by using the proposed LE and HDF strategies is also provided, followed by the algorithm comparison.

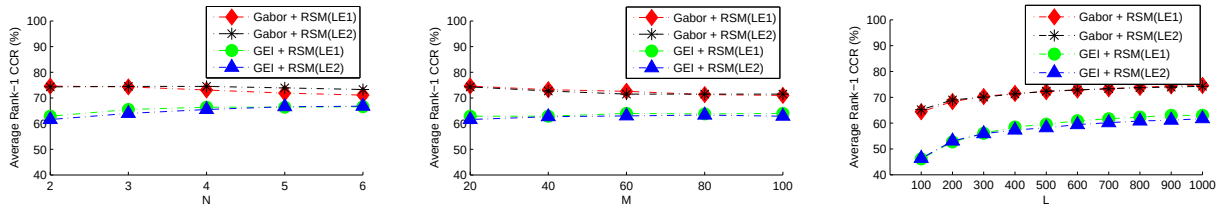
A. Datasets

The USF dataset [5] is a large outdoor gait database, consisting of 122 subjects. A number of covariates are presented: camera viewpoints (left/right), shoes (type A/type B), surface types (grass/concrete), carrying conditions (with/without a briefcase), elapsed time (May/Nov.), and clothing. Based on a gallery with 122 subjects, there are 12 pre-designed experiments for algorithm evaluation, which are summarized in Table I. The second gait database, namely, the OU-ISIR-B dataset [16], was recently constructed for studying the effect of clothing. The evaluation set includes 48 subjects walking on a treadmill with up to 32 types of clothes combinations, as listed in Table II. The gallery consists of the 48 subjects in standard clothes (i.e., type 9: regular pants + full shirt), whereas the probe set includes 856 gait sequences of the same 48 subjects in the other 31 clothes types. Several images from the USF and OU-ISIR-B datasets are shown in Fig. 5. For both datasets, the gallery is used for training. In this work, we employ two templates separately, namely, GEI (with the default size 128×88 pixels) and downsampled Gabor-filtered GEI (referred to as Gabor, with size 320×352 pixels).

We employ rank-1/rank-5 Correct Classification Rate (CCR) to evaluate the performance. Considering the random nature of our method, the results of different runs may vary to some extent. We repeat all the experiments 10 times and report the statistics (mean, std, maxima and minima) in Table III and Table VII for both datasets. The small std values (listed in Table III and Table VII) of the 10 runs indicate the stability of our method. Therefore, throughout the rest of the paper, we only report the mean values.

B. Parameter Settings and Time Complexity Analysis

There are only 3 parameters in our method, namely, the dimension of random subspace N, the base classifier number L, and the number of projection directions M for LE1/LE2. As discussed in Section II-B, we should set N to a small number for great generalization ability $P(N)$. The classifier number L should be large, since our classifier ensemble solution is based on *the law of large numbers* [27] (see Section II-C). Like most subspace learning methods, the performance should not be sensitive to the number of projection directions M for LE1/LE2 (unless it is extremely small). On the USF dataset (Exp. A-L) we check the average performance sensitivity to N, M, and L, based on Gabor and GEI templates, respectively. By empirically setting $L = 1000$



(a) (b) (c)

Fig. 6. On the performance sensitivity to the parameters on the USF dataset: (a) N is the dimension of the random subspace; (b) M is the number of projection directions of LE1/LE2; (c) L is the classifier number

TABLE III
PERFORMANCE STATISTICS IN TERMS OF RANK-1 CCRS (%) FOR 10 RUNS ON THE USF DATASET

-	maxima	minima	std	mean
GEI + RSM	52.44	50.73	0.62	51.45
GEI + RSM(LE1)	64.08	62.36	0.49	63.01
GEI + RSM(LE2)	62.79	59.81	0.91	61.72
GEI + RSM-HDF	70.90	69.20	0.50	70.01
Gabor + RSM	67.58	66.30	0.46	67.06
Gabor + RSM(LE1)	75.29	73.89	0.49	74.56
Gabor + RSM(LE2)	74.95	73.45	0.36	74.27
Gabor + RSM-HDF	82.12	80.30	0.53	81.17

TABLE IV
RUNNING TIME (SECONDS) ON THE USF DATASET

-	Training Time	Query Time Per Seq.
GEI + RSM(LE1)	91.42	0.58
GEI + RSM(LE2)	28.66	0.26
Gabor + RSM(LE1)	320.09	0.60
Gabor + RSM(LE2)	32.79	0.31

and $M = 20$, we run our methods with N set within the range $[2, 6]$. The results in Fig. 6(a) indicate that the performance is not sensitive to N. Based on $L = 1000$, and $N = 2$, we conduct experiments with $M = [20, 40, 60, 80, 100]$. The results in Fig. 6(b) suggest that the performance is stable for M with different values. By setting $N = 2$ and $M = 20$, we can also see from Fig. 6(c) that the performance is not decreasing with respect to the increasing number of classifiers. These observations are consistent with our expectation that the performance is not sensitive to these 3 parameters. For the rest of this paper, we only report the results based on $N = 2$, $M = 20$ and $L = 1000$.

We analyze the time complexity for training the L LE1/LE2-based classifiers. For a LE1-based classifier in Algorithm 1, it takes $O(nNN_1^2)$ (resp. $O(cNN_1^2)$) for S_w^l (resp. S_b^l) and $O(N_1^3)$ for eigenvalue decomposition. N_1 is the template's row number (i.e., $N_1 = 128$ for GEI, $N_1 = 320$ for Gabor), while n and c are the number of training samples and classes, respectively. Since, in our case, $n > c$ and $n > N_1$, the time complexity for generating L LE1-based classifiers can be written as $O(LnNN_1^2)$. For a LE2-based classifier in Algorithm 2, it takes $O(c^2S)$ for the QR decomposition, where $S = NN_1$. Calculating S_B^l and S_W^l requires $O(c^3)$ and $O(c^2n)$, while calculating Z and Y requires $O(nSc)$ and $O(Sc^2)$. Solving the eigenvalue decomposition problem of $(S_W^l)^{-1}S_B^l$ takes $O(c^3)$ and the final solution V^l is obtained by matrix multiplication, which takes $O(cSM)$. Since in our case $n > c$ and $S > c$, the time complexity for generating L LE2-based classifiers is $O(LnSc)$, which can also be written as $O(LnNN_1c)$. We run the matlab code of our method on a PC with an Intel Core i5 3.10 GHz processor and 16GB RAM. For the USF dataset, we report the training/query time (with $L = 1000$) in Table IV. Note the classifiers are trained in an offline manner and can be used to identify probe sets with various covariates. It is clear that LE2 is very efficient when the dimension is large. For example, based on Gabor templates, LE2 only takes about $1/10$ of LE1's training time.

C. Performance Gain Analysis and Algorithms Comparison

Based on GEI and Gabor templates, we evaluate the effect of suppressing the ratio of false votes $\sum_{i=1}^{c-1} V_{false}^i$ to true votes V_{true}, as defined in (10) by using LE and HDF. For a probe set with K gait sequences, according to (10) we define $\hat{\Gamma}$ as

$$\hat{\Gamma} = \text{median}\left\{ \frac{\sum_{i=1}^{c-1} V_{false}^i}{V_{true} + \epsilon} \right\}_{k=1}^{K}, \quad (17)$$

which is used to reflect the general ratio over the whole probe set. We set $\epsilon = 1$ to avoid the trivial results. Over the 12 probe sets (A-L) on the USF dataset, the distribution of $\hat{\Gamma}$ and the performance are reported in Fig. 7 and Table V. We can observe that, LE1/LE2 can reduce $\hat{\Gamma}$ to some extent, and RSM(LE1)/RSM(LE2) is less sensitive to covariates such as viewpoint, shoe, and briefcase (A-C, H-J). On the other hand, RSM-HDF can significantly suppress $\hat{\Gamma}$ and yields competitive accuracies in tackling the hard problems caused by walking surface and elapsed time (D-G, K-L).

On the USF dataset, we also compare our method Gabor+RSM-HDF with the recently published works, with the rank-1/rank-5 CCRs reported in Table VI. These works include Baseline [5], Hidden Markov Models(HMM) [31], GEI+Fusion [4], CSA+DATER [8], DNGR [13], Matrix-based Marginal Fisher Analysis (MMFA) [21], GTDA [9], Linearization of DLLE (DLLE/L) [7], TRIMAP [10], Image-to-Class [11], Gabor-PDF+LGSR [12], CGI+Fusion [14], Sparse Reconstruction based Metric Learning (SRML) [23], and Sparse Bilinear Discriminant Analysis (SBDA) [22]. From Table VI, we can see that in terms of rank-1 CCRs, our method outperforms other algorithms on all the 12 probe sets. Our method has an average rank-1 CCR more than 10% higher than the second best method (i.e., Gabor-PDF+LGSR [12]), and also the highest average rank-5 CCR. It is significantly superior than others on the challenging tasks D-G, and K-L, which are under the influences of walking surface, elapsed time and the combination of other covariates. Although these walking conditions may significantly corrupt the gait features, our proposed HDF scheme (based on LE1 and LE2) can still suppress Γ of (10), leading to competitive accuracies. We notice that our method only has 42% rank-1 CCRs for probe sets K-L. In these cases, elapsed time is coupled with other covariates like walking surface, clothing, and shoe, as listed in Table I. These walking conditions may significantly increase the number of irrelevant features m, which would result in a lower $P(N)$ in (6). According to Section II-C, a lower $P(N)$ would lead to a higher Γ in (10), which contributes negatively to

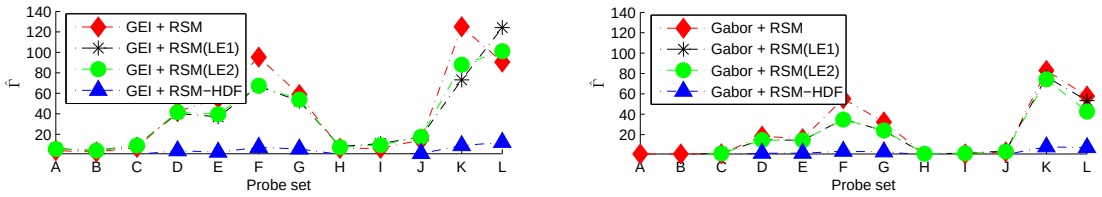


Fig. 7. Over the 12 probe sets on the USF dataset, the distribution of $\hat{\Gamma}$ (i.e., the general ratio of false votes to true votes)

TABLE V
RANK-1 CCRS (%) OF OUR METHODS ON THE USF DATASET

Exp.	A	B	C	D	E	F	G	H	I	J	K	L	Avg.
GEI + RSM	89	93	82	24	27	16	16	83	69	54	18	9	51.36
GEI + RSM(LE1)	95	94	86	42	50	23	35	88	88	72	26	19	62.88
GEI + RSM(LE2)	96	94	82	40	46	27	29	87	83	72	19	18	61.70
GEI + RSM-HDF	98	95	88	54	60	37	44	90	93	83	33	21	70.16
Gabor + RSM	96	94	87	47	45	24	38	96	97	85	25	27	67.13
Gabor + RSM(LE1)	100	95	93	62	63	42	50	97	96	89	23	29	74.65
Gabor + RSM(LE2)	98	94	93	60	58	39	47	97	97	92	34	31	74.09
Gabor + RSM-HDF	100	95	94	73	73	55	64	97	99	94	42	42	81.15

TABLE VI
ALGORITHMS COMPARISON IN TERMS OF RANK-1/RANK-5 CCRS (%) ON THE USF DATASET

Exp.	A	B	C	D	E	F	G	H	I	J	K	L	Avg.
Rank-1 CCRs													
Baseline [5]	73	78	48	32	22	17	17	61	57	36	3	3	40.96
HMM [31]	89	88	68	35	28	15	21	85	80	58	17	15	53.54
GEI + Fusion [4]	90	91	81	56	64	25	36	64	60	60	6	15	57.66
CSA + DATER [8]	89	93	80	44	45	25	33	80	79	60	18	21	58.51
DNGR [13]	85	89	72	57	66	46	41	83	79	52	15	24	62.81
MMFA [21]	89	94	80	44	47	25	33	85	83	60	27	21	59.90
GTDA [9]	91	93	86	32	47	21	32	95	90	68	16	19	60.58
DLLE/L [7]	90	89	81	40	50	27	26	65	67	57	12	18	51.83
TRIMAP [10]	92	94	86	44	52	27	33	78	74	65	21	15	59.66
Image-to-Class [11]	93	89	81	54	52	32	34	81	78	62	12	9	61.19
Gabor-PDF + LGSR [12]	95	93	89	62	62	39	38	94	91	78	21	21	70.07
CGI + Fusion [14]	91	93	78	51	53	35	38	84	78	64	3	9	61.69
SRML [23]	93	94	85	52	52	37	40	86	85	68	18	15	66.50
SBDA [22]	93	94	85	51	50	29	36	85	83	68	18	24	61.35
Gabor + RSM-HDF (Ours)	**100**	**95**	**94**	**73**	**73**	**55**	**64**	**97**	**99**	**94**	**42**	**42**	**81.15**
Rank-5 CCRs													
Baseline [5]	88	93	78	66	55	42	38	85	78	62	12	15	64.54
GEI + Fusion [4]	94	94	93	78	81	56	53	90	83	82	27	21	76.23
CSA + DATER [8]	96	96	94	74	79	53	57	93	91	83	40	36	77.86
DNGR [13]	96	94	89	85	81	68	69	96	95	79	46	39	82.05
MMFA [21]	98	98	94	76	76	57	60	95	93	84	48	39	79.90
GTDA [9]	98	**99**	97	68	68	50	56	95	**99**	84	40	40	77.58
DLLE/L [7]	95	96	93	74	78	50	53	90	90	83	33	27	71.83
TRIMAP [10]	96	**99**	95	75	72	54	58	93	88	85	43	36	77.75
Image-to-Class [11]	97	98	93	81	74	59	55	94	95	83	30	33	79.17
Gabor-PDF + LGSR [12]	99	94	96	**89**	**91**	64	64	**99**	98	92	39	45	85.31
CGI + Fusion [14]	97	96	94	77	77	56	58	98	97	86	27	24	79.12
SBDA [22]	98	98	94	74	79	57	60	95	95	84	40	40	79.93
Gabor + RSM-HDF (Ours)	**100**	98	**98**	85	84	**73**	**79**	98	**99**	**98**	**55**	**58**	**88.59**

the performance. Nevertheless, experimental results suggest our method is robust to most covariates in the outdoor environment.

D. In Tackling the Clothing Challenges

Clothing was deemed as the most challenging covariate [15], and there are only a few works that have studied the effect of various clothes types. Recently, Hossain et al. built the OU-ISIR-B dataset [16] with 32 combinations of clothes types, as shown Table II. Based on an additional training set that covers all the possible clothes types, they proposed an adaptive part-based method [16] for clothing-invariant gait recognition. On this dataset, based on Gabor templates, we evaluate our methods RSM(LE1), RSM(LE2) and RSM-HDF. The statistics of our methods over 10 runs are reported in Table VII. Compared with the part-based method [16], Gabor+RSM-HDF can yield a much higher accuracy, as shown in Table VIII. It is worth noting that different from [16], our method does not require the training set that covers all the possible clothes types and can generalize well to unseen clothes types.

We also study the effect of different clothes types, and the rank-1 CCRs for 31 probe clothes types are reported in Fig. 8. For most of the clothes types, our method can achieve more than 90% rank-1 accuracies. However, the performance can be affected

TABLE VII
PERFORMANCE STATISTICS IN TERMS OF RANK-1 CCRS (%) FOR 10 RUNS ON THE OU-ISIR-B DATASET

-	maxima	minima	std	mean
Gabor + RSM(LE1)	89.49	86.92	0.76	87.92
Gabor + RSM(LE2)	87.97	86.80	0.36	87.52
Gabor + RSM-HDF	91.00	90.07	0.32	90.72

TABLE VIII
ALGORITHMS COMPARISON IN TERMS OF RANK-1 CCRS (%) ON THE OU-ISIR-B DATASET

Part-based method [16]	Gabor + RSM-HDF (Ours)
63.9	90.7

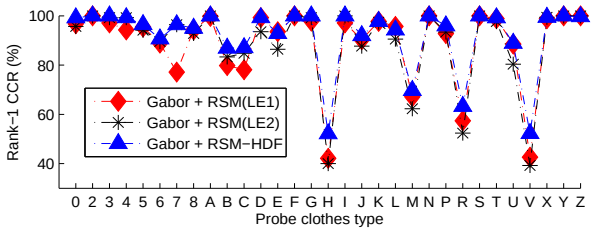


Fig. 8. Performance distribution with respect to 31 probe clothes types on the OU-ISIR-B dataset

with several clothes types that cover large parts of the human body. In this case, a large number of irrelevant features m would result in a higher Γ in (10), which would hamper the performance (as discussed in Section VI-C). Specifically, the results are less satisfactory when the following 3 clothes types are encountered: 1) clothes type R, (i.e., raincoat) with a rank-1 CCR of 63.3%; 2) clothes type H, (i.e., casual pants + down jacket) with a rank-1 CCR of 52.1%; 3) clothes type V, (i.e., skirt + down jacket) with a rank-1 CCR of 52.2%. Nevertheless, in general the results suggest that our method is robust to clothing.

VII. CONCLUSION

In this paper, we model the effect of covariates as a partial feature corruption problem with unknown locations and propose a RSM-based classifier ensemble solution. The theoretical basis suggests that its insensitivity to a large number of covariates in ideal cases. To tackle the hard problems in real cases, we then propose two strategies, i.e., LE and HDF, to suppress the ratio of false votes to true votes before the majority voting. Experimental results suggest that our method is less sensitive to the most challenging covariates like clothing, walking surface, and elapsed time. Our method has only 3 parameters, to which the performance is not sensitive. It can be trained within minutes and perform real-time recognition in less than 1 second, which suggests that it is practical in real-world applications.

ACKNOWLEDGEMENTS

The authors would like to thank the support from the Royal Society's International Exchange Programme (IE120092) and the constructive comments from the anonymous reviewers.

REFERENCES

[1] P.K. Larsen, E.B. Simonsen, and N. Lynnerup, "Gait Analysis in Forensic Medicine", *J Forensic Sci,* vol.53, pp.1149-1153, 2008.
[2] I. Bouchrika, M. Goffredo, J. Carter, and M. S. Nixon, "On Using Gait in Forensic Biometrics," *J Forensic Sci,* Vol.56, no.4, pp.882-889, 2011.
[3] D. Cunado, M. S. Nixon, and J. Carter, "Automatic extraction and description of human gait models for recognition purposes," *CVIU,* vol.90, no.1, pp. 1-41, 2003.
[4] J. Han and B. Bhanu, "Individual Recognition using Gait Energy Image," *T-PAMI,* vol.28, no.2, pp.316-322, 2006.
[5] S. Sarkar, P.J. Phillips, Z. Liu, I.R. Vega, P . Grother, and K.W. Bowyer, "The HumanID Gait Challenge Problem: Data Sets, Performance, and Analysis," *T-PAMI,* vol.27, no.2, pp.162-177, 2005.
[6] 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," *T-IFS,* vol.7, no.5, pp.1511-1521, 2012
[7] X. Li, S. Lin, S. Yan, and D. Xu, "Discriminant Locally Linear Embedding With High-Order Tensor Data," *T-SMC-B,* vol.38, no.2, pp.342-352, 2008
[8] D. Xu, S. Yan, D. Tao, L. Zhang, X. Li, and H.J. Zhang, "Human Gait Recognition With Matrix Representation," *T-CSVT,* vol.16, no.7, pp.896-903, 2006.
[9] D. Tao, X. Li, X. Wu, and S.J. Maybank, "General Tensor Discriminant Analysis and Gabor Features for Gait Recognition," *T-PAMI,* vol.29, no.10, pp.1700-1715, 2007.
[10] C. Chen, J. Zhang, and R. Fleischer, "Distance Approximating Dimension Reduction of Riemannian Manifolds," *T-SMC-B,* vol.40, no.1, pp.208-217, 2010.
[11] Y. Huang, D. Xu, and T. Cham, "Face and Human Gait Recognition Using Image-to-Class Distance," *T-CSVT,* vol.20, no.3, pp.431-438, 2010.
[12] D. Xu, Y. Huang, Z. Zeng, and X. Xu, "Human Gait Recognition Using Patch Distribution Feature and Locality-Constrained Group Sparse Representation," *T-IP,* vol.21, no.1, pp.316-326, 2012.
[13] Z. Liu and S. Sarkar, "Improved Gait Recognition by Gait Dynamics Normalization," *T-PAMI,* vol.28, no.6, pp.863-876, 2006.
[14] C. Wang, J. Zhang, L. Wang, J. Pu, and X. Yuan, "Human Identification Using Temporal Information Preserving Gait Template," *T-PAMI,* vol.34, no.11, pp. 2164-2176, 2012.
[15] D. Matovski, M. S. Nixon, S. Mahmoodi, and J. Carter, "The Effect of Time on Gait Recognition Performance," *T-IFS,* vol.7, no.2, pp.543-552, 2012.
[16] M.A. Hossain, Y. Makihara, J. Wang, and Y. Yagi, "Clothing-invariant Gait Identification using Part-based Clothing Categorization and Adaptive Weight Control," *PR,* Vol.43, No.6, pp.2281-2291, 2010.
[17] T. K. Ho, "The Random Subspace Method for Constructing Decision Forests," *T-PAMI,* vol.20, no.8, pp.832-844, 1998.
[18] X. Wang and X. Tang, "Random Sampling for Subspace Face Recognition," *IJCV,* Vol. 70, No. 1, pp 91-104, 2006
[19] Y. Guan, C.-T. Li, and Y. Hu, "Random Subspace Method for Gait Recognition," *ICMEW,* pp.284-289, 2012
[20] Y. Guan, C.-T. Li, and Y. Hu, "Robust Clothing-Invariant Gait Recognition," *IIH-MSP,* pp.321-324, 2012
[21] 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," *T-IP,* vol.16, no.11, pp.2811-2821, 2007.
[22] Z. Lai, Y. Xu, Z. Jin, and D. Zhang, "Human Gait Recognition via Sparse Discriminant Projection Learning," *T-CSVT,* 2014
[23] J. Lu, G. Wang, and P. Moulin, "Human Identity and Gender Recognition From Gait Sequences With Arbitrary Walking Directions," *T-IFS,* vol.9, no.1, pp.51-61, 2014
[24] J. Yang, D. Zhang, A.F. Frangi, and J. Yang, "Two-dimensional PCA: a New Approach to Appearance-based Face Representation and Recognition," *T-PAMI,* vol.26, no.1, pp.131-137, Jan. 2004.
[25] C. Walck, "Handbook on Statistical Distributions for Experimentalists," University of Stockholm Press, 2007
[26] J. Kittler, M. Hatef, R. P.W. Duin, and J. Matas, "On Combining Classifiers," *T-PAMI,* vol.20, no.3, pp.226-239, 1998
[27] C. M. Grinstead, J. L. Snell, "Introduction to Probability," American Mathematical Society, 1997.
[28] M. Li, and B. Yuan, "2D-LDA: A Statistical Linear Discriminant Analysis for Image Matrix," *PRL,* vol.26, pp.527-532, 2005.
[29] J. Ye, Q. Li, H. Xiong, H. Park, R. Janardan, and V. Kumar, "IDR/QR: An Incremental Dimension Reduction Algorithm via QR Decomposition," *T-KDE,* vol.17, no.9, pp.1208-1222, 2005.
[30] L. N. Trefethen and D. Bau, "Numerical Linear Algebra," Society for Industrial and Applied Mathematics, 1997.
[31] A. Kale, A. Sundaresan, A.N. Rajagopalan, N.P. Cuntoor, A.K. Roy-Chowdhury, V. Kruger, and R. Chellappa, "Identification of Humans using Gait," *T-IP,* vol.13, no.9, pp.1163-1173, 2004.