Boosting Discriminant Learners for Gait Recognition using MPCA Features

Haiping Lu*, K.N. Plataniotis* and A.N. Venetsanopoulos†

*The Edward S. Rogers Sr. Department of Electrical and Computer Engineering

University of Toronto, M5S 3G4, Canada

†Ryerson University, Toronto, Ontario, M5B 2K3, Canada

{haiping, kostas, anv}@comm.toronto.edu

Abstract

This paper proposes a boosted linear discriminant analysis (LDA) solution on features extracted by the multilinear principal component analysis (MPCA) to enhance gait recognition performance. Three dimensional gait objects are projected in the MPCA space first to obtain low-dimensional tensorial features. Then, lower-dimensional vectorial features are obtained through discriminative feature selection. These feature vectors are then fed into a LDA-style booster, where several regularized and weakened LDA learners work together to produce a strong learner through a novel feature weighting and sampling process. The LDA learner employs a simple nearest-neighbor classifier with a weighted angle distance measure for classification. The experimental results on the NIST/USF "Gait Challenge" data sets shows that the proposed solution has successfully improved the gait recognition performance and outperformed several state-of-the-art gait recognition algorithms.

Index Terms

Multilinear principal component analysis, linear discriminant analysis, boosting, gait recognition, biometrics

A preliminary version of this paper was presented at the 2007 Biometrics Symposium, Baltimore, Maryland, September 11-13, 2007.

1 INTRODUCTION

Automated human identification at a distance is important in visual surveillance and monitoring applications in security-sensitive environments such as airports, banks, shopping malls, parking lots and large civic structures [1], [2]. However, many conventional biometrics, such as iris, face and fingerprint, require the person to be recognized to be in close distance or even in contact with the capturing device. At a distance, these biometrics are usually not available in high enough resolution for recognition purposes.

Gait, the style of walking of an individual, is an emerging behavioral biometric that offers the potential for vision-based recognition at a distance [3]–[6]. In 1975 [7], Johansson used point light displays to show humans' ability to distinguish human locomotion from other motion patterns. Later, experiments demonstrate the capability of identifying familiar individuals or the gender of a person [8], [9]. Nonetheless, research on gait recognition from video sequences are only receiving significant attentions recently. Vision-based gait recognition is particularly attractive in human identification at a distance because gait capture is unobtrusive, requiring no cooperation or attention of the observed subject, and gait is difficult to hide [5], [10].

There are two approaches to gait recognition: the model-based approach [11]–[13], where human body structure is explicitly modeled, and the appearance-based approach [5], [6], [10], [14]–[18], where gait is treated as a sequence of holistic binary patterns (silhouettes)[1]. The appearance-based approach has been more successful working on practical data [10]. Appearance-based approaches take binary gait silhouette sequences extracted from raw gait sequences [19]–[21] as the input. These gait silhouette sequences are naturally three-dimensional objects, also called third-order tensors, and the three dimensions are the spatial row, column and the temporal modes [22]. These tensor objects are in a very high-dimensional tensor space. To apply traditional linear feature extraction algorithms such as the Principal Component Analysis (PCA) and Linear Discriminant Analysis (LDA) on these tensorial data, they need to be first reshaped (vectorized) into vectors in a very high-dimensional space. This reshaping not only results in high computation and memory demand, but also breaks the structure and correlation in the original data. This problem has motivated the development of multilinear subspace learning algorithms operating

1. It should be noted that although the EigenGait approach [18] makes use of silhouettes as well, a motion-based recognition approach is taken where features are extracted from the image self-similarity plots rather than from the silhouettes directly.

directly on the gait sequences in their tensorial representation rather than their vectorized forms. In particular, the multilinear PCA (MPCA) algorithm [22] aims to determine a multilinear projection that projects the original tensor objects into a lower-dimensional tensor subspace while preserving the variation in the original data as much as possible. For gait recognition, a number of discriminative features in the projected tensor space can be selected. The MPCA-based gait recognition algorithm has achieved better overall performance when compared with the state-of-the-art gait recognition algorithms.

Although progresses have been made in gait recognition, it remains a very challenging problem. A person's gait can be affected by many factors, such as viewing angles, walking surfaces and shoes. Similar to face patterns, the distribution of gait patterns is expected to be nonlinear and complex. Furthermore, the gait data in training and those in testing may be captured under different conditions and this makes generalization very difficult, as studied in the Gait Challenge problem [10]. There are many methods proposed in the literature to handle complex and nonlinear patterns. The ensemble-based machine learning method named boosting is a very promising one that offers good generalization capability. Traditional boosting design works through the combination of a set of weak classifiers repeatedly trained on weighted training samples [23], [24], which tends to be an adaptive feature selection process [25]. Feature extraction is not a concern in these boosters and the requirement of an appropriate weak learner in boosting has restricted its applicability [24], [26]. A recent work in [27] has broken this limitation by proposing a boosting algorithm that puts the learning focus on the feature extractor rather than the classifier so that the new boosting scheme works with LDA-style learners. The effectiveness of the boosting scheme proposed in [27] has been demonstrated on the problem of face recognition. A cross-validation mechanism is employed to weaken the LDA learner and the pairwise class discriminant distribution (PCDD) is introduced for interaction between the booster and the learner.

In this paper, the boosting work in [27] is enhanced and extended so it can be successfully applied to the problem of gait recognition through combination with the recent development of MPCA [22]. It should be noted that, to the best of the authors' knowledge, this is the first work that has applied boosting to gait recognition, although boosting has been well studied for face recognition [27]–[29]. In the proposed processing scheme, MPCA [22] first produces EigenTensorGaits (ETGs) in a lower-dimensional tensor space and then only a number of discriminative ETGs are selected as the input to the LDA-based booster. There are two main advantages in this scheme. On one hand, the MPCA feature extractor applied before the booster reduces the processing cost greatly (in both training and testing) such that the very-high dimensional tensorial gait data can be handled efficiently. On the other hand, the number of selected ETGs provides another way (in addition to the cross-validation mechanism in [27]) to control the weakness of the LDA learner. In addition, in order to improve the generalization performance further, a regularization mechanism is incorporated since the within-class scatter of gait patterns under the capturing conditions in testing is expected to be larger than that of gait patterns in training. Furthermore, the training sample selection scheme in the original LDA-style boosting scheme proposed in [27] tends to prevent the inclusion of "difficult" (hard to classify correctly) samples in subsequent boosting steps. Therefore, a new training sample selection method is introduced in this paper to include more "difficult" samples in subsequent boosting steps to get better boosting results.

The rest of the paper is organized as follows. Section 2 briefly reviews the MPCA-based gait feature extraction method introduced in [22]. Section 3 proposes the LDA-based boosting algorithm operating on MPCA features for enhancing gait recognition performance. In Section 4, experimental results on the NIST/USF "Gait Challenge" data sets are presented and the proposed algorithm is compared with the state-of-the-art gait recognition algorithms to illustrate the effectiveness of the proposed solution. Finally, conclusions are drawn in Section 5.

2 Review of MPCA-based gait feature extraction

MPCA [22] is a multilinear subspace learning method that extracts features directly from tensorial representation of multi-dimensional objects. In this section, the notations are introduced and the MPCA-based gait feature extraction algorithm is briefly reviewed.

2.1 Notations

In this paper, vectors are denoted by lowercase boldface letters, e.g., $\mathbf{x}$; matrices by uppercase boldface, e.g., $\mathbf{U}$; and tensors by calligraphic letters, e.g., $\mathcal{A}$. Their elements are denoted with indices in brackets. Indices are denoted by lowercase letters and span the range from 1 to the uppercase letter of the index, e.g., $n = 1, 2, ..., N$. An N^{th}-order tensor is denoted as $\mathcal{A} \in \mathbb{R}^{I_1 \times I_2 \times ... \times I_N}$. It is addressed by N indices i_n, $n = 1, ..., N$, and each i_n addresses the n-mode of $\mathcal{A}$. The n-mode product of a tensor $\mathcal{A}$ by a matrix $\mathbf{U} \in \mathbb{R}^{J_n \times I_n}$, denoted by $\mathcal{A} \times_n \mathbf{U}$, is a tensor with entries [30]–[32]:

$$(\mathcal{A} \times_n \mathbf{U})(i_1, ..., i_{n-1}, j_n, i_{n+1}, ..., i_N) = \sum_{i_n} \mathcal{A}(i_1, ..., i_N) \cdot \mathbf{U}(j_n, i_n). \tag{1}$$

A rank-1 tensor $\mathcal{A}$ equals to the outer product of N vectors:

$$\mathcal{A} = \mathbf{u}^{(1)} \circ \mathbf{u}^{(2)} \circ ... \circ \mathbf{u}^{(N)}, \tag{2}$$

which means that

$$\mathcal{A}(i_1, i_2, ..., i_N) = \mathbf{u}^{(1)}(i_1) \cdot \mathbf{u}^{(2)}(i_2) \cdot ... \cdot \mathbf{u}^{(N)}(i_N) \tag{3}$$

for all values of indices [22], [30], [32].

2.2 Gait feature extraction through MPCA

In the MPCA-based gait feature extraction algorithm proposed in [22], a gait sample is a half cycle of gait silhouette sequences[2], represented naturally as a third-order tensor. The procedures described in [22] are followed to obtain these gait samples, where the foreground pixels in the lower-half of the silhouettes are counted and the minimums of the foreground pixel number sequence partition a gait sequence into half cycles. There are two types of gait data sets in a typical gait recognition problem: the gallery and the probe [10]. Gait samples in the gallery set are labeled with their identities and they are used as training data, while the probe set contains the test data, which are gait samples of unknown identities that need to be matched against those included in the gallery set.

The input to MPCA are third-order gallery gait samples $\{\mathcal{X}_1, ..., \mathcal{X}_M \in \mathbb{R}^{I_1 \times I_2 \times I_3}\}$, where M is the total number of gait samples in the gallery set. The MPCA algorithm solves for a multilinear transformation

$$\{ \tilde{\mathbf{U}}^{(n)} \in \mathbb{R}^{I_n \times P_n}, n = 1, 2, 3\}, \tag{4}$$

2. Another choice is to use full cycles as gait samples, which results in larger sample size in the time mode but fewer samples available for both training and test, while asymmetry between two adjacent half cycles could be potentially useful for discrimination in this case. In addition, half cycles may not always be an appropriate choice for gait samples. For example, when a luggage is carried on one side, full cycles are more appropriate to be used as gait samples. Thus, it will be worthwhile to study the effects of this choice on the gait recognition performance. However, this issue is out of the scope of this paper and it is left for future works since this paper focuses on the incorporation of the boosting scheme in gait recognition.

where $P_n < I_n$ for $n = 1, 2, 3$, that maps the original gait tensor space $\mathbb{R}^{I_1} \bigotimes \mathbb{R}^{I_2} \bigotimes \mathbb{R}^{I_3}$ into a lower-dimensional tensor subspace $\mathbb{R}^{P_1} \bigotimes \mathbb{R}^{P_2} \bigotimes \mathbb{R}^{P_3}$:

$$\mathcal{Y}_m = \mathcal{X}_m \times_1 \tilde{\mathbf{U}}^{(1)^T} \times_2 \tilde{\mathbf{U}}^{(2)^T} \times_3 \tilde{\mathbf{U}}^{(3)^T}, m = 1, ..., M, \tag{5}$$

such that the total tensor scatter

$$\Psi_{\mathcal{Y}} = \sum_{m=1}^{M} \| \mathcal{Y}_m - \bar{\mathcal{Y}} \|_F^2, \tag{6}$$

is maximized, where $\bar{\mathcal{Y}} = \frac{1}{M}\sum_{m=1}^{M} \mathcal{Y}_m$ is the average of the training samples [22], i.e., the mean sample. This MPCA problem is solved through an iterative alternating projection method in [22].

The MPCA projection matrices $\{\tilde{\mathbf{U}}^{(n)}, n = 1, 2, 3\}$ can be viewed as $\prod_{n=1}^{3} P_n$ so-called EigenTensorGaits (ETGs) [22]:

$$\tilde{\mathcal{U}}_{p_1 p_2 p_3} = \tilde{\mathbf{u}}_{p_1}^{(1)} \circ \tilde{\mathbf{u}}_{p_2}^{(2)} \circ \tilde{\mathbf{u}}_{p_3}^{(3)}, \tag{7}$$

where $\tilde{\mathbf{u}}_{p_n}^{(n)}$ is the p_n^{th} column of $\tilde{\mathbf{U}}^{(n)}$. Since MPCA is unsupervised and the class information is not considered in feature extraction, not all ETGs are useful for recognition purposes. Therefore, in [22], a number of discriminative ETGs are selected according to their class discriminability $\Gamma_{p_1 p_2 p_3}$, where $\Gamma_{p_1 p_2 p_3}$ for the eigentensor $\tilde{\mathcal{U}}_{p_1 p_2 p_3}$ is defined as

$$\Gamma_{p_1 p_2 p_3} = \frac{\sum_{c=1}^{C} M_c \cdot \left[\bar{\mathcal{Y}}_c(p_1, p_2, p_3) - \bar{\mathcal{Y}}(p_1, p_2, p_3)\right]^2}{\sum_{m=1}^{M} \left[\mathcal{Y}_m(p_1, p_2, p_3) - \bar{\mathcal{Y}}_{c_m}(p_1, p_2, p_3)\right]^2}, \tag{8}$$

where C denotes the number of classes (subjects), M_c denotes the number of gait samples for class (subject) c, and c_m denotes the class label for the m^{th} gallery gait sample $\mathcal{X}_m$. $\mathcal{Y}_m$ is the feature tensor of $\mathcal{X}_m$ in the projected MPCA subspace, the mean feature tensor $\bar{\mathcal{Y}} = \frac{1}{M}\sum_m \mathcal{Y}_m$ and the class mean feature tensor $\bar{\mathcal{Y}}_c = \frac{1}{M_c}\sum_{m, c_m = c} \mathcal{Y}_m$. For the ETG selection, the entries in $\mathcal{Y}_m$ are arranged into a feature vector $\mathbf{y}_m$ according to $\Gamma_{p_1 p_2 p_3}$ in descending order. Only the first $H_{\mathbf{y}}$ entries of $\mathbf{y}_m$ are kept for subsequent recognition task [22]. It should be noted that discriminability is only considered in the ETG selection process, while the selected ETG features are extracted in an unsupervised way by MPCA.

3 Boosting LDA-style learners on MPCA features for gait recognition

Figure 1 illustrates the proposed new gait recognition method. Input tensorial gait samples are projected through MPCA and a number of discriminative EigenTensorGaits (ETGs) are selected to obtain gait feature vectors as in [22]. The extracted discriminative feature vectors are then fed into a new LDA-based booster built upon the booster in [27] for learning and classification. This section presents the proposed LDA-based booster in detail.

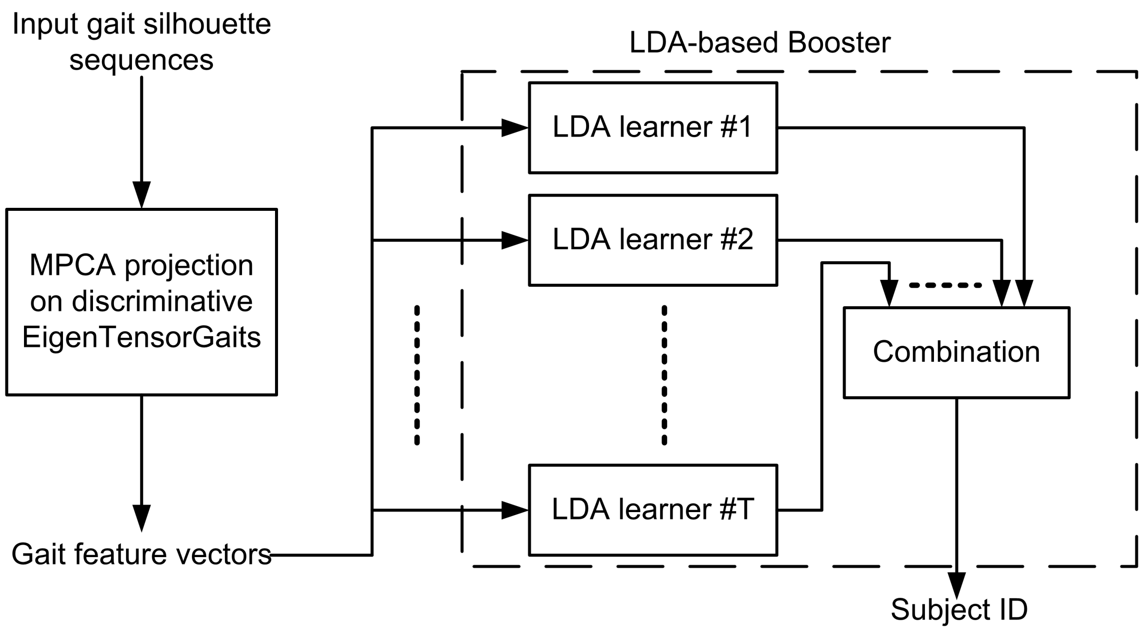


Fig. 1. Illustration of gait recognition through LDA-based boosting on MPCA features.

3.1 The boosting scheme

The pseudo-code implementation of the proposed MPCA+boosting scheme is summarized in Algorithm 1. As in [27], the AdaBoost.M2 algorithm [23] is followed here. AdaBoost.M2 aims to extend the communication between the boosting algorithm and the weak learner by allowing the weak learner to generate more expressive hypotheses (a set of "plausible" labels rather than a single label) indicating a "degree of plausibility", i.e., a hypothesis h takes a sample $\mathbf{y}$ and a class label c as the inputs and produces a "plausibility" score $h(\mathbf{y}, c) \in [0, 1]$ as the output. To achieve its objective, the AdaBoost.M2 introduces a sophisticated error measure pseudo-loss $\hat{\epsilon}_t$ with respect to the mislabel distribution $\mathbf{D}_t(m, c)$ in [23]. Thus, $\mathbf{D}_t$ is an $M \times C$ matrix. A mislabel is a pair (m, c), where m is the index of a training sample and c is an incorrect label associated with the sample $\mathbf{y}_m$. Let B be the set of all mislabels:

$$B = \{(m, c) : m = 1, ..., M, c \neq c_m\}. \tag{9}$$

The mislabel distribution is initialized as

$$\mathbf{D}_1(m,c) = \frac{1}{M \cdot (C-1)} \tag{10}$$

for $(m,c) \in B$. Accordingly, the weak learner produces a hypothesis

$$h_t : \mathbb{R}^{H_{\mathbf{y}}} \times C \to [0,1], \tag{11}$$

where $h(\mathbf{y},c)$ measures the degree to which it is believed that c is the correct label for $\mathbf{y}$. The pseudo-loss $\hat{\epsilon}_t$ of the hypothesis h_t with respect to $\mathbf{D}_t(m,c)$ is defined to measure the goodness of h_t and it is given by [23]:

$$\hat{\epsilon}_t = \frac{1}{2} \sum_{(m,c)\in B} \mathbf{D}_t(m,c) \left(1 - h_t(\mathbf{y}_m, c_m) + h_t(\mathbf{y}_m, c)\right). \tag{12}$$

The introduction of the mislabel distribution enhances the communication between the learner and the booster, so that the AdaBoost.M2 can focus the weak learner not only on hard-to-classify samples, but also on the incorrect labels that are the hardest to discriminate [23].

Another distribution $\mathbf{d}_t(m)$, named as pseudo sample distribution in [27], is derived from $\mathbf{D}_t(m,c)$ as

$$\mathbf{d}_t(m) = \sum_{c \neq c_m} \mathbf{D}_t(m,c). \tag{13}$$

Thus, $\mathbf{d}_t$ is an $M \times 1$ vector.

For the communication between the booster and the learner, the following "pairwise class discriminant distribution" (PCDD) $\hat{\mathbf{A}}_t \in \mathbb{R}^{C \times C}$ was introduced in [27]

$$\hat{\mathbf{A}}_t(c_a, c_b) = \frac{1}{2} \left(\sum_{m: c_m = c_a, c_{m_t} = c_b} \mathbf{d}_t(m) + \sum_{m: c_m = c_b, c_{m_t} = c_a} \mathbf{d}_t(m) \right), \tag{14}$$

where

$$c_{m_t} = \arg\max_c h_t(\mathbf{y}_m, c) \tag{15}$$

and the diagonal of $\hat{\mathbf{A}}_t$ is set to zeros.

3.2 The LDA learner

In building the LDA learner, the approach in [27] is adopted with several enhancement:

1) In [27], R samples per class are used as the input to the LDA learner in order to get weaker but more diverse LDA learners. R random samples per class are taken for the first boosting step and the hardest R (with the largest $\mathbf{d}(m)$) samples per class are selected for

Algorithm 1 The pseudo-code implementation of the LDA-based booster.

Input: The gallery gait feature vectors $\{\mathbf{y}_m, m = 1, ..., M\}$ with class labels $\mathbf{c} \in \mathbb{R}^M$, the LDA learner described in Sec. 3.2, the number of samples for LDA training S, the maximum number of iterations T.

Algorithm:

i. Initialize $\mathbf{D}_1(m, c) = \frac{1}{M(C-1)}$, $\hat{\mathbf{A}}_1(c_a, c_b) = \frac{1}{C^2}$, $\mathbf{D}_1(m, c_m) = 0$, $\hat{\mathbf{A}}_1(c_a, c_a) = 0$, and S samples are selected to form the initial training set $\{\mathbf{y}_s, s = 1, ..., S\}_1$, with the first $\lfloor S/C \rfloor$ or $\lceil S/C \rceil$ samples from each class, where $\lfloor \cdot \rfloor$ and $\lceil \cdot \rceil$ are the floor and ceil functions, respectively.

ii. Do for t=1:T:

1) Get $\mathbf{V}_t$ from $\mathbf{S}_{B_t}$ and $\mathbf{S}_{W_t}$ constructed from $\{\mathbf{y}_s, s = 1, ..., S\}_t$ and project $\{\mathbf{y}_m\}$ to $\{\mathbf{z}_m\}$.
2) Get hypothesis $\{h_t(\mathbf{y}_m, c) \in [0, 1]\}$ by applying the nearest neighbor classifier with the MAD measure [22] on $\{\mathbf{z}_m\}$.
3) Calculate ϵ_t, the pseudo-loss of h_t, from (12).
4) Set $\beta_t = \epsilon_t/(1 - \epsilon_t)$.
5) Update $\mathbf{D}_t$:

$$\mathbf{D}_{t+1}(m, c) = \mathbf{D}_t(m, c)\beta_t^{\frac{1}{2}(1+h_t(\mathbf{y}_m, c_m)-h_t(\mathbf{y}_m, c))},$$

and normalize it:

$$\mathbf{D}_{t+1}(m, c) = \frac{\mathbf{D}_{t+1}(m, c)}{\sum_m \sum_c \mathbf{D}_{t+1}(m, c)}.$$

6) Update $\mathbf{d}_{t+1}(m)$, $\hat{\mathbf{A}}_{t+1}$ and $\{\mathbf{y}_s\}_{t+1}$ accordingly.

Output: The final hypothesis:

$$h_{fin}(\mathbf{y}) = \arg\max_c \sum_{t=1}^{T} \left(\log \frac{1}{\beta_t}\right) h_t(\mathbf{y}, c)$$

subsequent steps. Let $\{\mathbf{y}_s, s = 1, ..., S\}$ denote the selected samples, where $S = R \times C$ for the sample selection scheme in [27].

This sample selection scheme has a problem that is against the design principal of boosting, where the hardest samples should be selected for subsequent learning. Figure 8(a) gives a typical example of the average weights of the samples selected according to the method in [27], denoted as “OldSel”. From the figure, the hardness of the selected samples (i.e., the weight) decreases quickly after the first a few boosting steps. The reason is that it is not true in most cases that the hardest samples will be distributed evenly among all the classes. In fact, it is common that some classes have more hard samples while some classes have more easy samples. Therefore, in this paper, a new sample selection scheme is proposed to enhance the sample selection process in boosting, as described below:

a) For each class, select the hardest sample to result in C samples added to the pool of training sample for subsequent learning.
b) Select the hardest $S - C$ samples among all the rest samples, regardless of their class labels so that together with the C samples selected in the previous step, S samples are chosen for subsequent boosting.

The average weights of the samples selected according to the sample selection scheme proposed above are shown in Figure 8(a) as well, denoted as “NewSel”. As seen from the figure, the new sample selection scheme results in samples with much larger weights selected compared to the scheme in [27].

2) For the between-class scatter matrix $\mathbf{S}_B$, the pairwise between-class scatter in [33] is used instead of that used in [27] in this paper for its simplicity and easy computation:

$$\mathbf{S}_B = \sum_{c_a=1}^{C-1} \frac{M_{c_a}}{M} \sum_{c_b=c_a+1}^{C} \frac{M_{c_b}}{M} \tilde{\mathbf{A}}_t(c_a, c_b)(\bar{\mathbf{y}}_{c_a} - \bar{\mathbf{y}}_{c_b})(\bar{\mathbf{y}}_{c_a} - \bar{\mathbf{y}}_{c_b})^T, \tag{16}$$

where

$$\bar{\mathbf{y}}_c = \frac{1}{M_c} \sum_{s}^{c_s=c} \mathbf{y}_s \tag{17}$$

is the mean for class c.

3) For the within-class scatter matrix, a regularized version of that in [27] is used:

$$\mathbf{S}_W = \sum_s \mathbf{d}(s)(\mathbf{y}_s - \bar{\mathbf{y}}_{c_s})(\mathbf{y}_s - \bar{\mathbf{y}}_{c_s})^T + \eta \cdot \mathbf{I}_{H_\mathbf{y}}, \tag{18}$$

where η is a regularization parameter to increase the estimated within-class scatter and $\mathbf{I}_{H_\mathbf{y}}$ is an identity matrix of size $H_\mathbf{y} \times H_\mathbf{y}$. The regularization term is added because in the gait recognition problem, the actual within-class scatter of gait sequences captured under various conditions is expected to be greater than the within-class scatter that can be estimated from the gallery set, which is captured under a single condition.

With these definitions, the projection $\mathbf{V}$ is then to be solved to maximize the ratio of the between-class scatter to the within-class scatter. The solution is

$$\mathbf{V} = \arg\max_{\mathbf{V}} \frac{|\mathbf{V}^T \mathbf{S}_B \mathbf{V}|}{|\mathbf{V}^T \mathbf{S}_W \mathbf{V}|} = [\mathbf{v}_1 \ \mathbf{v}_2 \ ... \ \mathbf{v}_{H_\mathbf{z}}], \tag{19}$$

where $\{\mathbf{v}_{h_\mathbf{z}}, h_\mathbf{z} = 1, ..., H_\mathbf{z}\}$ is the set of generalized eigenvalues of $\mathbf{S}_B$ and $\mathbf{S}_W$ corresponding to the $H_\mathbf{z}$ ($\leq C-1$) largest generalized eigenvalues $\{ \lambda_{h_\mathbf{z}}, \ h_\mathbf{z} = 1, ..., H_\mathbf{z}\}$:

$$\mathbf{S}_B \mathbf{v}_{h_\mathbf{z}} = \lambda_{h_\mathbf{z}} \mathbf{S}_W \mathbf{v}_{h_\mathbf{z}}, h_\mathbf{z} = 1, ..., H_\mathbf{z}. \tag{20}$$

Thus, the LDA feature vector $\mathbf{z}_m$ is obtained as $\mathbf{z}_m = \mathbf{V}^T \mathbf{y}_m$ for the input to a classifier.

The nearest neighbor classifier (NNC), which assigns label c to the test sample $\mathbf{y}$ if c is the class label of the sample nearest to $\mathbf{y}$, is used with the modified angle distance (MAD) measure defined in [22], which is found to have the best performance for MPCA-based algorithms in gait recognition. The MAD between two feature vectors $\mathbf{z}_a$ and $\mathbf{z}_b$ is calculated as

$$d(\mathbf{z}_a, \mathbf{z}_b) = \frac{-\sum_{h=1}^{H_\mathbf{z}} \mathbf{z}_a(h_\mathbf{z}) \cdot \mathbf{z}_b(h_\mathbf{z})}{\sqrt{\lambda_{h_\mathbf{z}} \sum_{h_\mathbf{z}=1}^{H_\mathbf{z}} \mathbf{z}_a(h_\mathbf{z})^2 \sum_{h_\mathbf{z}=1}^{H_\mathbf{z}} \mathbf{z}_b(h_\mathbf{z})^2}}. \tag{21}$$

The calculated distances between a sample and the C class means are matched to the interval $[0, 1]$ as required by the AdaBoost.M2 algorithm.

3.3 Discussions

It should be noted that besides the algorithmic difference, the proposed solution has an important difference in design with that in [27]. Direct application of the algorithm in [27] on the gait recognition problem requires the vectorization of the tensorial input $\{\mathcal{X}_m\}$ to $\{\mathbf{x}_m\}$. For a gait sample of typical size $128 \times 88 \times 20$, the resulted vectors are of size $225,280 \times 1$. In contrast, the LDA-based learners in the proposed booster take the gait feature vectors extracted by MPCA $\{\mathbf{y}_m \in \mathbb{R}^{H_\mathbf{y}}, m = 1, ..., M\}$, rather than the original data $\{ \mathbf{x}_m \in \mathbb{R}^{I_1 \times I_2 \times I_3}, \ m = 1, ..., M\}$. The proposed scheme has two benefits:

1) The number of selected discriminative ETGs, which is the gait feature vector dimension $H_\mathbf{y}$, gives us one more degree (besides the number of samples used for LDA learners) to control the weakness of the LDA learners. Similar to the case of PCA+LDA, where the recognition performance is often affected by the number of principal components for input to LDA, $H_\mathbf{y}$ affects the recognition performance of LDA on the MPCA features as well, as observed in [22]. Therefore, by choosing a value of $H_\mathbf{y}$ that is not optimal for a single

LDA learner, the obtained LDA learner is weakened. On the other hand, the LDA learner cannot be made "too weak" either. Otherwise, the boosting scheme will not work.

2) Using feature vectors of dimensionality $H_{\mathbf{y}}$ instead of the original high-dimensional data as the booster input is computationally advantageous. Since boosting is an iterative algorithm with T rounds, the computational cost is about T times of that of a single learner with the same input, both in training and testing. When the booster works on lower-dimensional features extracted by MPCA, it becomes much more efficient since it needs to deal with low-dimensional vectors only in each round. For instance, the dimension of the input vectors to the booster is around 200 in this paper, which is much smaller than the dimension $17,154$ for face data in [27] and the original gait data dimension $225,280$, Therefore, the computational cost is reduced significantly this way.

4 Experimental results

This section evaluates the proposed solution of LDA-based boosting on MPCA features (B-LDA-MPCA) through the following studies:

1) The comparison of gait recognition performance against the state-of-the-art gait recognition algorithms.
2) The effects of the gait feature vector dimension $H_{\mathbf{y}}$ for input to LDA learners, the LDA feature vector dimension $H_{\mathbf{z}}$, the number of LDA training samples for LDA learner input S, and the regularization parameter η on boosting recognition performance.
3) The effectiveness of the new sample selection scheme proposed in this paper in improving the booster performance.

4.1 The data sets

The NIST/USF "Gait Challenge" data sets version 1.7 [10], [34] is chosen to carry out the gait recognition experiments. All the recognition results reported and compared in this paper are obtained from this database. It consists of 452 sequences from 74 subjects walking in elliptical paths in front of the camera, with two viewpoints (left or right), two shoe types (A or B) and

TABLE 1

The characteristics of the gait data from the NIST/USF "Gait Challenge" data set version 1.7.

Gait data set	Number of sequences	Difference from the gallery
Gallery (GAR)	71	-
A (GAL)	71	View
B(GBR)	41	Shoe
C(GBL)	41	Shoe, view
D(CAR)	70	Surface
E(CBR)	44	Surface, shoe
F(CAL)	70	Surface, view
G(CBL)	44	Surface, shoe, view

two surface types (grass or concrete)[3]. The gallery set contains 71 sequences (one sequence from each subject). Seven experiments, corresponding to seven probe sets, are designed for human identification. Subjects are unique in the gallery and each probe set. There are no common sequences between the gallery set and any of the probe sets. Also, all the probe sets are distinct. The number of sequences in each probe set and the difference from the gallery set are summarized in Table 1. The capturing condition for each probe set is indicated in brackets after the probe name, where C, G, A, B, L, R, stand for cement surface, grass surface, shoe type A, shoe type B, left view, and right view, respectively.

The procedures described in [22] are used to obtain the third-order tensorial gait samples $\{\mathcal{X}_m\}$ of size $128 \times 88 \times 20$, with five examples shown as unfolded images in Fig. 2. There are 731 gait samples in the Gallery set and each subject has an average of roughly 10 samples available. MPCA is applied to get the gait feature vectors $\{\mathbf{y}_m\}$ for the input to the booster. The

3. There is a newer version 2.1 available, which is of much larger size with two additional differences in briefcase carrying condition and time (including clothing). Version 1.7 is chosen in this work because this version is widely used in the research community as well and the performance on it is far from saturated [4], [5], [14], [16], [17], [22]. In addition, version 1.7 is much smaller than version 2.1 so the computational demand is much lower in experimental evaluation.

rank 1 and rank 5 identification rates are used for recognition performance evaluation, where rank k results report the percentage of probe subjects whose true match in the gallery set was in the top k matches [10].

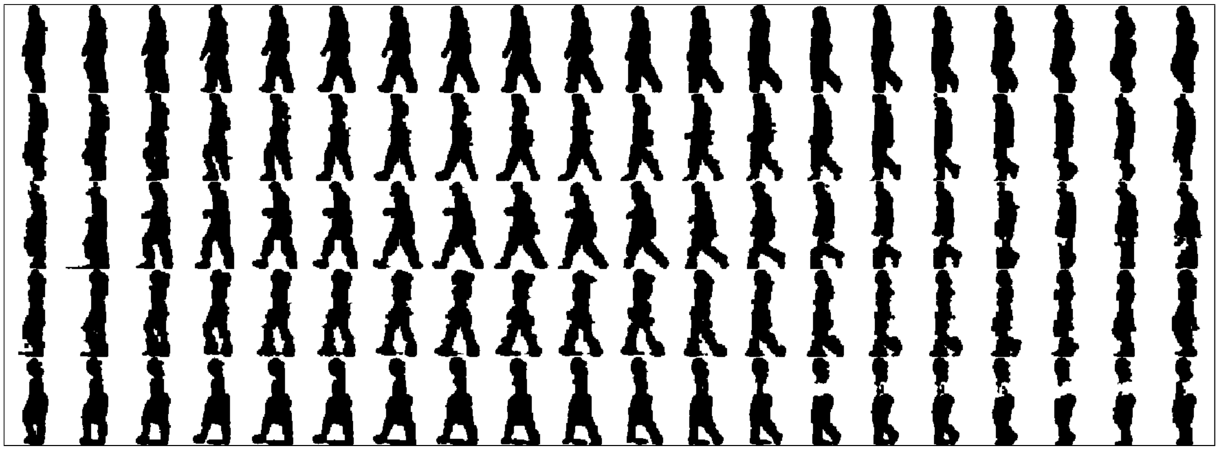


Fig. 2. Five gait silhouette samples shown by concatenating the frames.

4.2 Comparison of gait recognition results with the state-of-the-art algorithms

There are several parameters in the proposed B-LDA-MPCA solution: T, $H_{\mathbf{y}}$, $H_{\mathbf{z}}$, S and η. As in other boosting algorithms [27], the optimal determination of these parameters is still open problem to be solved. The maximum number of iterations is set to $T = 60$, as in [27]. Several values of the other four parameters are tested empirically, as listed below:

- $H_{\mathbf{y}}$: 160, 170, 180, 190, 200, 210, 220.
- $H_{\mathbf{z}}$: 35, 40, 45, 50, 55, 60, 65, 70.
- S: 142, 165, 188, 213, 234, 257, 284.
- η: 1, 10^{-1}, 10^{-2}, 10^{-3}, 10^{-4}.

For the proposed B-LDA-MPCA algorithm, the parameters producing the highest rank-1 identification rates averaged over all seven probes in the test are $H_{\mathbf{y}} = 200$, $H_{\mathbf{z}} = 55$, $S = 213$, and $\eta = 10^{-1}$, which are considered to be the best performing set of parameters in this paper. The evolutions of various identification rates over the boosting steps are shown in Fig. 3 with this set of parameters. In the figure legends, "R1" and "R5" stand for the average rank 1 and rank 5 identification rates, respectively. The identification rate obtained from the single learner in each step is denoted as 'Sgl' and that obtained from the aggregated learners is denoted by

'Bst'. From the figure, the effectiveness of the boosting scheme is observed. There are about 20% boost in both rank 1 and rank 5 identification rates through the combination of single learners.

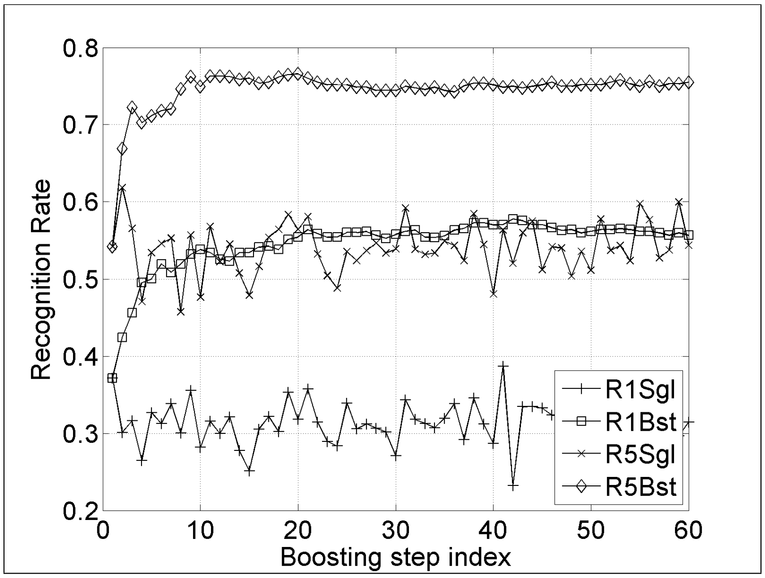


Fig. 3. The evolutions of various identification rates over the boosting steps with the best parameter set tested.

Tables 2 and 3 show the detailed rank 1 and rank 5 gait recognition results, respectively, of B-LDA-MPCA on each probe set, obtained with the best parameter settings tested. In addition, the tables also include the results from the baseline algorithm [10], [34] as well as the following state-of-the-art gait recognition algorithms: the Hidden Markov Model (HMM) framework [4], the linear time normalization (LTN) algorithm [16], the Gait Energy Image (GEI) algorithm [17], and the MPCA+LDA algorithm [22]. The average identification rates over the seven probes are also shown for comparison of the overall performance. The best results for each column are highlighted by boldface font in the table.

From the results, the B-LDA-MPCA algorithm has achieved the best rank 1 and rank 5 recognition results on all probes except the rank 1 identification rate on probe B and the rank 5 identification rate on probe D, demonstrating its superior recognition performance. Compared to the MPCA+LDA algorithm, the B-LDA-MPCA algorithm has improved the rank 1 identification rate by an average of 4% and the rank 5 identification rate by an average of 1%. The greatest improvement in rank 1 identification rate is 13% on probe F, and the greatest improvement in rank 5 identification rate is 10% on probe E. In particular, in rank 1 identification rates, the

TABLE 2

Comparison of the gait recognition results on the NIST/USF “Gait Challenge” data sets version 1.7: the rank 1 identification rate (%).

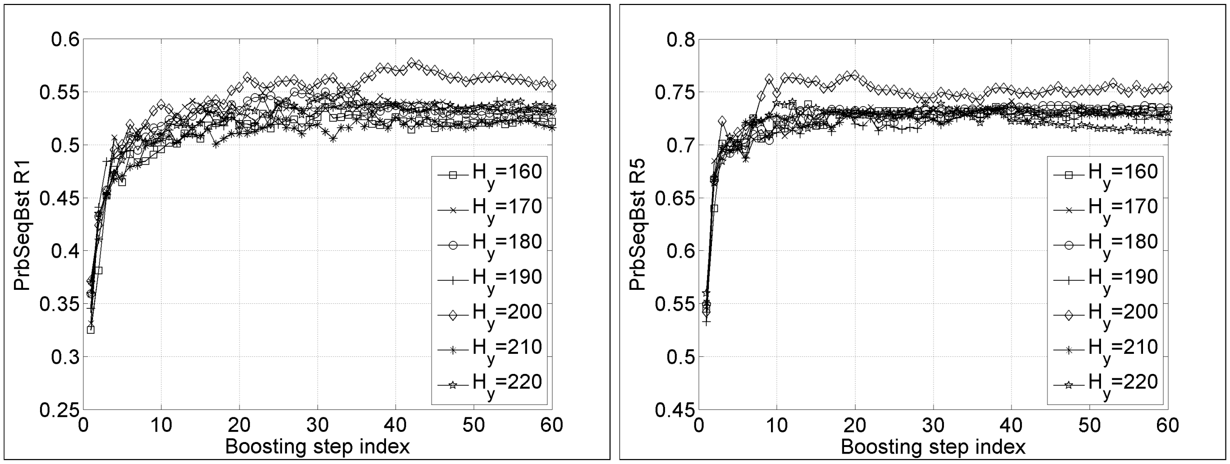
Probe	A	B	C	D	E	F	G	Average
Baseline	79	66	56	29	24	30	10	42
HMM	99	**89**	78	35	29	18	24	53
LTN	94	83	78	33	24	17	21	50
GEI	**100**	85	80	30	33	21	29	54
MPCA+LDA	99	88	**83**	36	29	21	21	54
B-LDA-MPCA	**100**	88	**83**	**39**	**34**	**34**	**30**	**58**

TABLE 3

Comparison of the gait recognition results on the NIST/USF “Gait Challenge” data sets version 1.7: the rank 5 identification rate (%).

Probe	A	B	C	D	E	F	G	Average
Baseline	96	81	76	61	55	46	33	64
HMM	**100**	90	90	65	65	60	50	74
LTN	99	85	83	65	67	58	48	72
GEI	**100**	85	88	55	55	41	48	67
MPCA+LDA	**100**	**93**	88	**71**	60	59	60	76
B-LDA-MPCA	**100**	**93**	**93**	67	**70**	**63**	**61**	**77**

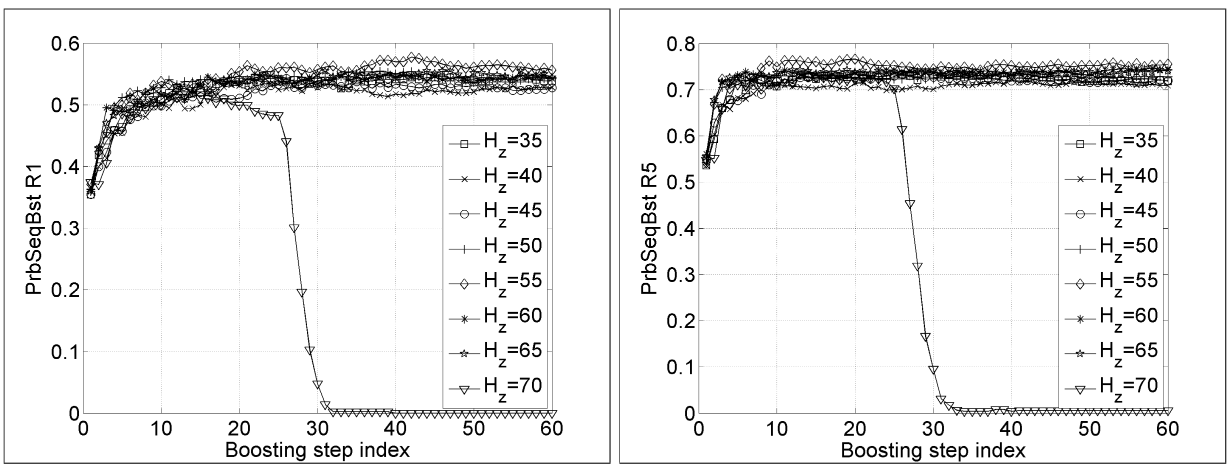
performance improvement on the more difficult probes, D, E, F and G, are more significant than the improvement on the easier probes, A, B and C, showing that the B-LDA-MPCA algorithm indeed generalizes better than the MPCA+LDA algorithm.



(a) Rank 1 identification rate.

(b) Rank 5 identification rate

Fig. 4. The effects of $H_{\mathbf{y}}$ on the gait recognition performance of the proposed method.



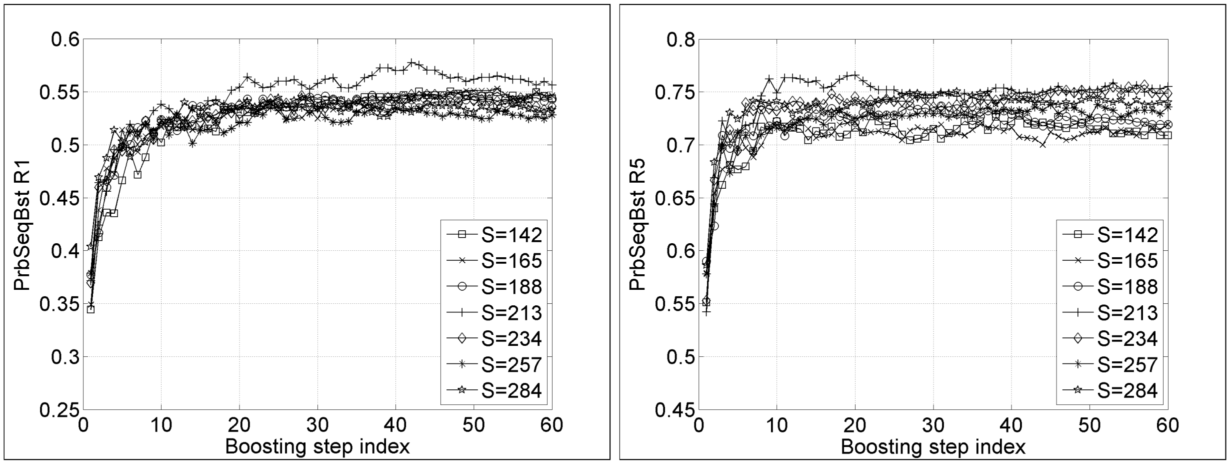
(a) Rank 1 identification rate.

(b) Rank 5 identification rate

Fig. 5. The effects of $H_{\mathbf{z}}$ on the gait recognition performance of the proposed method.

4.3 The effects of $H_{\mathbf{y}}$, $H_{\mathbf{z}}$, S and η on boosting

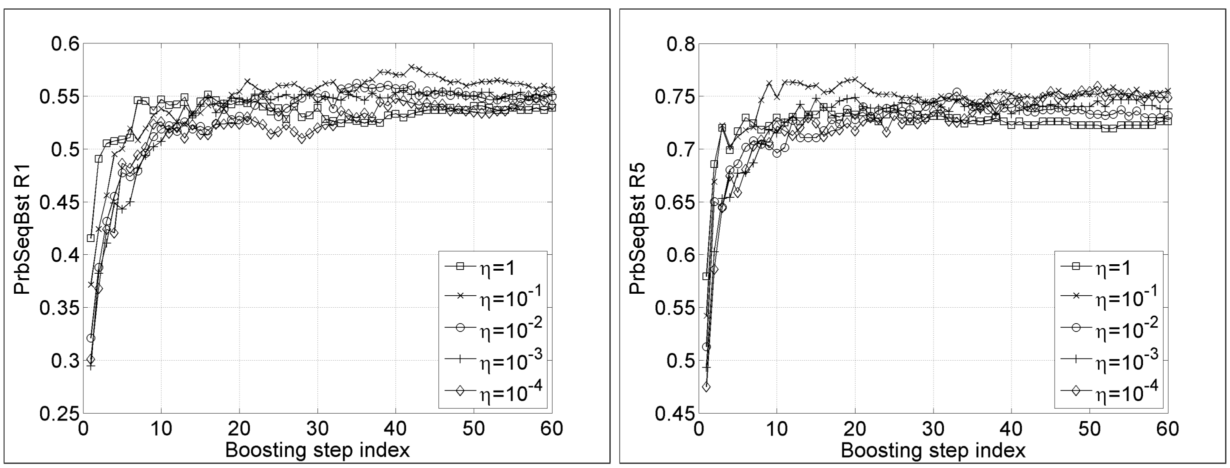
The effects of $H_{\mathbf{y}}$, $H_{\mathbf{z}}$, S and η on the average rank 1 and rank 5 gait recognition performance of the proposed method are shown in Figs. 4, 5, 6 and 7, respectively. The effects of a parameter are shown by varying only the parameter of interest while fixing all the other parameters since it is not possible to show the results of all possible parameter combinations. Only a number of values are tested for each parameter, as specified in Sec. 4.2. The fixed set of parameters is chosen to be the best set described in Sec. 4.2.



(a) Rank 1 identification rate.

(b) Rank 5 identification rate

Fig. 6. The effects of S on the gait recognition performance of the proposed method.



(a) Rank 1 identification rate.

(b) Rank 5 identification rate

Fig. 7. The effects of η on the gait recognition performance of the proposed method.

The proposed method introduces an additional learner weakness control mechanism by $H_{\mathbf{y}}$. From [22], $H_{\mathbf{y}} = 170$ gives the best gait recognition performance with the MAD measure and the NNC classifier. From Fig. 4, the weaker learners with $H_{\mathbf{y}} = 200$ give much better boosting results than the stronger learners with $H_{\mathbf{y}} = 170$. This confirms that $H_{\mathbf{y}}$ can improve the boosting performance through controlling the weakness of the learners.

The dimensionality of the LDA features $H_{\mathbf{z}}$ affects the recognition performance of the proposed solution as well. Since $C = 71$, the maximum dimensionality of the features extracted by

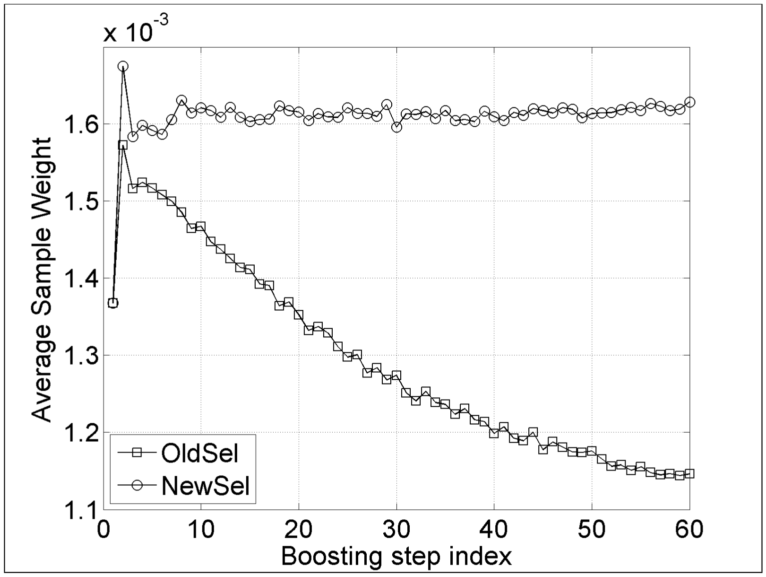
LDA learners is $C - 1 = 70$. Nonetheless, as pointed out in [27], if $H_{\mathbf{z}} = 70$, the resulted LDA learner will be very strong, deteriorating the performance of the booster. From Fig. 5, it can be seen that the value of $H_{\mathbf{z}}$ giving the best performance is a medium value. It is also evident from the figure that the strong learner with $H_{\mathbf{z}} = 70$ collapsed around the 30^{th} boosting step, as expected in boosting [27]. This set of experiments demonstrate that appropriate weakness is again required and the best boosting performance can not be reached with too strong or too weak learners.

The value of S determines the number of training samples for the LDA learners. As discussed in [27], the diversity of the LDA learners is necessary to ensure good boosting performance. Therefore, by choosing only a subset of the available training samples, the diversity among learners at different boosting steps is enhanced. On the other hand, S needs to be sufficiently large to enable learners to achieve a certain classification accuracy. Figure 6 illustrates the effects of S discussed here, showing that an appropriate choice of S is neither too small nor too large.

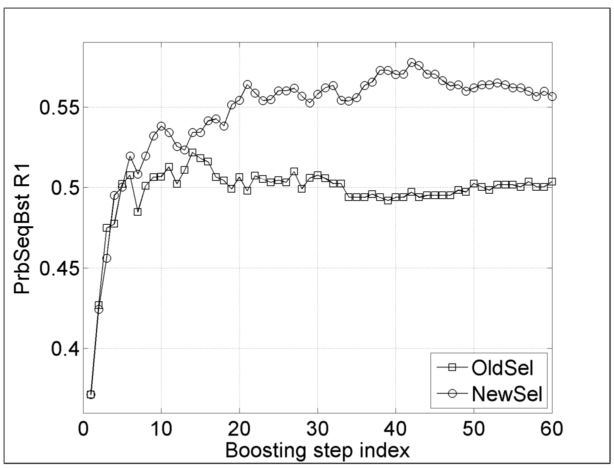
The effects of regularization are depicted in Fig. 7, where it is shown that an appropriate regularization parameter η does result in better generalization. This study confirms that gait recognizer can benefit from making use of the fact that the within-class scatter of gait patterns under various capturing conditions is greater than that under the same capturing condition.

4.4 The effectiveness of the proposed sample selection scheme for LDA learners in boosting the recognition performance

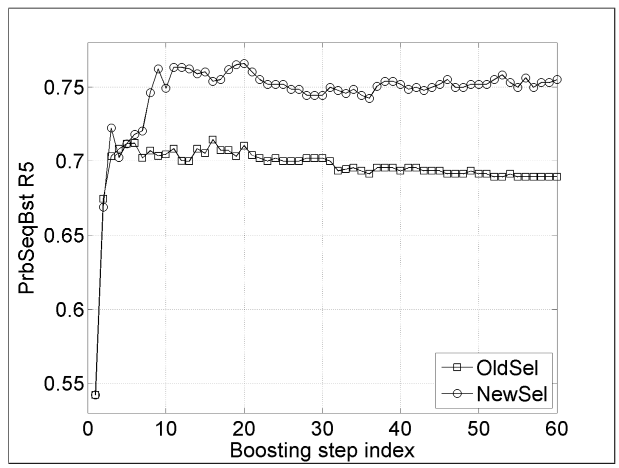
Figure 8 demonstrates the effectiveness of the new sample selection scheme proposed in Sec. 3.2. As discussed in Sec. 3.2 and illustrated in Fig. 8(a), the proposed scheme selects samples with much larger weights for subsequent boosting steps, compared with the scheme in [27]. Thus, the new scheme focuses more on the difficult samples, which agrees the working principle behind boosting. The effects of the new sample selection scheme on the recognition performance are shown in Figs. 8(b) and 8(c), where the corresponding average rank 1 and rank 5 identification rates are compared, respectively. From the figure, it can be seen that the proposed new sample selection scheme results in approximately 5% improvement in both rank 1 and rank 5 identification rates.



(a) The average weights of selected samples.



(b) The rank 1 identification rates



(c) The rank 5 identification rates

Fig. 8. Comparison of the proposed sample selection scheme against the sample selection scheme in [27].

5 CONCLUSIONS

This paper proposes a gait recognition solution through combining the MPCA algorithm [22] and the ensemble-based discriminant learning method in [27]. The MPCA algorithm in [22] is used to extract features from tensorial gait data and a subset of the extracted features are fed into an enhanced LDA-style booster. This scheme gives another way of learner weakness control in addition to computational efficiency. The LDA learner in [27] is modified by adopting a simpler weighted pairwise between-class scatter matrix and introducing a regularization term in the within-class scatter matrix so that the gait challenge due to various capturing conditions is taken into account. Furthermore, a new sample selection scheme of the LDA-based booster is proposed to concentrate more on the "difficult" samples in the boosting process. Experiments carried out on the gait challenge data sets show that the proposed scheme is effective in boosting the gait recognition performance and outperforms several state-of-the-art gait recognition algorithms.

ACKNOWLEDGMENTS

The authors would like to thank the anonymous reviewers for their insightful comments. The authors would also like to thank Prof. Sudeep Sarkar from the University of South Florida for kindly providing us with the Gait Challenge data sets. This work is partially supported by the Ontario Centres of Excellence through the Communications and Information Technology Ontario Partnership Program and the Bell University Labs - at the University of Toronto.

REFERENCES

[1] A. K. Jain, A. Ross, and S. Prabhakar, "An introduction to biometric recognition," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 14, no. 1, pp. 4–20, Jan. 2004.

[2] H. Lu, J. Wang, and K. N. Plataniotis, "A review on face and gait recognition: System, data and algorithms," in *Advanced Signal Processing: Theory and Implementation for Sonar, Radar, and Non-Invasive Medical Diagnostic Systems*, 2nd ed., S. Stergiopoulos, Ed. Boca Raton, Florida: CRC Press, 2009, pp. 303–330, ISBN: 978-1-4200-6238-0.

[3] M. S. Nixon and J. N. Carter, "Automatic recognition by gait," *Proceedings of the IEEE*, vol. 94, no. 11, pp. 2013–2024, Nov. 2006.

[4] A. Kale, A. N. Rajagopalan, A. Sunderesan, N. Cuntoor, A. Roy-Chowdhury, V. Krueger, and R. Chellappa, "Identification of humans using gait," *IEEE Transactions on Image Processing*, vol. 13, no. 9, pp. 1163–1173, Sep. 2004.

[5] N. V. Boulgouris, D. Hatzinakos, and K. N. Plataniotis, "Gait recognition: a challenging signal processing technology for biometrics," *IEEE Signal Processing Magazine*, vol. 22, no. 6, pp. 78–90, Nov. 2005.

[6] L. Wang, T. Tan, H. Ning, and W. Hu, "Silhouette analysis-based gait recognition for human identification," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 25, no. 12, pp. 1505–1518, Dec. 2003.

[7] G. Johansson, "Visual motion perception," *Science American*, vol. 232, no. 6, pp. 76–88, 1975.

[8] J. Cutting and L. Kozlowski, "Recognizing friends by their walk: Gait perception without familiarity cues," *Bulletin of the Psychonomic Society*, vol. 9, no. 5, pp. 353–356, 1977.

[9] C. Barclay, J. Cutting, and L. Kozlowski, "Temporal and spatial factors in gait perception that influence gender recognition," *Perception and Psychophysics*, vol. 23, no. 2, pp. 145–152, 1978.

[10] S. Sarkar, P. J. Phillips, Z. Liu, I. Robledo, P. Grother, and K. W. Bowyer, "The human ID gait challenge problem: Data sets, performance, and analysis," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27, no. 2, pp. 162–177, Feb. 2005.

[11] H. Lu, K. N. Plataniotis, and A. N. Venetsanopoulos, "A layered deformable model for gait analysis," in *Proc. IEEE International Conference on Automatic Face and Gesture Recognition*, Apr. 2006, pp. 249 – 254.

[12] C. Y. Yam, M. S. Nixon, and J. N. Carter, "Automated person recognition by walking and running via model-based approaches," *Pattern Recognition*, vol. 37, no. 5, pp. 1057–1072, May 2004.

[13] H. Lu, K. N. Plataniotis, and A. N. Venetsanopoulos, "A full-body layered deformable model for automatic model-based gait recognition," *EURASIP Journal on Advances in Signal Processing: Special Issue on Advanced Signal Processing and Pattern Recognition Methods for Biometrics*, vol. 2008, 2008, article ID 261317, 13 pages, doi:10.1155/2008/261317.

[14] H. Lu, K. N. Plataniotis, and A. N. Venetsanopoulos, "Uncorrelated multilinear discriminant analysis with regularization and aggregation for tensor object recognition," *IEEE Transactions on Neural Networks*, vol. 20, no. 1, pp. 103–123, Jan. 2009.

[15] H. Lu, K. N. Plataniotis, and A. N. Venetsanopoulos, "Multilinear principal component analysis of tensor objects for recognition," in *Proc. International Conference on Pattern Recognition*, vol. 2, August 2006, pp. 776 – 779.

[16] N. V. Boulgouris, K. N. Plataniotis, and D. Hatzinakos, "Gait recognition using linear time normalization," *Pattern Recognition*, vol. 39, no. 5, pp. 969–979, 2006.

[17] J. Han and B. Bhanu, "Individual recognition using gait energy image," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 28, no. 2, pp. 316–322, Feb. 2006.

[18] R. Cutler, C. Benabdelkader, and L. Davis, "Motion-based recognition of people in eigengait space," in *Proc. IEEE International Conference on Automatic Face and Gesture Recognition*, May 2002, pp. 254–259.

[19] H. Lu, K. N. Plataniotis, and A. N. Venetsanopoulos, "Coarse-to-fine pedestrian localization and silhouette extraction for the gait challenge data sets," in *Proc. IEEE Conference on Multimedia and Expo*, Jul. 2006, pp. 1009–1012.

[20] J. Migdal and W. E. L. Grimson, "Background subtraction using markov thresholds," in *IEEE Workshop on Motion and Video Computing*, January 2005.

[21] L. Lee, G. Dalley, and K. Tieu, "Learning pedestrian models for silhouette refinement," in *Proc. IEEE Conference on Computer Vision*, Oct. 2003, pp. 663–670.

[22] H. Lu, K. N. Plataniotis, and A. N. Venetsanopoulos, "MPCA: Multilinear principal component analysis of tensor objects," *IEEE Transactions on Neural Networks*, vol. 19, no. 1, pp. 18–39, Jan. 2008.

[23] Y. Freund and R. E. Schapire, "Experiments with a new boosting algorithm," in *Proc. the Thirteenth International Conference on Machine Learning*, 1996, pp. 148–156.

[24] R. E. Schapire, "The boosting approach to machine learning: An overview," in *MSRI Workshop on Nonlinear Estimation and Classification*, D. D. Denison, M. H. Hansen, C. Holmes, B. Mallick, and B. Yu, Eds. Springer, 2003. [Online]. Available: http://stat.haifa.ac.il/ goldensh/DM/msri.pdf

[25] P. Viola and M. J. Jones, "Robust real-time face detection," *International Journal of Computer Vision*, vol. 57, no. 2, pp. 137–154, May 2004.

[26] M. Skurichina and R. P. W. Duin, "Bagging, boosting and the random subspace method for linear classifiers," *Pattern Analysis & Applications*, vol. 5, no. 2, pp. 121–135, 2002.

[27] J. Lu, K. N. Plataniotis, A. N. Venetsanopoulos, and S. Z. Li, "Ensemble-based discriminant learning with boosting for face recognition," *IEEE Transactions on Neural Networks*, vol. 17, no. 1, pp. 166–178, Jan. 2006.

[28] P. Yang, S. Shan, W. Gao, S. Z. Li, and D. Zhang, “Face recognition using ada-boosted gabor features,” in *Proc. IEEE International Conference on Automatic Face and Gesture Recognition*, May 2004, pp. 356–361.

[29] S. Z. Li, R. Chu, S. Liao, and L. Zhang, “Illumination invariant face recognition using near-infrared images,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 29, no. 4, pp. 627–639, Apr. 2007.

[30] L. D. Lathauwer, B. D. Moor, and J. Vandewalle, “A multilinear singualr value decomposition,” *SIAM Journal of Matrix Analysis and Applications*, vol. 21, no. 4, pp. 1253–1278, 2000.

[31] B. W. Bader and T. G. Kolda, “Algorithm 862: Matlab tensor classes for fast algorithm prototyping,” *ACM Transactions on Mathematical Software*, vol. 32, no. 4, pp. 635–653, Dec. 2006.

[32] L. D. Lathauwer, B. D. Moor, and J. Vandewalle, “On the best rank-1 and rank-$(R_1, R_2, ..., R_N)$ approximation of higher-order tensors,” *SIAM Journal of Matrix Analysis and Applications*, vol. 21, no. 4, pp. 1324–1342, 2000.

[33] M. Loog, R. P. W. Duin, and R. Haeb-Umbach, “Multiclass linear dimension reduction by weighted pairwise fisher criteria,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 23, no. 7, pp. 762–766, Jul. 2001.

[34] P. J. Phillips, S. Sarkar, I. Robledo, P. Grother, and K. W. Bowyer, “Multilinear principal component analysis of tensor objects for recognition,” in *Proc. International Conference on Pattern Recognition*, vol. 1, August 2002, pp. 385–388.