UNCORRELATED MULTILINEAR DISCRIMINANT ANALYSIS WITH REGULARIZATION FOR GAIT RECOGNITION

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

ABSTRACT

This paper proposes a novel uncorrelated multilinear discriminant analysis (UMLDA) algorithm for the challenging problem of gait recognition. A tensor-to-vector projection (TVP) of tensor objects is formulated and the UMLDA is developed using TVP to extract uncorrelated discriminative features directly from tensorial data. The small-sample-size (SSS) problem present when discriminant solutions are applied to the problem of gait recognition is discussed and a regularization procedure is introduced to address it. The effectiveness of the proposed regularization is demonstrated in the experiments and the regularized UMLDA algorithm is shown to outperform other multilinear subspace solutions in gait recognition.

1. INTRODUCTION

Gait recognition [1, 2], the identification of individuals in video sequences by the way they walk, has gained significant attention recently. This interest is strongly motivated by the need for automated person identification system, visual surveillance at a distance in security-sensitive environments such as banks, airports and large civic structures, where other biometrics such as fingerprint, face or iris information can not be utilized [3]. Furthermore, night vision capability is usually impossible with other biometrics due to the limited signature in the IR image [3]. Gait is a complex spatio-temporal biometric that can address these problems effectively.

Binary gait silhouette sequences are taken as the input in most of the gait recognition algorithms proposed in the literature. The binary sequences are three-dimensional objects naturally represented as third-order tensors in a very high-dimensional tensor space, with the spatial row, column and the temporal modes for the three dimensions. To deal with these tensor objects directly, classical vector-based linear feature extraction algorithms such as the Principal Component Analysis (PCA) and Linear Discriminant Analysis (LDA) need to reshape (vectorize) the input into vectors in a very high-dimensional space, resulting in high computation and memory demand. Furthermore, the input reshaping breaks the structure and correlation in the original data and thus the redundancy and structure in the original data is not fully utilized.

Lately, multilinear subspace algorithms operating directly on the gait sequences in their tensor representation rather than their vectorized versions have been proposed. The multilinear PCA (MPCA) framework [4] attempts 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 and it has achieved good results when applied to the gait recognition problem. Nonetheless, MPCA is an unsupervised method and the class information is not used in the feature extraction process. This motivated research towards the development of supervised multilinear methodologies. A number of such solutions have been introduced recently. The multilinear discriminant analysis (MDA) proposed in [5] maximizes a tensor-based scatter ratio criterion, but unfortunately the algorithm does not converge and performs poorly on tensorial gait data [4]. In [6], a so-called general tensor discriminant analysis (GTDA) algorithm is proposed by maximizing a scatter difference criterion. Although the algorithm converges, its direct application on tensorial gait data results in poor performance [6]. All these three methodologies are based on the tensor-to-tensor projection (TTP). The so-called Discriminant Tensor Rank-one Decomposition (DTROD) algorithm [7, 8], which uses the scatter difference criterion, obtains a number of rank-one projections from the residues of the original tensor data and it can be viewed as a tensor-to-vector projection (TVP). This "greedy" approach, originally proposed in [9], is a heuristic development without theoretical justification and systematic determination of parameter settings.

In this paper, a novel uncorrelated multilinear discriminant analysis (UMLDA) is proposed to extract uncorrelated discriminative features directly from tensorial data based on the Fisher's discrimination criterion. In the next section, basic notations and multilinear algebra are introduced and the tensor-to-vector projection (TVP) is formulated as a number of elementary multilinear projections (EMPs). The UMLDA is then derived in Section 3 and the small-sample-size (SSS) problem in gait recognition is analyzed and a regularization procedure is introduced to tackle this problem. Finally, the experimental results are shown in Sec. 4 to demonstrate the effectiveness of the proposed methods and conclusions are drawn in Sec. 5.

2. MULTILINEAR ALGEBRA BASICS

2.1. Notations and basic multilinear algebra

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., $n = 1, 2, ..., N$. Throughout this paper, the discussion is restricted to real-valued vectors, matrices and tensors since the targeted application (holistic gait recognition using binary silhouettes) involve real data only.

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: $(\mathcal{A} \times_n \mathbf{U})(i_1, ..., i_{n-1}, j_n, i_{n+1}, ..., i_N) = \sum_{i_n} \mathcal{A}(i_1, ..., i_N) \cdot \mathbf{U}(j_n, i_n)$. The scalar product of two tensors $\mathcal{A}, \mathcal{B} \in \mathbb{R}^{I_1 \times I_2 \times ... \times I_N}$ is defined as: $< \mathcal{A}, \mathcal{B} >= \sum_{i_1} \sum_{i_2} ... \sum_{i_N} \mathcal{A}(i_1, i_2, ..., i_N) \cdot \mathcal{B}(i_1, i_2, ..., i_N)$. The "$n$-mode vectors" of $\mathcal{A}$ are defined as the I_n-dimensional vectors obtained from $\mathcal{A}$ by varying the index i_n while keeping all the other indices fixed. A rank-1 tensor $\mathcal{A}$ equals to the outer product of N vectors: $\mathcal{A} = \mathbf{u}^{(1)} \circ \mathbf{u}^{(2)} \circ ... \circ \mathbf{u}^{(N)}$, which means that $\mathcal{A}(i_1, i_2, ..., i_N) = \mathbf{u}^{(1)}(i_1) \cdot \mathbf{u}^{(2)}(i_2) \cdot ... \cdot \mathbf{u}^{(N)}(i_N)$ for all values of indices. Unfolding $\mathcal{A}$ along the n-mode is denoted as $\mathbf{A}_{(n)} \in \mathbb{R}^{I_n \times (I_1 \times ... \times I_{n-1} \times I_{n+1} \times ... \times I_N)}$. The column vectors of $\mathbf{A}_{(n)}$ are the n-mode vectors of $\mathcal{A}$.

2.2. Tensor-to-Vector projection for classification

The classification of tensor objects in this paper is determined through a multilinear projection from a tensor space to a vector space. Firstly, the projection from a tensor to a scalar is considered. A tensor $\mathcal{X} \in \mathbb{R}^{I_1 \times I_2 \times ... \times I_N}$ is projected to a point y as:

$$y = \mathcal{X} \times_1 \mathbf{u}^{(1)^T} \times_2 \mathbf{u}^{(2)^T} ... \times_N \mathbf{u}^{(N)^T}, \qquad (1)$$

which can also be written as the inner product:

$$y =< \mathcal{X}, \mathbf{u}^{(1)} \circ \mathbf{u}^{(2)} \circ ... \circ \mathbf{u}^{(N)} > . \qquad (2)$$

Let $\mathcal{U} = \mathbf{u}^{(1)} \circ \mathbf{u}^{(2)} \circ ... \circ \mathbf{u}^{(N)}$, then $y =< \mathcal{X}, \mathcal{U} >$. Such a multilinear projection $\{\mathbf{u}^{(1)^T}, \mathbf{u}^{(2)^T}, ..., \mathbf{u}^{(N)^T}\}$, hereafter named an elementary multilinear projection (EMP), is the projection of a tensor on a single multilinear projection direction, and it consists of one projection vector in each mode.

The projection of a tensor object $\mathcal{X}$ to $\mathbf{y} \in \mathbb{R}^P$ in a P-dimensional vector space consists of P EMPs

$$\{\mathbf{u}_p^{(1)^T}, \mathbf{u}_p^{(2)^T}, ..., \mathbf{u}_p^{(N)^T}\}, p = 1, ..., P, \qquad (3)$$

which can be written concisely as $\{\mathbf{u}_p^{(n)^T}, n = 1, ..., N\}_{p=1}^P$. This tensor-to-vector multilinear projection is therefore written as

$$\mathbf{y} = \mathcal{X} \times_{n=1}^N \{\mathbf{u}_p^{(n)^T}, n = 1, ..., N\}_{p=1}^P, \qquad (4)$$

where the p^{th} component of $\mathbf{y}$ is obtained from the p^{th} EMP as:

$$\mathbf{y}(p) = \mathcal{X} \times_1 \mathbf{u}_p^{(1)^T} \times_2 \mathbf{u}_p^{(2)^T} ... \times_N \mathbf{u}_p^{(N)^T}. \qquad (5)$$

3. UNCORRELATED MULTILINEAR DISCRIMINANT ANALYSIS WITH REGULARIZATION FOR GAIT RECOGNITION

In a typical tensor object classification task, a set of M training tensor object samples $\{\mathcal{X}_1, \mathcal{X}_2, ..., \mathcal{X}_M\}$ is available. For the convenience of discussion, the mean of these samples is assumed to be zero, without loss of generality. Each tensor object $\mathcal{X}_m \in \mathbb{R}^{I_1 \times I_2 \times ... \times I_N}$ assumes values in the tensor space $\mathbb{R}^{I_1} \bigotimes \mathbb{R}^{I_2} ... \bigotimes \mathbb{R}^{I_N}$, where I_n is the n-mode dimension of the tensor. The objective of uncorrelated multilinear discriminant analysis (UMLDA) is to find a set of P EMPs $\{\mathbf{u}_p^{(n)} \in \mathbb{R}^{I_n \times 1}, n = 1, ..., N\}_{p=1}^P$ mapping from the original tensor space $\mathbb{R}^{I_1} \bigotimes \mathbb{R}^{I_2} ... \bigotimes \mathbb{R}^{I_N}$ into a vector subspace $\mathbb{R}^P$ (with $P < \prod_{n=1}^N I_n$):

$$\mathbf{y}_m = \mathcal{X}_m \times_{n=1}^N \{\mathbf{u}_p^{(n)^T}, n = 1, ..., N\}_{p=1}^P, m = 1, ..., M, \qquad (6)$$

such that the Fisher's discriminant criterion is maximized in each EMP direction, subject to the constraint that the P coordinate vectors $\{\mathbf{g}_p \in \mathbb{R}^M, p = 1, ..., P\}$ are uncorrelated. The m^{th} component of the p^{th} coordinate vector $\mathbf{g}_p$, $\mathbf{g}_p(m)$, is the projection of the m^{th} sample $\mathcal{X}_m$ on the p^{th} EMP $\{\mathbf{u}_p^{(n)^T}, n = 1, ..., N\}$:

$$\mathbf{g}_p(m) = \mathcal{X}_m \times_{n=1}^N \{\mathbf{u}_p^{(n)^T}, n = 1, ..., N\}. \qquad (7)$$

The objective function for the p^{th} EMP can be written, in terms of the between-class scatter $S_{B_p}^{\mathbf{y}}$ and the within-class scatter $S_{W_p}^{\mathbf{y}}$ of the p^{th} projected features $\{y_{m_p}, m = 1, ..., M\}$, where y_{m_p} is the projection of the m^{th} sample by the p^{th}

EMP, as following:

$$\{\mathbf{u}_p^{(n)^T}, n = 1, ..., N\} = \arg\max \frac{\mathbf{S}^{\mathbf{y}}_{B_p}}{\mathbf{S}^{\mathbf{y}}_{W_p}}, \quad (8)$$
$$\texttt{subject to } \mathbf{g}_p^T \mathbf{g}_q = \delta_{pq},\ p, q = 1, ..., P,$$

where $\mathbf{S}^{\mathbf{y}}_{B_p} = \sum_{c=1}^{C} N_c(\bar{y}_{c_p} - \bar{y}_p)^2$, $\bar{y}_p = \frac{1}{M}\sum_m y_{m_p} = 0$, $\mathbf{S}^{\mathbf{y}}_{W_p} = \sum_{m=1}^{M}(y_{m_p} - \bar{y}_{c_{m}p})^2$, $\bar{y}_{c_p} = \frac{1}{N_c}\sum_{m,c_m=c} y_{m_p}$, and δ_{pq} is the Kronecker delta (defined as 1 for $p = q$ and as 0 otherwise).

The P EMPs $\{\mathbf{u}_p^{(n)^T}, n = 1, ..., N\}_{p=1}^{P}$ are determined as follows:

Step 1: The first EMP $\{\mathbf{u}_1^{(n)^T}, n = 1, ..., N\}$ is obtained by maximizing $\frac{\mathbf{S}^{\mathbf{y}}_{B_1}}{\mathbf{S}^{\mathbf{y}}_{W_1}}$ without any constraint.

Step 2: The $p^{th}(p > 1)$ EMP $\{\mathbf{u}_p^{(n)^T}, n = 1, ..., N\}$ is obtained by maximizing $\frac{\mathbf{S}^{\mathbf{y}}_{B_p}}{\mathbf{S}^{\mathbf{y}}_{W_p}}$ subject to the constraint that $\mathbf{g}_p^T\mathbf{g}_q = 0$ for $q = 1, ..., p-1$.

In the following, the two-step UMLDA solution will be described in detail. The procedures are summarized in the pseudo-code in Fig. 1.

3.1. Determine the first EMP

The problem of projecting the tensor samples onto a line where the projected samples are well separated is considered first. Through the first EMP $\{\mathbf{u}_1^{(n)^T}, n = 1, ..., N\}$, a corresponding set of M samples $\{y_{m_1}, m = 1, ..., M\}$ is obtained with the objective of maximizing the Fisher's criterion $J_1(\{\mathbf{u}_1^{(n)^T}, n = 1, ..., N\}) = \frac{\mathbf{S}^{\mathbf{y}}_{B_1}}{\mathbf{S}^{\mathbf{y}}_{W_1}}$.

As in the case of other multilinear algorithms [4–8], there is currently no way to simultaneously obtain in all modes those projection vectors $\{\mathbf{u}_1^{(1)}, \mathbf{u}_1^{(2)}, ...\mathbf{u}_1^{(N)}\}$ that maximizes $J_1(\{\mathbf{u}_1^{(n)^T}, n = 1, ..., N\})$. The commonly used alternating projection principle is used instead. In other words, the projection vector is solved one by one and while solving the projection vector in a particular mode n^*, the projection vectors in all the other modes $\{n \neq n^*\}$ are assumed to be known and fixed, based on some projection initialization procedure.

When $\{\mathbf{u}_1^{(n)}, n \neq n^*\}$ is given, the tensor samples are projected in these $(N-1)$ modes $\{n \neq n^*\}$ first to obtain $\tilde{\mathbf{y}}_{m_1}^{(n^*)} = \mathcal{X}_m \times_1 \mathbf{u}_1^{(1)^T} ... \times_{n^*-1} \mathbf{u}_1^{(n^*-1)^T} \times_{n^*+1} \mathbf{u}_1^{(n^*+1)^T} ... \times_N \mathbf{u}_1^{(N)^T}$, $\tilde{\mathbf{y}}_{m_1}^{(n^*)} \in \mathbb{R}^{I_{n^*}}$. Thus, the problem becomes a classical LDA problem, with input samples $\{\tilde{\mathbf{y}}_{m_1}^{(n^*)}, m = 1, ..., M\}$ and the projection to be solved is given by $\mathbf{u}_1^{(n^*)}$. In the input space, the between-class scatter $\tilde{\mathbf{S}}_{B_1}^{(n^*)}$ and within-class

---

Input: A set of tensor samples $\{\mathcal{X}_m \in \mathbb{R}^{I_1 \times I_2 \times ... \times I_N}, m = 1, ..., M\}$ with class labels $\mathbf{c} \in \mathbb{R}^M$, the desired feature vector length P, the maximum number of iterations K and a small number ϵ for testing convergence.

Output: The P EMPs $\{\mathbf{u}_p^{(n)^T}, n = 1, ..., N\}_{p=1}^{P}$ that best separate classes in the projected space and the feature vectors $\{\mathbf{y}_m \in \mathbb{R}^P, m = 1, ..., M\}$ of the input tensor samples.

Algorithm:

Step 1 (The first EMP) :

- For $n = 1, ..., N$, initialize $\mathbf{u}_{1_{(0)}}^{(n)} \in \mathbb{R}^{I_n}$.
- For $k = 1 : K$
  - For $n = 1 : N$
    - Calculate $\tilde{\mathbf{y}}_{m_1}^{(n)} = \mathcal{X}_m \times_1 \mathbf{u}_{1_{(k-1)}}^{(1)^T} ... \times_{n-1} \mathbf{u}_{1_{(k-1)}}^{(n-1)^T} \times_{n+1} \mathbf{u}_{1_{(k-1)}}^{(n+1)^T} ... \times_N \mathbf{u}_{1_{(k-1)}}^{(N)^T}$, for $m = 1, ..., M$.
    - Calculate $\tilde{\mathbf{S}}_{B_1}^{(n)}$ and $\tilde{\mathbf{S}}_{W_1}^{(n)}$. Set $\mathbf{u}_{1_{(k)}}^{(n)}$ to be the (unit) eigenvector of $\tilde{\mathbf{S}}_{W_1}^{(n)^{-1}} \tilde{\mathbf{S}}_{B_1}^{(n)}$ associated with the largest eigenvalue.
  - If $\| \mathbf{u}_{1_{(k)}}^{(n)} - \mathbf{u}_{1_{(k-1)}}^{(n)} \|_2 < \epsilon$ for all n, set $\mathbf{u}_1^{(n)} = \mathbf{u}_{1_k}^{(n)}$ for all n and break.
- The coordinate vector $\mathbf{g}_1$ is obtained with $\mathbf{g}_1(m) = \mathcal{X}_m \times_1 \mathbf{u}_1^{(1)^T} \times_2 \mathbf{u}_1^{(2)^T} ... \times_N \mathbf{u}_1^{(N)^T}$.

Step 2 (The rest EMPs) :

For $p = 2 : P$

- For $n = 1, ..., N$, initialize $\mathbf{u}_{p_{(0)}}^{(n)} \in \mathbb{R}^{I_n}$.
- For $k = 1 : K$
  - For $n = 1 : N$
    - Calculate $\tilde{\mathbf{y}}_{m_p}^{(n)} = \mathcal{X}_m \times_1 \mathbf{u}_{p_{(k-1)}}^{(1)^T} ... \times_{n-1} \mathbf{u}_{p_{(k-1)}}^{(n-1)^T} \times_{n+1} \mathbf{u}_{p_{(k-1)}}^{(n+1)^T} ... \times_N \mathbf{u}_{p_{(k-1)}}^{(N)^T}$, for $m = 1, ..., M$.
    - Calculate $\mathbf{R}_p^{(n)}$, $\tilde{\mathbf{S}}_{B_p}^{(n)}$ and $\tilde{\mathbf{S}}_{W_p}^{(n)}$. Set $\mathbf{u}_{p_{(k)}}^{(n)}$ to be the (unit) eigenvector of $\left(\tilde{\mathbf{S}}_{W_p}^{(n)}\right)^{-1} \mathbf{R}_p^{(n)} \tilde{\mathbf{S}}_{B_p}^{(n)}$ associated with the largest eigenvalue.
  - If $\| \mathbf{u}_{p_{(k)}}^{(n)} - \mathbf{u}_{p_{(k-1)}}^{(n)} \|_2 < \epsilon$ for all n, set $\mathbf{u}_p^{(n)} = \mathbf{u}_{p_k}^{(n)}$ for all n and break.
- The coordinate vector $\mathbf{g}_p$ is obtained with $\mathbf{g}_p(m) = \mathcal{X}_m \times_1 \mathbf{u}_p^{(1)^T} \times_2 \mathbf{u}_p^{(2)^T} ... \times_N \mathbf{u}_p^{(N)^T}$.

Step 3 (Projection) :

The feature vector after projection is obtained as $\mathbf{y}_m = \mathcal{X}_m \times_{n=1}^{N} \{\mathbf{u}_p^{(n)^T}, n = 1, ..., N\}_{p=1}^{P}$, for $m = 1, ..., M$, or $\mathbf{y}_m$ can be obtained with $\mathbf{y}_m(p) = \mathbf{g}_p(m)$.

Fig. 1. The pseudo-code implementation of the UMLDA algorithm.

scatter $\tilde{\mathbf{S}}_{W_1}^{(n^*)}$ are defined as

$$\tilde{\mathbf{S}}_{B_1}^{(n^*)} = \sum_{c=1}^{C} N_c(\bar{\tilde{\mathbf{y}}}_{c_1}^{(n^*)} - \bar{\tilde{\mathbf{y}}}_1^{(n^*)})(\bar{\tilde{\mathbf{y}}}_{c_1}^{(n^*)} - \bar{\tilde{\mathbf{y}}}_1^{(n^*)})^T, \quad (9)$$

$$\tilde{\mathbf{S}}_{W_1}^{(n^*)} = \sum_{m=1}^{M} (\tilde{\mathbf{y}}_{m_1}^{(n^*)} - \bar{\tilde{\mathbf{y}}}_{c_{m1}}^{(n^*)})(\tilde{\mathbf{y}}_{m_1}^{(n^*)} - \bar{\tilde{\mathbf{y}}}_{c_{m1}}^{(n^*)})^T \quad (10)$$

where $\bar{\tilde{\mathbf{y}}}_{c_1}^{(n^*)} = \frac{1}{N_c}\sum_{m,c_m=c}\tilde{\mathbf{y}}_{m_1}^{(n^*)}$, and $\bar{\tilde{\mathbf{y}}}_1^{(n^*)} = \frac{1}{M}\sum_m \tilde{\mathbf{y}}_{m_1}^{(n^*)} = \mathbf{0}$. Thus, the $\mathbf{u}_1^{(n^*)}$ that maximizes the Fisher's criterion $\frac{\mathbf{u}_1^{(n^*)^T}\tilde{\mathbf{S}}_{B_1}^{(n^*)}\mathbf{u}_1^{(n^*)}}{\mathbf{u}_1^{(n^*)^T}\tilde{\mathbf{S}}_{W_1}^{(n^*)}\mathbf{u}_1^{(n^*)}}$ in the projected space is obtained as the eigenvector of $\tilde{\mathbf{S}}_{W_1}^{(n^*)^{-1}}\tilde{\mathbf{S}}_{B_1}^{(n^*)}$ associated with the largest eigenvalue (provided that $\tilde{\mathbf{S}}_{W_1}^{(n^*)}$ is nonsingular). Starting with initialized $\{\mathbf{u}_1^{(n)}\}$, this procedure is repeated for each mode in sequence until a maximum number of iterations K is reached or the EMP converges, i.e., $\| \mathbf{u}_{1_{(k)}}^{(n)} - \mathbf{u}_{1_{(k-1)}}^{(n)} \|_2 < \epsilon$ for all n, where ϵ is a small number chosen empirically.

3.2. Determine the p^{th} EMP given the first $(p-1)$ EMPs

Now, assuming that the first $(p-1)$ EMPs are available, the p^{th} EMP is to be determined so that the scatter ratio $\frac{S_{B_p}^{\mathbf{y}}}{S_{W_p}^{\mathbf{y}}}$ is maximized, subject to the constraint that the projection by the p^{th} EMP is uncorrelated with the projections by the first $(p-1)$ EMPs.

An alternating projection approach is considered. With given $\{\mathbf{u}_p^{(n)}, n \neq n^*\}$, the tensor samples are projected in these $(N-1)$ modes first to obtain $\tilde{\mathbf{y}}_{m_p}^{(n^*)} = \mathcal{X}_m \times_1 \mathbf{u}_p^{(1)^T} ... \times_{n-1} \mathbf{u}_p^{(n-1)^T} \times_{n-1} \mathbf{u}_p^{(n-1)^T} \times_{n+1} \mathbf{u}_p^{(n+1)^T} ... \times_N \mathbf{u}_p^{(N)^T}$, $\tilde{\mathbf{y}}_{m_p}^{(n^*)} \in \mathbb{R}^{I_{n^*}}$. Let $\tilde{\mathbf{Y}}_p^{(n^*)} \in \mathbb{R}^{I_{n^*}\times M}$ be a matrix with its m^{th} column to be $\tilde{\mathbf{y}}_{m_p}^{(n^*)}$, then the p^{th} coordinate vector is obtained as $\mathbf{g}_p = \tilde{\mathbf{Y}}_p^{(n^*)^T}\mathbf{u}_p^{(n^*)}$. The constraint that $\mathbf{g}_p$ is uncorrelated with $\{\mathbf{g}_q, q = 1, ..., p-1\}$ can be written as

$$\mathbf{u}_p^{(n^*)^T}\tilde{\mathbf{Y}}_p^{(n^*)}\mathbf{g}_q = 0, q = 1, ..., p-1. \quad (11)$$

Thus, $\mathbf{u}_p^{(n^*)}$ can be determined by solving the following optimization problem:

$$\mathbf{u}_p^{(n^*)} = \arg\max \frac{\mathbf{u}_p^{(n^*)^T}\tilde{\mathbf{S}}_{B_p}^{(n^*)}\mathbf{u}_p^{(n^*)}}{\mathbf{u}_p^{(n^*)^T}\tilde{\mathbf{S}}_{W_p}^{(n^*)}\mathbf{u}_p^{(n^*)}}, \quad (12)$$

$$\texttt{subject to } \mathbf{u}_p^{(n^*)^T}\tilde{\mathbf{Y}}_p^{(n^*)}\mathbf{g}_q = 0, q = 1, ..., p-1,$$

where $\tilde{\mathbf{S}}_{B_p}^{(n^*)} = \sum_{c=1}^{C} N_c(\bar{\tilde{\mathbf{y}}}_{c_p}^{(n^*)} - \bar{\tilde{\mathbf{y}}}_p^{(n^*)})(\bar{\tilde{\mathbf{y}}}_{c_p}^{(n^*)} - \bar{\tilde{\mathbf{y}}}_p^{(n^*)})^T$, $\tilde{\mathbf{S}}_{W_p}^{(n^*)} = \sum_{m=1}^{M}(\tilde{\mathbf{y}}_{m_p}^{(n^*)} - \bar{\tilde{\mathbf{y}}}_{c_{m_p}}^{(n^*)})(\tilde{\mathbf{y}}_{m_p}^{(n^*)} - \bar{\tilde{\mathbf{y}}}_{c_{m_p}}^{(n^*)})^T$, $\bar{\tilde{\mathbf{y}}}_{c_p}^{(n^*)} = \frac{1}{N_c}\sum_{m,c_m=c}\tilde{\mathbf{y}}_{m_p}^{(n^*)}$, and $\bar{\tilde{\mathbf{y}}}_p^{(n^*)} = \frac{1}{M}\sum_m \tilde{\mathbf{y}}_{m_p}^{(n^*)} = \mathbf{0}$.

The solution is given by the following theorem:

Theorem 1 *The solution to the problem (12) is the (unit) generalized eigenvector corresponding to the largest generalized eigenvalue of the following generalized eigenvalue problem:*

$$\mathbf{R}_p^{(n^*)}\tilde{\mathbf{S}}_{B_p}^{(n^*)}\mathbf{u} = \lambda\tilde{\mathbf{S}}_{W_p}^{(n^*)}\mathbf{u}, \quad (13)$$

where

$$\mathbf{R}_p^{(n^*)} = \mathbf{I}_{I_{n^*}} - \tilde{\mathbf{Y}}_p^{(n^*)}\mathbf{G}_{p-1}\left(\mathbf{G}_{p-1}^T\tilde{\mathbf{Y}}_p^{(n^*)^T}\tilde{\mathbf{S}}_{W_p}^{(n^*)^{-1}}\tilde{\mathbf{Y}}_p^{(n^*)}\mathbf{G}_{p-1}\right)^{-1}\mathbf{G}_{p-1}^T\tilde{\mathbf{Y}}_p^{(n^*)^T}\tilde{\mathbf{S}}_{W_p}^{(n^*)^{-1}}, \quad (14)$$

$$\mathbf{G}_{p-1} = [\mathbf{g}_1 \quad \mathbf{g}_2 \quad ...\mathbf{g}_{p-1}] \in \mathbb{R}^{M\times(p-1)}, \quad (15)$$

and $\mathbf{I}_{I_{n^*}}$ is an identity matrix of size $I_{n^*} \times I_{n^*}$.

Proof The proof is not included due to space limitation.

3.3. Intialization

The iterative determination of each EMP $\{\mathbf{u}_p^{(n)}, n = 1, ..., N\}$ requires initialization. In this paper, the EMP initialization procedure is as follows: Let $\{\mathbf{v}_{j_n}^{(n)}, j_n = 1, ..., J_n\}$, where $J_n = min\{I_n, C-1\}$, be the J_n projection bases obtained by applying the classical LDA on the n-mode vectors (treating each n-mode vector as a sample), with corresponding eigenvalues as $\lambda_{j_n}^{(n)}$. From these bases, a total number of $\prod_{n=1}^{N} J_n$ candidate EMPs are obtained by considering all possible combinations, and each of them is associated with a discrimination score $D_{j_1 j_2 ... j_N} = \prod_{n=1}^{N}\lambda_{j_n}^{(n)}$. These EMP candidates are ordered according to $D_{j_1 j_2 ... j_N}$ in descending order and the p^{th} candidate is taken sequentially as the initialization for the p^{th} EMP.

3.4. Regularized UMLDA for gait recognition

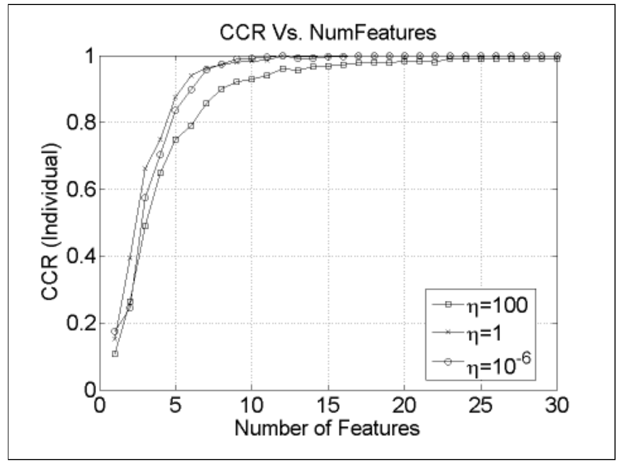
Although multilinear subspace solutions for gait recognition usually do not have the numerical small-sample-size (SSS) problems associated with traditional discriminant methodologies, the SSS problem does exist in gait recognition as well, especially when iterative discriminant analysis methods are considered. Since there is a large number of possible EMPs which can be used for the projection of gait tensors, when the number of samples per class is small, it is always possible to find some EMPs such that the projected features of the same class has almost zero within-class scatter. In the simulation studies reported here, it has been observed that when iterations maximize the scatter ratio, they tend to decrease the within-class scatter towards zero, severely overfitting the training data. However, in the challenging problem of gait recognition, large within-class scatter should be expected. Therefore, a regularization term is introduced to prevent the iterative procedure to shrink the within-class scatter while focusing on maximizing the between-class scatter, i.e., $\tilde{\mathbf{S}}_{W_p}^{(n^*)R} = \tilde{\mathbf{S}}_{W_p}^{(n^*)} + \eta \cdot \mathbf{I}_{I_{n^*}}$ is used instead of $\tilde{\mathbf{S}}_{W_p}^{(n^*)}$ in (13), where η is the regularization parameter and it is determined empirically in this paper. In addition, in computing the matrix inverse in (14), a small term $(10^{-6} \cdot \mathbf{I}_{p-1})$ is added, where $\mathbf{I}_{p-1}$ is a $(p-1) \times (p-1)$ identity matrix, in order to get better conditioned matrix for the inverse computation.

4. EXPERIMENTAL RESULTS

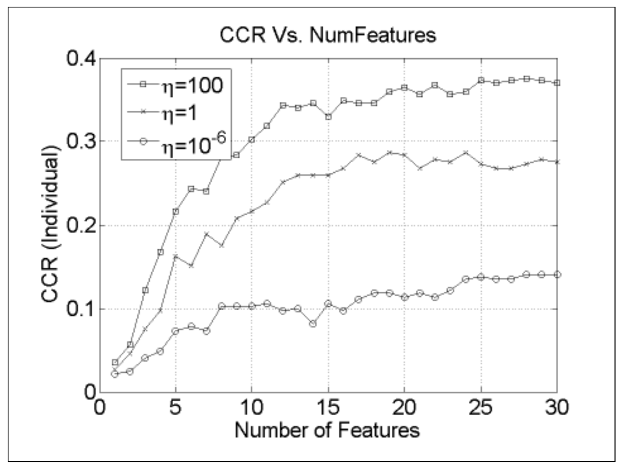
In this section, i) the compactness of the projected features obtained by UMLDA is illustrated on the problem of tensorial gait classification, and ii) the effectiveness of the regularization is demonstrated. The gait recognition experiments are carried out on the USF HumanID "Gait Challenge" data sets version 1.7 [1] for preliminary evaluation. The human gait sequences in these data sets were captured under different conditions (walking surfaces, shoe types and viewing angles). The gallery set contains 71 sequences (subjects) and seven experiments (probe sets) are designed for human identification. Gait samples (half gait cycles) of size $64 \times 44 \times 20$ are obtained following the procedures in [4] and Fig. 2 shows two examples as unfolded images. There are 725 gait samples in the Gallery set and each subject has an average of roughly 10 samples available. The nearest mean classifier and the L1 distance measure are used in the following experiments for preliminary testing, and the correct classification rate (CCR) is used for performance evaluation. In all the experiments, we set $K = 10$ and $\epsilon = 10^{-6}$.

Fig. 2. Two gait silhouette samples (unfolded).

In the first experiment, the small-sample-size problem in gait recognition is illustrated. The first five samples of each sequence (355 in total) from the gallery set are used as the training data and the rest 370 samples are used as the test data. Since the test data and the training data are captured under the same condition, the classification performance is expected to be good. However, the UMLDA performs poorly in this experiment, which is due to the SSS problem explained above. The results obtained with the regularized UMLDA (R-UMLDA) on the training samples and test samples are shown in Figs. 3(a) and 3(b), respectively, with $\eta = 100$, 1 and 10^{-6}. From the figures, it can been seen that although a stronger regularization results in less compact clusters on the training set, it has better classification results on the test set, indicating better generalization.



(a) The classification of training samples.



(b) The classification of test samples.

Fig. 3. The SSS problem in gait recognition and the effects of regularization.

Next, the R-UMLDA with $\eta = 100$ is applied on the whole gallery samples to extract $P = 70$ features and compared against the MPCA and DTROD algorithms in gait recognition. The regularization parameter $\eta = 100$ is empirically selected here for illustration. It is not optimized and a systematic way to set η will be investigated in future work. The classification results on the gallery set are shown in Fig. 4(a), where the R-UMLDA outperforms the others significantly, showing the R-UMLDA results in more compact and well-separated clusters in the projected space. The averaged recognition results for the probe samples and probe sequences [1] from the seven probe sets (probes A to G) are shown in Figs. 4(b) and 4(c), respectively, where R-UMLDA is the best performing algorithm in the figures. In particular, the first a few features extracted by R-UMLDA

[1]The matching score of a sequence is obtained as the average matching score of its samples.

are very powerful. Nevertheless, a current limitation of the R-UMLDA is that the discriminability of features extracted by R-UMLDA drops to small values as P increases, e.g., the scatter ratio is less than 1 after the 11^{th} feature and it is less than 0.5 after the 24^{th} feature in this experiment. Therefore, the number of discriminative features is limited and further research needs to be done to solve this problem.

5. CONCLUSIONS

In this paper, a novel uncorrelated multilinear discriminant analysis (UMLDA) algorithm is proposed to extract uncorrelated discriminative features directly from tensorial data using the tensor-to-vector projection of tensor objects. A regularized UMLDA is further developed to tackle the small-sample-size problem in the challenging gait recognition problem. Experiments demonstrates the effectiveness of the regularization procedure and the R-UMLDA has achieved better gait recognition results than other multilinear subspace solutions.

6. REFERENCES

[1] 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 Trans. Pattern Anal. Machine Intell.*, vol. 27, no. 2, pp. 162–177, Feb. 2005.

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

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

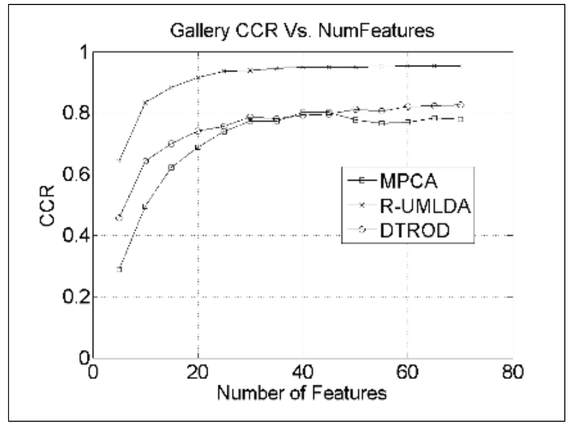
[4] H. Lu, K. N. Plataniotis, and A. N. Venetsanopoulos, "MPCA: Multilinear principal component analysis of tensor objects," *IEEE Trans. Neural Networks*, vol. 18, no. 6, Nov. 2007, to be published.

[5] S. Yan, D. Xu, Q. Yang, L. Zhang, X. Tang, and H. Zhang, "Multilinear discriminant analysis for face recognition," *IEEE Trans. Image Processing*, vol. 16, no. 1, pp. 212–220, Jan. 2007.

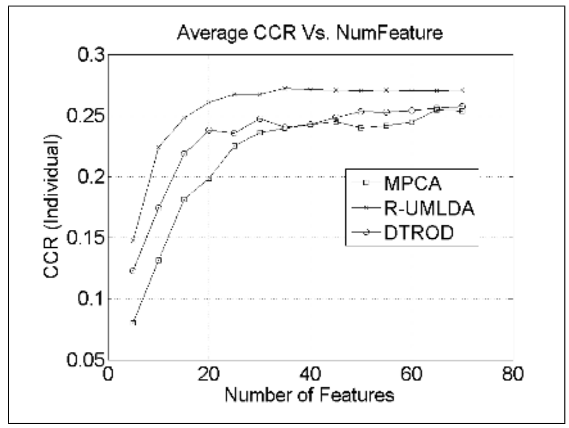
[6] D. Tao, X. Li, X. Wu, and S. J. Maybank, "General tensor discriminant analysis and gabor features for gait recognition," *IEEE Trans. Pattern Anal. Machine Intell.*, to appear.

[7] Y. Wang and S. Gong, "Tensor discriminant analysis for view-based object recognition," in *Proc. Int. Conf. on Pattern Recognition*, August 2006, vol. 3, pp. 33 – 36.

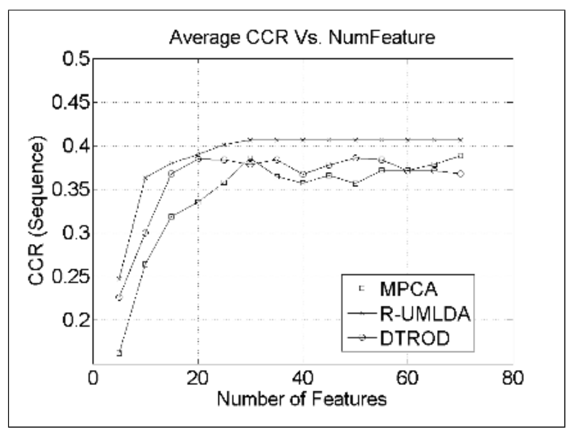
[8] D. Tao, X. Li, X. Wu, and S. J. Maybank, "Elapsed time in human gait recognition: A new approach," in *Proc. IEEE Int. Conf. on Acoustics, Speech and Signal Processing*, Apr. 2006, vol. 2, pp. 177 – 180.



(a) The classification results of the gallery samples.



(b) The average CCRs of probe samples



(c) The average CCRs of probe sequences

Fig. 4. Performance comparison of MPCA, R-UMLDA and DTROD.

[9] T. G. Kolda, "Orthogonal tensor decompositions," *SIAM Journal of Matrix Analysis and Applications*, vol. 23, no. 1, pp. 243–255, 2001.