ELAPSED TIME IN HUMAN GAIT RECOGNITION: A NEW APPROACH

Dacheng Tao¹, Xuelong Li¹, Xindong Wu², and Steve Maybank¹

School of Computer Science and Information Systems, Birkbeck, University of London, UK
 Department of Computer Science, University of Vermont, USA.
 {dacheng, xuelong, sjmaybank}@dcs.bbk.ac.uk; xwu@cs.uvm.edu

Abstract

Human gait is an effective biometric source for human identification and visual surveillance; therefore human gait recognition becomes to be a hot topic in recent research. However, the elapsed time problem, which is in its infancy, still receives poor performance. In this paper, we introduce a novel discriminant analysis method to improve the performance. The new model inherits the merits from the tensor rank one analysis, which handles the small samples size problem naturally, and the linear discriminant analysis, which is optimal for classification. Although 2DLDA and DATR also benefit from these two methods, they cannot converge during the training procedure. This means they can be hardly utilized for practical applications. Based on a lot of experiments on elapsed time problem in human gait recognition, the new method is demonstrated to significantly outperform the existing appearance-based methods, such as the principle component analysis, the linear discriminant analysis, and the tensor rank one analysis.

1. Introduction

In visual information processing research, such as the biometrics [3][4], objects are always represented by 2ndorder or 3rd-order tensors, such as the human face [3] and the human gait [4][9]. However the current pattern classification methods, for example, the linear discriminant analysis (LDA) [5], cannot work well because the number of the training samples is much less than the dimensionality of the features, i.e. LDA encounters the under sample or small sample size (SSS) problem [5]. Moreover, LDA can only accept vectors as inputs and objects in face and gait recognition are always 2nd-order or 3rd-order tensors. Brutal vectorizing the 2ndorder or 3rd-order tensors into vectors loses the position information of the original datum [1][2]. So, to introduce the tensor representation for LDA and other pattern classification methods [10] is a suitable solution. Although Ye et al. [1] proposed the Two-Dimensional Linear Discriminant Analysis (2DLDA), which is a special case of the Discriminant Analysis with Tensor Representation (DATR) [2] developed by Yan *et al.*, both 2DLDA and DATR cannot converge during training. It is a vital problem for machine learning algorithms.

In this paper, we introduce a new approach for discriminant learning using tensor representations. We name it the tenor rank-one discriminant analysis (TR1DA). Its differences from the existing methods are given below:

• Unlike 2DLDA and DATR, which use one rank n tensor to describe an object, we use a series of rank-one tensors to represent an object. In fact, the rank n method is a special case of using a series of rank-one tensors, whose benefits were demonstrated by Shashua and Levin [6] in image coding and face recognition;

• Unlike 2DLDA and DATR, we do not directly use the Fisher discriminant criterion (FDC), but its modification style, the differential scatter discriminant criterion (DSDC) [7][8]. DSDC is an approximately convex function proved by a large number of experiments [7]. In this paper, we prove that DSDC is mathematically equivalent to the Fisher discriminant criterion (FDC) with a special requirement;

As an application of our novel learning machine, we apply the proposed TR1DA to the elapsed time problem [8] in human gait recognition. From the experimental results, TR1DA can outperform principle component analysis (PCA) [5] and LDA [5] significantly.

The notation for the learning problems considered in this paper is given as follows. The bold uppercase symbols represent tensor objects, such as X, Y, Z; normal uppercase symbols represent matrices, such as X, Y, Z; italic lowercase symbols represent vectors, such as x, y, z; normal lowercase symbols represent scale numbers, such as x,y,z; i, j, k, l represent the index numbers in a vector, matrix, or tensor; M, N represent dimensions of a space; and m, n represent the sizes of a set.

The organization of the rest paper is as follows: In Section 2, we very briefly introduce PCA, LDA and TR1A. In Section 3, DSDC, TR1DA, and the convergence checking criterion are given. The experimental results are listed in Section 4 and Section 5 draws the conclusion.

2. Related Work

2.1. Principle Component Analysis (PCA)

Given a set of *n* objects, x_i , $1 \le i \le n$, $x_i \in \mathbb{R}^N$, PCA [5] diagonalizes the covariance matrix *C* according to:

$$C = \frac{1}{n} \sum_{i=1}^{n} (x_i - m) (x_i - m)^T.$$
(1)

where $m = \left(\sum_{i=1}^{n} x_i\right) / n$ is the mean vector. The eigenspace U of C is spanned by the first K eigenvectors with the largest eigenvalues, $U = [v_1, \dots, v_K]$. For an object x, it is transformed by,

$$y = U^T \left(x - m \right). \tag{2}$$

For recognition, the prototype x_p and the test object x_t are projected onto U to get y_p and y_t by (2). The class is found by minimizing the distance $\varepsilon = \|y_t - y_p\|$.

2.2. Linear Discriminant Analysis (LDA)

PCA is optimal for representation. LDA [5][3] tries to find the subspace that best discriminates different classes. It is spanned by U to maximize the ratio between the *between-class scatter* matrix S_b and the *within-class scatter* matrix S_w . S_b and S_w are defined as follows:

$$\begin{cases} S_{b} = \sum_{i=1}^{c} n_{i} (m_{i} - m) (m_{i} - m)^{T} \\ S_{w} = \sum_{i=1}^{c} \sum_{j=1}^{n_{i}} (x_{j}^{i} - m_{i}) (x_{j}^{i} - m_{i})^{T} \end{cases}$$
(3)

where $n = \sum_{i=1}^{c} n_i$, n_i is the number of objects in the i^{th} class, $m_i = \left(\sum_{j=1}^{n_i} x_j^i\right) / n_i$ is the mean vector for class C_i , $m = \sum_{i=1}^{c} m_i$ is the total mean vector, and x_j^i is the j^{th} sample in the i^{th} class. The subspace of LDA is spanned by a set of vectors $U = [u_1, \dots, u_{c-1}]$, according to:

$$U = \underset{U}{\arg\max} \frac{\left\| U^{T} S_{b} U \right\|}{\left\| U^{T} S_{w} U \right\|} .$$
(4)

The optimal projection direction U can be computed from the eigenvectors of $S_w^{-1}S_b$. For recognition, the linear discriminant function for the class prototype x_p and test object x_t is computed to minimize $\varepsilon = \left\| U^T \left(x_t - x_p \right) \right\|$. Generally, S_w will become singular, and LDA vectors are difficult to compute.

2.3. Tensor Rank-One Analysis (TR1A)

Shashua [6] demonstrated the efficiency of TR1A in image representation and classification compared with PCA. TR1A can accept an M^{th} -order tensor as input for further processing. Given an M^{th} -order tensor **X**, TR1A wants to find the optimal solution with a minimal possible r to minimize ε according to:

$$\varepsilon = \sum_{r=1}^{R} \varepsilon_r \tag{5}$$

where *R* is the number of rank-one tensors and ε_r is the r^{th} reconstruction error. The r^{th} reconstruction error ε_r is defined by an iterative method:

 $\mathbf{X}_{1}^{k} = \mathbf{X}^{k} \tag{6}$

$$\varepsilon_{1} = \sum_{k=1}^{n} \left\| \mathbf{X}_{1}^{k} - \lambda_{1}^{k} \prod_{j=1}^{M} u_{1}^{j} \right\|$$
(7)

$$\mathbf{X}_{r}^{k} = \mathbf{X}_{r-1}^{k} - \lambda_{r-1}^{k} \prod_{j=1}^{M} u_{r-1}^{j}$$
(8)

$$\varepsilon_r = \sum_{k=1}^n \left\| \mathbf{X}_r^k - \lambda_r^k \prod_{j=1}^M u_r^j \right\|$$
(9)

where u_r^j is the base vector for decomposition with $\|u_r^j\| = 1$, and $\lambda_r^k = \mathbf{X}_r^k \prod_{j=1 \times j}^M u_r^j$ is a series of scalars for optimal reconstruction. For recognition, the prototype \mathbf{X}^p for each individual class in the database and the test tensor \mathbf{X}^t to be classified are projected onto the bases to get the prototype weight vector $\lambda_p^k |_{k=1}^R$ and the test weight vector $\lambda_t^k |_{k=1}^R$. The test tensor class is found by minimizing the distance $\varepsilon = \|\lambda_t^k \|_{k=1}^R - \lambda_p^k \|_{k=1}^R$.

3. Tensor Rank-One Discriminant Analysis

The relationship between TR1A and TR1DA equals to the relationship between PCA and LDA. TR1A is optimal for representation and TR1DA is optimal for classification.

3.1. Differential Scatter Discriminant Criterion

We define the DSDC criterion in this paper, because of the weak convergence property in the previous work. The new criterion can converge quickly, and is given by:

$$u = \underset{u}{\arg\max} \left(u^T S_b u - \zeta u^T S_w u \right).$$
(10)

where ζ is a tuning parameter, u is the projection vector, and S_b, S_w are the between- and within- class scatter matrices, respectively. DSDC is relevant to LDA and it can access the optimal solution with some proper tuning parameter ζ , because LDA is defined by:

$$\max u^T S_b u; \ s.t. \ u^T S_w u = 1, \tag{11}$$

which equals to:

$$\min_{u} -\frac{1}{2}u^{T}S_{b}u; \ s.t. \ u^{T}S_{w}u = 1.$$
(12)

The Lagrangian equation of the problem is:

$$L = -\frac{1}{2}u^{T}S_{b}u + \frac{1}{2}\zeta(u^{T}S_{w}u - 1)$$

= $-\frac{1}{2}(u^{T}S_{b}u - \zeta u^{T}S_{w}u) - \frac{1}{2},$ (13)

which equals to (11) with a proper tuning parameter ζ .

3.2. Tensor Rank-One Discriminant Analysis

Because TR1DA includes many variables, we first define all the variables used in this section. $\mathbf{X}_{i,j}^k$ is the *i*th object in the *j*th class in the *k*th training iteration. For *k*=1, we have $\mathbf{X}_{i,j}^1 = \mathbf{X}_{i,j}$. Moreover, $\mathbf{X}_{i,j}^k$ is an *M*th-order tensor. $\mathbf{M}_j^k = \left(\sum_{i=1}^{n_j} \mathbf{X}_{i,j}^k\right) / n_j$ is the *j*th class mean tensor in the *k*th training iteration and $\mathbf{M}^k = \left(\sum_{j=1}^{c} \mathbf{M}_j^k\right) / c$ is the total mean tensor of all objects in the *k*th training iteration. u_i^j is the *j*th direction base vector for decomposition in the *i*th training iteration. With these definitions, TR1DA can be defined by the following equations:

$$\mathbf{X}_{i,j}^1 = \mathbf{X}_{i,j}, \ 1 \le i \le n_j \text{ and } 1 \le j \le c .$$
(14)

$$\underset{u_{l}^{l}|_{l=1}^{M}}{\operatorname{arg\,max}} \begin{pmatrix} \sum_{i=1}^{c} \left(\mathbf{M}_{i}^{1} - \mathbf{M}^{1} \right) \prod_{l=1}^{M} \left(u_{l}^{l} \right)^{T} \\ \left(\left(\mathbf{M}_{i}^{1} - \mathbf{M}^{1} \right) \prod_{l=1}^{M} \left(u_{l}^{l} \right)^{T} \right)^{T} \\ -\zeta \sum_{i=1}^{c} \sum_{j=1}^{n_{i}} \left(\left(\mathbf{X}_{j}^{1,i} - \mathbf{M}_{i}^{1} \right) \prod_{l=1}^{M} \left(u_{l}^{l} \right)^{T} \right) \\ \left(\left(\mathbf{X}_{j}^{1,i} - \mathbf{M}_{i}^{1} \right) \prod_{l=1}^{M} \left(u_{l}^{l} \right)^{T} \right) \end{pmatrix} \end{pmatrix}$$
(15)
$$\mathbf{Y}^{k} = \mathbf{Y}^{k-1} = 2^{k-1} \prod_{j=1}^{M} \left(u_{l}^{k-1} - \mathbf{Y}^{k-1} \prod_{j=1}^{M} u_{j}^{k-1} \right)$$
(16)

$$\mathbf{X}_{i,j}^{k} = \mathbf{X}_{i,j}^{k-1} - \lambda_{i,j}^{k-1} \prod_{d=1}^{k} u_{d}^{d} u_{k-1}^{d}, \lambda_{i,j}^{k-1} = \mathbf{X}_{i,j}^{k-1} \prod_{d=1}^{k} u_{k-1}^{d}$$
(16)

$$\underset{\substack{u_{k}^{l}|_{l=1}^{l}}{\operatorname{arg\,max}} \left\{ \begin{array}{c} \sum_{i=1}^{c} \left[n_{i} \left(\left(\mathbf{M}_{i}^{k} - \mathbf{M}^{k} \right) \prod_{l=1}^{m} \left[\left(u_{k}^{l} \right)^{T} \right) \right] \\ \left(\left(\left(\mathbf{M}_{i}^{k} - \mathbf{M}^{k} \right) \prod_{l=1}^{M} \left[\left(u_{k}^{l} \right)^{T} \right)^{T} \right) \right] \\ -\zeta \sum_{i=1}^{c} \sum_{j=1}^{n_{i}} \left[\left(\left(\mathbf{X}_{j,i}^{k} - \mathbf{M}_{i}^{k} \right) \prod_{l=1}^{M} \left[\left(u_{k}^{l} \right)^{T} \right) \right] \\ \left(\left(\mathbf{X}_{j,i}^{k} - \mathbf{M}_{i}^{k} \right) \prod_{l=1}^{M} \left[\left(u_{k}^{l} \right)^{T} \right] \right] \end{array} \right)$$
(17)

From the definition of the problem in (14), (15), (16), and (17), we know that TR1DA can be calculated by a

greedy approach, because of the lack of the closed form solution for the problem. The greedy approach can be seen from Figure 1. The calculation of $\mathbf{X}_{i,j}^{r}$ is based on the given $\mathbf{X}_{i,j}^{r-1}$ and $u_{r-1}^{d} |_{d=1}^{M}$. With the given $\mathbf{X}_{i,j}^{r-1}$ and $u_{r-1}^{d} |_{d=1}^{M}$, we calculate $\lambda_{i,j}^{r-1}$ via $\lambda_{i,j}^{r-1} = \mathbf{X}_{i,j}^{r-1} \prod_{d=1 \times d}^{M} u_{r-1}^{d}$. The projection base vectors $u_{r-1}^{d} |_{d=1}^{M}$ can be obtained by the alternating least square (ALS) method. In ALS, we can obtain the optimal base vector u_{r-1}^{d} with the given $u_{r-1}^{i} |_{l \le i \le M}^{i \ne d}$. We can conduct the procedure iteratively to obtain $u_{r-1}^{d} |_{d=1}^{M}$. The detailed procedure for TR1DA is given in Table 1.

Table 1. TR1DA algorithm.

Input: Training tensors $\mathbf{X}_{i,j}, 1 \le i \le n_j, 1 \le j \le c$ and the							
number of rank-one tensors R .							
Output: The base vectors $u_j^i, 1 \le i \le M, 1 \le j \le R$							
constrained by $\left\ \boldsymbol{u}_{j}^{i}\right\ = 1$.							
Step 0. Set $\mathbf{X}_{i,j}^{1} = \mathbf{X}_{i,j}, 1 \le i \le n_{j}, 1 \le j \le c$.							
Step 1. For $k = 1$ to R {							
Step 2. For $l = 1$ to M {							
Step 3. Calculate $\mathbf{X}_{i,j}^{k} = \mathbf{X}_{i,j}^{k-1} - \lambda_{i,j}^{k-1} \prod_{d=1 \times d}^{M} u_{k-1}^{d}$ with $k > 1$							
or $\mathbf{X}_{i,j}^1 = \mathbf{X}_{i,j}$ and set projection base vectors							
$u_k^d \mid_{1 \le d \le M}^{d \ne l} = 1$. Here, $\lambda_{i,j}^{k-1} = \mathbf{X}_{i,j}^{k-1} \prod_{d=1 \le d}^M u_{k-1}^d$. With the							
given $\mathbf{X}_{i,j}^k$, calculate the class mean tensor \mathbf{M}_j^k and the							
total mean tensor \mathbf{M}^k .							
Step 4. Optimize u_k^l according to (17) with the given							
$u_k^d \Big _{1 \le d \le M}^{d \ne l} = 1$.							

{//For loop in Step 2.									
Step	5.	Conduct	Steps	2-4	for	a	few	iterations	until
convergence.									

}//For loop in Step 1.

According to the algorithm described in Table 1, we can obtain $u_{r-1}^d |_{d=1}^M$ iteratively. For TR1DA, we use the coordinate value $\lambda_{i,j}^{k-1} = \mathbf{X}_{i,j}^{k-1} \prod_{d=1 \times d}^M u_{k-1}^d$ to represent **X**, where $\mathbf{X}^1 = \mathbf{X}$ and $\mathbf{X}_{i,j}^k = \mathbf{X}_{i,j}^{k-1} - \lambda_{i,j}^{k-1} \prod_{d=1 \times d}^M u_{k-1}^d$. For recognition, the prototype \mathbf{X}^p for each individual class in the database and the test tensor \mathbf{X}^t to be classified are projected onto the bases to get the prototype weight vector $\lambda_p^k |_{k=1}^R$ and the test weight vector $\lambda_i^k |_{k=1}^R$. The test tensor class is found by minimizing $\varepsilon = \left\| \lambda_t^k |_{k=1}^R - \lambda_p^k |_{k=1}^R \right\|$.

Unlike the existing tensor extension of discriminant analysis, the calculation procedure of the projection base vectors in TR1DA can converge. The following method can check the convergence of TR1DA. In the step 2, 3, and 4 of the algorithm, we can check the convergence through $\left\| \left(u_k^l \right)_t \cdot \left(u_k^l \right)_{t+1} \right\| - 1 \right\| \le \varepsilon$ with ε , which is a small value. If $\left\| \left(u_k^l \right)_t \cdot \left(u_k^l \right)_{t+1} \right\| = 1$, the calculated projection direction in the t^{th} iteration is equivalent to the $(t+I)^{\text{th}}$ iteration.

4. Experimental Results

To demonstrate the ability of TR1DA, we utilize it to solve the elapsed time problem in gait recognition. Here, we first introduce our experimental data (*gallery* and *probe*) sets [4] and then report the performances of our TR1DA algorithm against PCA, LDA, and TR1A.

Our experiments are carried out on the USF HumanID outdoor gait database [4]. The database consists of 1,870 sequences from 122 subjects, and for each subject, the covariates are up to the following five: change in *viewpoints* (Left or Right), change in *shoe type* (A or B), change in *walking surface* (Grass or Concrete), change in *carrying condition* (carrying a Briefcase or No Briefcase), and change in *elapsed time* (May or November) between sequences being compared.

Average gait (AG) stands for the mean image (pixel by pixel) of silhouettes **over a gait cycle** within a sequence. AG is robust against any errors in individual frames, so we choose AG to represent a *gait cycle*, thus one sequence yields several AGs and each AG's number depends on the gait cycle's number in this sequence. In the following experiments, the AGs are utilized as the *original data* for the *elapsed time* gait recognition. All experimental results are based on non-overlapping *gallery* and *probe* sets.

We aim to compare the performances of PCA, LDA, TR1A, and TR1DA for the most difficult gait recognition covariate - *elapsed time*. So, we choose our *gallery* set (May, C, A, L, NB) from the *May* data and choose the *probe* set from the *November* data.

We are facing eight pattern classification problems with the test sets of: CAL(<u>Nov.</u>, C, A, L, NB), CBL(<u>Nov.</u>, C, B, L, NB), GAL(<u>Nov.</u>, G, A, L, NB), GBL(<u>Nov.</u>, G, B, L, NB), CBAL(<u>Nov.</u>, C, A, L, BF), CBBL(<u>Nov.</u>, C, B, L, BF), GBAL(<u>Nov.</u>, G, A, L, BF), and GBBL(<u>Nov.</u>, G, B, L, BF).

All experimental results are shown in Table 2, which demonstrates the advantages of the proposed method by dimensionality-precision curves. According to the Table, both TR1DA and TR1A outperform the other two schemes consistently, and the former is even better than the later.

 Table 2. Experimental Results (%)

rubie 2. Emperimental results (70)										
ID.	Probe Set	PCA	LDA	TR1A	TR1DA					
1.	(Nov., G, B, L, BF)	9.5	9.0	23.5	23.5					
2.	(Nov., G, A, L, BF)	10.0	13.0	15.0	16.0					
3.	(Nov., C, B, L, BF)	11.3	7.5	17.7	21.0					
4.	(Nov., C, A, L, BF)	9.7	16.1	16.0	17.2					
5.	(Nov., G, B, L, NB)	11.0	7.5	20.5	24.6					
6.	(Nov., G, A, L, NB)	14.7	17.8	17.8	18.8					
7.	(Nov., C, B, L, NB)	13.4	14.0	18.1	21.3					
8.	(Nov., C, A, L, NB)	18.3	18.8	14.0	16.1					

5. Conclusion

A pattern classification method, named the tensor rank-one discriminant analysis (TR1DA), has been proposed in this paper. Through a series of experiments for elapsed time problem in gait recognition, TR1DA demonstrates its benefits compared with principle component analysis, linear discriminant analysis, and tensor rank one analysis.

References

[1] J. Ye, R. Janardan, and Q. Li, "Two-Dimensional Linear Discriminant Analysis", *NIPS*, 2005.

[2] S. Yan, D. Xu, Q. Yang, L. Zhang, X. Tang, and H. Zhang, "Discriminant Analysis with Tensor Representation", *CVPR*, vol. 1, pp. 526,532, 2005.

[3] W. Zhao, R. Chellappa, and P. Phillips, "Subspace Linear Discriminant Analysis for Face Recognition", Technical Report CAR-TR-914, Center for Automation Research, University of Maryland, 1999.

[4] S. Sarkar, P. Phillips, Z. Liu, I. Robledo, P. Grother, and K. W. Bowyer, "The Human ID Gait Challenge Problem: Data Sets, Performance, and Analysis", *IEEE TPAMI*, vol. 27, no. 2, pp. 162-177, 2005.

[5] K. Fukunaga, Statistical Pattern Recognition, Academic Press, 1990.

[6] A. Shashua and A. Levin, "Linear Image Coding for Regression and Classification Using the Tensor-rank Principle", *CVPR*, vol. 1, pp. 42-49, 2001.

[7] D. Tao, "Differential Scatter Biased Discriminant Analysis for Relevance Feedback in Image Retrieval," Unpublished, Dec, 2003.

[8] Q. Liu, X. Tang, H. Lu, and S. Ma, "Kernel Scatter-Difference Based Discriminant Analysis for Face Recognition", *ICPR*, vol. 2, pp. 419-422, 2004.

[9] L. Wang, T. Tan, H. Ning, and W. Hu, "Silhouette Analysis-based Gait Recognition for Human Identification," *IEEE TPAMI*, vol. 25, no. 12, pp. 1,505-1,518, 2003.

[10] D. Tao, X. Li, W. Hu, S.J. Maybank, and X. Wu, "Supervised tensor learning," *IEEE ICDM*, pp. 450-457, 2005.