Multilinear Principal Component Analysis of Tensor Objects for Recognition

Haiping Lu, K.N. Plataniotis and A.N. Venetsanopoulos
Bell Canada Multimedia Laboratory*
The Edward S. Rogers Sr. Department of Electrical and Computer Engineering
University of Toronto, M5S 3G4, Canada
{haiping, kostas, anv}@dsp.toronto.edu

Abstract

In this paper, a multilinear formulation of the popular Principal Component Analysis (PCA) is proposed, named as multilinear PCA (MPCA), where the input can be not only vectors, but also matrices or higher-order tensors. It is a natural extension of PCA and the analogous counterparts in MPCA to the eigenvalues and eigenvectors in PCA are defined. The proposed MPCA has wide range of applications as a higher-order generalization of PCA. As an example, MPCA is applied to the problem of gait recognition using a novel representation called EigenTensorGait. A gait sequence is divided into half gait cycles and each half cycle, represented as a 3rd-order tensor, is considered as one data sample. Experiments show that the proposed MPCA performs better than the baseline algorithm in human identification on the Gait Challenge data sets.

1. Introduction

Real data in pattern recognition and computer vision are often very high-dimensional. However, they can be characterized mostly by a more compact representation, e.g., subspace representation. Dimensionality reduction of such data is an important preprocessing step in many statistical pattern recognition problems. Principal Component Analysis (PCA) is a traditional linear technique for dimensionality reduction. On the other hand, while real world data, such as images and video sequences, are often multidimensional and naturally represented as 2nd-order (matrices) or higher-order tensors, PCA reshapes them into vectors in a very high-dimensional space and thus it suffers from the *curse of dimensionality*.

Multilinear algebra has recently received broad attention as many researchers become aware of the problem above and start to represent data in their natural form. Yang *et al.* recently proposed a Two-Dimensional PCA by constructing an image covariance matrix using the original image matrices [13]. He *et al.* proposed Tensor Subspace Analysis for face recognition, considering each image as a second-order tensor [3]. Three-mode PCA was studied in as early as 1983 [4], with recent application to gender recognition [2]. However, a more general formulation of PCA to higher-order tensors is not available to our knowledge.

Also, multilinear algebra has been applied to multifactor analysis. Bilinear models were proposed to separate style and content in [10]. In [11], a method was proposed for multilinear analysis of image ensembles to account for multiple factors in image formation. The method was claimed to subsume PCA but a centering step is missing. In [12], this approach is named as multilinear PCA but it is different from the proposed MPCA in this paper since images are still represented as vectors. Similarly in [8], gait image sequences are modeled by three components and silhouette images are represented as (high-dimensional) vectors.

In this paper, a new multilinear PCA (MPCA) algorithm is proposed. It is a natural extension of PCA to the multilinear case. The multilinear singular values and eigentensors are defined analogous to eigenvalues and eigenvectors in PCA, respectively. MPCA is particularly useful in applications where the data samples are naturally represented as matrices or higher-order tensors, instead of vectors. As an example, MPCA is applied to the problem of gait recognition, where gait sequences are represented naturally as 3rd-order tensors. The spatial row space, column space and the time space account for the 3 modes. Thus, a sample data set to be analyzed is a 4th-order tensor, with the sample space as the 4th-mode. Each half gait cycle is treated as a data sample, with spatial and temporal normalization. Experiments on the Gait Challenge data sets [9] show that the proposed MPCA algorithm outperforms the baseline algorithm.

*The authors would like to thank Prof. Sarkar from the USF for providing the Gait Challenge data sets. Support provided by the Communications and Information Technology Ontario Partnership Program and the Bell University Labs - at the University of Toronto is also acknowledged.

2. Multilinear Principal Component Analysis

2.1. Basic multilinear algebra

The notational conventions in [1] are used in this paper. Indices are denoted by lowercase letters and span the range from 1 to the uppercase letter of the index, e.g., $n = 1, 2, ..., N$. We denote vectors by lowercase boldface letters, e.g., $\mathbf{x}$; matrices by uppercase boldface, e.g., $\mathbf{U}$; and tensors by calligraphic letters, e.g., $\mathcal{A}$.

An Nth-order tensor is denoted as: $\mathcal{A} \in \mathbb{R}^{I_1 \times I_2 \times ... \times I_N}$. 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 defined by a tensor with entries: $(\mathcal{A} \times_n \mathbf{U})_{i_1 ... i_{n-1} j_n i_{n+1} ... i_N} = \sum_{i_n} a_{i_1 ... i_N} u_{j_n i_n}$. 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: $< \mathcal{A}, \mathcal{B} >= \sum_{i_1} \sum_{i_2} ... \sum_{i_N} a_{i_1 i_2 ... i_N} b_{i_1 i_2 ... i_N}$. The Frobenius norm of a tensor $\mathcal{A}$ is defined as $\| \mathcal{A} \| = \sqrt{< \mathcal{A}, \mathcal{A} >}$. The n-rank of $\mathcal{A}$, denoted by $R_n = rank_n(\mathcal{A})$, is the dimension of the vector space spanned by the n-mode vectors. Higher-order SVD (HOSVD) exists for tensors. Any tensor $\mathcal{A}$ can be expressed as the product: $\mathcal{A} = \mathcal{S} \times_1 \mathbf{U}^{(1)} \times_2 \mathbf{U}^{(2)} \times ... \times_N \mathbf{U}^{(N)}$, where $\mathcal{S} \in \mathbb{R}^{I_1 \times I_2 \times ... \times I_N}$ is the core tensor of which the subtensors $\mathcal{S}_{i_n = \alpha}$ have the property of all-orthogonality and ordering based on the Frobenius-norms $\| \mathcal{S}_{i_n = \alpha} \|$, and $\mathbf{U}^{(n)} = \left(\mathbf{u}_1^{(n)} \mathbf{u}_2^{(n)} ... \mathbf{u}_{I_n}^{(n)} \right)$ is a unitary $I_n \times I_n$ matrix. For more details, please refer to [5].

For convenience of discussion in making analogy between linear PCA and MPCA, we propose to use the following convention in forming the sample data tensor: for tensor objects that are $(N-1)$-th order, the sample data tensor is an Nth-order tensor of I_N samples formed with mode-N representing the sample space and mode-1 to mode-$(N-1)$ representing the $(N-1)$ modes of a tensor sample.

2.2. PCA with tensor notation

PCA chooses a dimensionality reducing linear projection that maximizes the scatter (variance) of all projected samples. It involves the SVD of the centered (mean 0) training data matrix $\mathbf{X}^c$ as following, with tensor notation:

$$\mathbf{X}^c = \mathbf{U}\mathbf{S}\mathbf{V}^T = \mathbf{U}^{(1)}\mathbf{S}\mathbf{U}^{(2)^T} = \mathbf{S} \times_1 \mathbf{U}^{(1)} \times_2 \mathbf{U}^{(2)}. \quad (1)$$

Here, in conformance with our convention, the columns of the centered data matrix $\mathbf{X}^c$ represent data points (samples) and each column of $\mathbf{X}^c$ represents a data sample (as vectors). Thus, the orthonormal basis vectors (principal components) are the columns of $\mathbf{U}^{(1)}$ of $\mathbf{X}^c$, i.e., the left singular vectors of $\mathbf{X}^c$, the variance in each direction is given by the corresponding singular values σ_k, lying along the diagonals of $\mathbf{S}$. The coordinates of the data in the basis defined by the principal components are $\mathbf{S}\mathbf{V}^T = \mathbf{S}\mathbf{U}^{(2)^T} = \mathbf{S} \times_2 \mathbf{U}^{(2)}$.

In PCA-based subspace analysis, $\mathbf{U}^{(1)}$ is truncated to $\tilde{\mathbf{U}}^{(1)} \in \mathbb{R}^{I_1 \times R_1}$ by keeping only the first R_1 columns and a centered input test sample (vector) $\mathbf{x}^c$ is projected to feature space as

$$\tilde{\mathbf{y}} = \tilde{\mathbf{U}}^{(1)^T} \mathbf{x}^c = \mathbf{x}^c \times_1 \tilde{\mathbf{U}}^{(1)^T}, \quad (2)$$

where $\tilde{\mathbf{U}}^{(1)^T}$ is the projection matrix. $\tilde{\mathbf{y}}$ is then classified by its distances to the columns of $\mathbf{X}^c \times_1 \tilde{\mathbf{U}}^{(1)^T}$.

2.3. The multilinear PCA formulation

The proposed MPCA formulation is a direct extension of PCA to the multilinear case: the input data samples to MPCA are centered as in PCA, the projection is orthonormal and the projected feature is a tensor of the same order as the sample with reduced dimension.

In PCA, the training data matrix is centered by subtracting the means of the data samples, the proposed MPCA does the same for the training tensor samples $\mathcal{X}$. $\mathcal{X}$ is centered by subtracting its mode-N mean $\bar{\mathcal{X}}^{(N)}$ from each mode-N slice (tensor sample) and the mode-N centered tensor $\mathcal{X}^c$ is decomposed using HOSVD as:

$$\mathcal{X}^c = \mathcal{S} \times_1 \mathbf{U}^{(1)} \times_2 \mathbf{U}^{(2)} ... \times_N \mathbf{U}^{(N)}. \quad (3)$$

The mode-n singular values of $\mathcal{X}^c$ are $\| \mathcal{S}_{i_n = \alpha} \|$ (symbolized by $\sigma_i^{(n)}$), where the norm is the Frobenius-norm and the subtensor $\mathcal{S}_{i_n = \alpha}$ is an n-mode slice of the tensor $\mathcal{S}$, obtained by fixing the nth index to α.

The basis vectors in PCA are contained in a single matrix after SVD, while it is not obvious what the basis tensors in MPCA are. Analogous to PCA, the basis tensors should have the same size as the tensor samples, i.e., order-$(N-1)$ tensors. As in SVD truncation for PCA, the HOSVD of $\mathcal{X}^c$ is truncated by keeping the first R_n ($R_n \leq I_n$) columns for the basis matrix $\mathbf{U}^{(n)}$ in each mode n ($1 \leq n \leq N-1$), to produce $\tilde{\mathbf{U}}^{(n)} \in \mathbb{R}^{I_n \times R_n}$. From the first $(N-1)$ basis matrices $\tilde{\mathbf{U}}^{(n)}$ ($1 \leq n \leq N-1$), we obtain $\prod_{n=1}^{N-1} R_n$ basis and each basis is a rank-1 $(N-1)$th-order tensor formed by the outer product $\mathbf{u}_{r_1}^{(1)} \circ \mathbf{u}_{r_2}^{(2)} ... \circ \mathbf{u}_{r_{N-1}}^{(N-1)}$, where r_n ranges from 1 to R_n, and $\mathbf{u}_{r_n}^{(n)}$ is the r_nth column of $\mathbf{U}^{(n)}$ [6]. The tensor projection using the collection of these $\prod_{n=1}^{N-1} R_n$ basis is denoted as

$$\tilde{\mathcal{P}}_{-N} = \times_1 \tilde{\mathbf{U}}^{(1)^T} \times_2 \tilde{\mathbf{U}}^{(2)^T} ... \times_{N-1} \tilde{\mathbf{U}}^{(N-1)^T}. \quad (4)$$

We name these basis tensors as eigentensors and for the modeling of a particular object such as face or gait, we call these eigentensors as EigenTensorObject (e.g., EigenTensorFace or EigenTensorGait), corresponding to similar terms in PCA.

In MPCA subspace analysis, a centered input tensor $\mathcal{Z}^c$ of order $(N-1)$ is projected to a lower-dimension feature tensor as

$$\tilde{\mathcal{Y}} = \mathcal{Z}^c \times \tilde{\mathcal{P}}_{-N} = \mathcal{Z}^c \times_1 \tilde{\mathbf{U}}^{(1)^T} ... \times_{N-1} \tilde{\mathbf{U}}^{(N-1)^T}. \quad (5)$$

$\tilde{\mathcal{Y}}$ is then classified based on its distances to the projections of the mode-N slices of $\mathcal{X}^c$ (the centered tensor samples).

From the above derivation, it is easy to verify that MPCA with vector samples ($N = 2$) is equivalent to PCA, i.e., MPCA subsumes PCA.

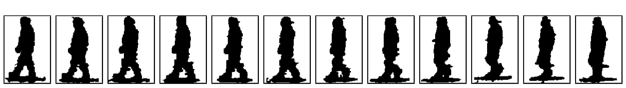
3. EigenTensorGait for gait recognition

The proposed MPCA is a general dimensionality reduction method for tensor object (multidimensional data) analysis. As an example, the application to gait recognition problem is presented here, using a higher-order tensor representation of gait silhouette sequences.

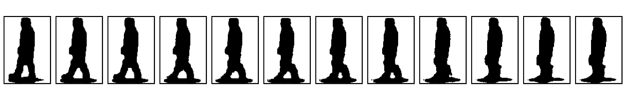
While in many recognition problems, a data sample is clearly defined, such as iris, face or fingerprint images, gait is a spatial-temporal biometric without obvious definition of a sample. The proposed method treats each half gait cycle as a data sample, which is a 3rd-order tensor. The spatial row space, column space and the time space account for the 3 modes. In turn, a whole data set to be analyzed will be a 4th-order tensor, with the addition of the sample space.

To obtain half gait cycles, a gait silhouette sequence is partitioned into half cycles in a similar way as in [9]. The numbers of foreground pixels are counted in the bottom half of silhouettes. The sequence of numbers are smoothed with a running average filter and the minimums in this number sequence partition the sequence into several half gait cycles.

Gait silhouette images are often noisy. A simple best rank-(R_1, R_2, R_3) approximation [6] of the silhouettes could reduce noise greatly. Therefore, after partitioning into half cycles, each half cycle is approximated by its best rank-$(10, 10, 3)$ approximation (ranks obtained through studying approximation errors). Figure 1 shows the approximated silhouettes, which are cleaner through visually comparison with the original silhouettes.



(a) The original silhouettes.



(b) The best rank-$(10, 10, 3)$ approximation of the silhouettes.

Figure 1. Best rank approximations.

To perform MPCA on gait tensor samples, we need to normalize the samples to a canonical size of $I_1^c \times I_2^c \times I_3^c$. Conventional image resizing algorithms are applied to normalization in spatial domain. While there are sophisticated algorithms available for temporal normalization, such as mapping a gait cycle to a unit circle using nonlinear interpolation [7], here we apply conventional image resizing algorithms, which are simple and fast in comparison, to the temporal domain normalization as well. We consider each horizontal (or lateral) slice as an image and resize this image (in mode 3) using conventional image resizing algorithms, e.g. nearest neighbor, bilinear, or bicubic interpolation.

Centering of all the half cycles obtained from the training data set results in the training tensor $\mathcal{X}^c$. The MPCA can then be applied for gait recognition. In classification, the sum of the absolute differences between the projected tensors is used as the distance measure, which is equivalent to the L_1 norm for vectors.

The basis gait tensors obtained through MPCA are named here as EigenTensorGait and this approach to gait analysis is called the EigenTensorGait approach, which can be extended easily to other recognition problems.

To obtain the matching score of a test sequence with N_p samples against a training sequence with N_g samples, each of the N_p samples is matched against all the N_g samples in a training sequence and the best matching score (minimum distance) is kept. The sum of the N_p best scores gives the matching score between the test and training sequences.

4. Experimental results

To evaluate the proposed MPCA method, the EigenTensorGait approach for gait recognition is tested on the USF's Gait Challenge data sets version 1.7 [9]. The human gait sequences in these data sets were captured under different conditions (walking surfaces, shoe types and viewing angles). The gallery set was used as the training set and the probe sets were test sequences containing sequences of unknown subjects to be identified. There are seven probe sets (A to G) available and 71 subjects in the gallery set. The capturing conditions of the probe sets are summarized in brackets in Table 1, where C,G,A,B,L,R, standing for cement surface, grass surface, shoe type A, shoe type B, left view, and right view respectively. The capturing condition of the gallery set is GAR.

The silhouette images are shrunk to half size of 64×44 and the half cycles are normalized temporally to 20 frames. Thus, each sample is a $64 \times 44 \times 20$ tensor, which is of size $56,320$ if vector representation is used. To find the combination of (R_1, R_2, R_3) in HOSVD truncation with the best performance, an exhaustive testing is not practical. The parameter space is sampled sparsely first and regions with good performance are then sampled densely. The best results obtained are with $R_1 = 19$, $R_2 = 11$ and $R_3 = 1$, and the projected feature is a $19 \times 11 \times 1$ third-order tensor.

The cumulative match characteristics (CMCs) [9] are used for performance measurement. The identification rates (P_I) at rank 1 and rank 5 are listed in Table 1 in comparison to the baseline algorithm. Rank k results report the percentage of probe subjects whose true match in the gallery set was in the top k matches. The CMC curves are depicted in Fig. 2. Comparison to the baseline results shows that the MPCA-based algorithm, called EigenTensorGait approach, achieves better overall recognition rate.

Table 1. MPCA identification performance.

Probe	P_I (%) at Rank 1		P_I (%) at Rank 5	
	Baseline	MPCA	Baseline	MPCA
A (GAL)	79	94	96	99
B (GBR)	66	76	81	83
C (GBL)	56	66	76	81
D (CAR)	29	27	61	64
E (CBR)	24	36	55	52
F (CAL)	30	15	46	53
G (CBL)	10	19	33	48
Average	42	48	64	68

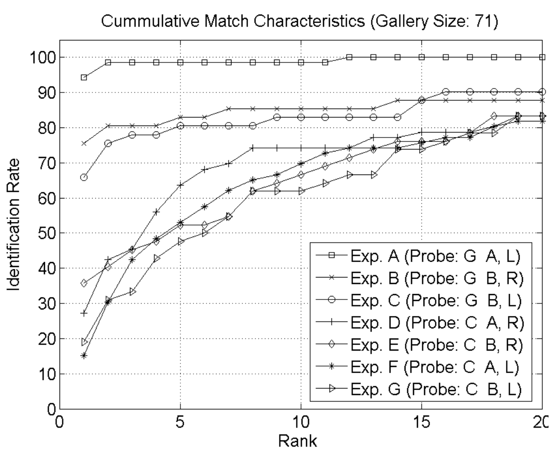


Figure 2. MPCA CMC curves.

5. Conclusions and future work

This paper proposes a new Multilinear Principal Component Analysis algorithm that is a natural multilinear extension of PCA, where multidimensional objects are represented naturally as higher-order tensors. MPCA subsumes PCA, which can be viewed as MPCA with $N = 2$.

By using a natural and simple representation for gait silhouettes as 3rd-order tensors, MPCA is applied to gait recognition as an example, called the EigenTensorGait approach. Gait silhouette sequences are partitioned into half-cycles, each of which is treated as a data sample. The best rank approximation is used to reduce noise and half cycles are normalized spatially and temporally for recognition. Experimental results show that this approach outperforms the baseline algorithm on the Gait Challenge data sets.

The proposed tensor representation of half gait cycles as data samples allows us to apply techniques popular in other recognition tasks, such as face recognition, to gait recognition. Future work also includes applying MPCA to other problems and exploring more multilinear extension of traditional pattern recognition methods.

References

[1] B. W. Bader and T. G. Kolda. Matlab tensor classes for fast algorithm prototyping. *Technical Report SAND2004-5187*, Oct. 2004.

[2] J. W. Davis and H. Gao. An expressive three-mode principal components model for gender recognition. *Journal of Vision*, (4):362–377, 2004.

[3] X. He, D. Cai, and P. Niyogi. Tensor subspace analysis. In *Advances in Neural Information Processing Systems 18 (NIPS)*, 2005.

[4] P. Kroonenberg. *Three-mode principal component analysis: theory and applications*. Leiden: DSWO Press, 1983.

[5] L. D. Lathauwer, B. D. Moor, and J. Vandewalle. A multilinear singualr value decomposition. *SIAM Journal of Matrix Analysis and Applications*, 21(4):1253–1278, 2000.

[6] 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*, 21(4):1324–1342, 2000.

[7] C. S. Lee and A. Elgammal. Gait style and gait content: bilinear models for gait recognition using gait re-sampling. In *Proc. IEEE Int. Conf. on Automatic Face and Gesture Recognition*, pages 147–152, May 2004.

[8] C. S. Lee and A. Elgammal. Towards scalable view-invariant gait recognition: Multilinear analysis for gait. In *Proc. Int. Conf. on Audio and Video-Based Biometric Person Authentication*, pages 395–405, July 2005.

[9] 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.*, 27(2):162–177, Feb. 2005.

[10] J. B. Tenenbaum and W. T. Freeman. Seperating style and content with bilinear models. *Neural Computation*, (12):1247–1283, 2000.

[11] M. A. O. Vasilescu and D. Terzopoulos. Multilinear analysis of image ensembles: Tensorfaces. In *Proc. 7th European Conference on Computer Vision*, pages 447–460, May 2002.

[12] M. A. O. Vasilescu and D. Terzopoulos. Multilinear independent components analysis. In *Proc. IEEE Computer Society Conf. on Computer Vision and Pattern Recognition*, volume I, pages 547–553, June 2005.

[13] J. Yang, D. Zhang, A. F. Frangi, and J. Yang. Two-dimensional pca: a new approach to appearance-based face representation and recognition. *IEEE Trans. Pattern Anal. Machine Intell.*, 26(1):131–137, Jan. 2004.