MPCA+DATER: A Novel Approach for Face Recognition Based on Tensor Objects

Ali. A. Shams Baboli, *Member, IEEE,* G. Rezai-rad, *Member, IEEE*, Aref. Shams Baboli

Abstract —A **novel approach proposed to solve the supervised dimensionality reduction problem by using tensor objects. MPCA used for unsupervised dimension reduction. Then discriminant analysis with Tensor Representation-(DATER) is proposed to find the best subspaces. It's important that both of those algorithms work in tensor space so the structure of the objects never broke. These algorithms are avoiding the curse of dimensionality by using higher-order tensors. At the end, the comprehensive experiments are provided on CMU-PIE, FERET databases.**

***Keywords*- Tensor objects, HOSVD, discriminant analysis with Tensor Representation, multilinear principal component analysis, subspace learning, features extraction.**

I. Introduction

Typical tensor object in machine vision or pattern recognition applications is actually in a high-dimensional tensor space. The extracted feature of an object has specific structures that are in the form of second or even higher order tensors [1]. Most previous work on dimensionality reduction transform all kind of input data into a 1-D vector, which ignores their underlying structure so these methods often suffer from curse of dimensionality [2] and leads us to the *small sample size problem.* Subspace learning is one of the most important directions in computer vision research [1] Most traditional algorithms, such as and *linear discriminant analysis* (LDA) [3] the *principal component analysis* (PCA) [4] input an image object as a 1-D vector. It is well understood that reshaping breaks the natural structure and correlation in the original data. Some recent works have started to consider an image as a second-order tensors rather than vectors first-order tensors for subspace learning. A 2-D PCA algorithm is proposed in [5] where gets the input images as a matrix and compute a covariance matrix. The classical LDA aims to maximize Fisher's discrimination criterion (FDC). In this paper as we mentioned before a method that uses the DATER after MPCA algorithms has been proposed in which both of those algorithms get the tensor objects that will give us the better accuracy.

Recently, there are many developments in the analysis of higher order tensors. The higher order singular value decomposition (HOSVD) solution was formulated in [6], [7]. A DTC algorithm is similar to an alternating least square (ALS) algorithm [8] for the eigen decomposition in each mode so it will be affected by the average of the data set. Reference [9] used a MPCA method based on HOSVD. There is also a recent work on discriminant analysis with Tensor Representation (DATER) in [10], [11]. Motivated by these two works first, we use MPCA algorithm for tensor image feature extraction and dimensionality reduction. MPCA is a multilinear algorithm that is reduced dimension in all modes and in each mode finds those bases that allow projected tensors to achieve most of the original tensors variation. After finding the best projection matrices, these bases are applied on input tensors and a new data set with a new size will be generated. This new data set will be the inputs of a DATER algorithm. DATER uses a novel criterion for dimensionality reduction, *discriminant tensor criterion* (DTC), which maximizes the interclass scatter and simultaneously time minimizes the intraclass scatter. After using DTC, a method to learn these interrelated discriminate subspaces iteratively with a novel *n-mode cluster-based discriminant analysis* approach is used.

As we know, MPCA+DATER avoid the *curse of dimensionality dilemma* by using higher order tensor for objects and *n-mode cluster-based discriminant analysis* approach. Due to using the DATER after applying the MPCA, this method is performed in a much lower-dimension feature space than DATER and the traditional vector-based methods. Also because of the structure of DATER, our approach can overcome the *small sample size problem*. From our knowledge, the available feature dimension of LDA is theoretically limited by the number of available classes in database but in our algorithm it is not limited. So it will give us the better recognition accuracy. As a result of all the above characteristics, we expect this novel method to be a better choice than LDA and more general algorithm than DATER for the pattern classification problems in image analysis and also overcome the small sample sizes and curse of dimensionality dilemma.

The rest of this paper is organized as follows. Section II introduces basic multilinear algebra notations and concepts. In Section III, the Initialization procedures of MPCA and introducing the DTC and *n-mode cluster-based discriminant analysis* that are used in DATER is discussed. Then, in Section IV, we present the face recognition experiments by encoding the image objects as

Ali Akbar Shams Baboli is with department of Electrical Engineering, Iran University of Science and Technology, Iran, Tehran, e-mail: ali.shams2222@gmail.com.

Gholamali rezai-rad is with department of Electrical Engineering, Iran University of Science and Technology, Iran, Tehran, e-mail: rezai@iust.ac.ir.

Aref Shams Baboli is with department of Electrical and computer Engineering, Noshirvani University of Technology, Iran, mazandaran, Babol, Aref.shams2222@gmail.com.

second or third-order tensors and compare them to traditional subspace learning algorithms and DATER algorithm. Finally, in Section V, the major point of this paper and the future work is summarized.

II. Multilinear Notations and Basic Multilinear Algebra

This section will review some basic multilinear concepts used in this paper and see an example for *n–mode unfolding* of a tensor. In this paper, vectors are denoted by lowercase boldface letters, such as, $\mathbf{x}$, $\mathbf{y}$. The bold uppercase symbols are used for representing matrices, such as $\mathbf{U}$, $\mathbf{S}$, and tensors by calligraphic letters, e.g. $\mathcal{A}$. An Nth-order tensor is denoted as $\mathcal{A} \in \mathbb{R}^{I_1 \times I_2 \times \ldots \times I_N}$. 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 $\mathbf{U}$, denoted by $\mathcal{A} \times_n \mathbf{U}$ is $(\mathcal{A} \times_n \mathbf{U})(i_1, \ldots, i_{n-1}, j_n, i_{n+1}, \ldots, i_N) = \sum_{i_n} \mathcal{A}(i_1, \ldots, i_N).\mathbf{U}(j_n, i_n)$.

The scalar product of two tensors $\mathcal{A}, \mathcal{B} \in \mathbb{R}^{I_1 \times I_2 \times \ldots \times I_N}$ is defined as $\prec \mathcal{A}, \mathcal{B} \succ = \sum_{i_1} \sum_{i_2} \ldots \sum_{i_N} \mathcal{A}(i_1, \ldots, i_N) . \mathcal{B}(i_1, \ldots, i_N)$and the Frobenius norm of $\mathcal{B}$ is defined as$\|\mathcal{B}\|_F = \sqrt{\prec \mathcal{B}, \mathcal{B} \succ}$ [6], [7].

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. Unfolding along the *n*-mode is denoted as $\mathbf{A}_{(n)} \in \mathbb{R}^{I_n \times (I_1 \times \ldots \times i_{n-1} \times i_{n+1} \times \ldots \times i_N)}$. The column vectors of $\mathbf{A}_{(n)}$ are the *n*-mode vectors of $\mathcal{A}$. "Fig. 1" illustrates three ways to unfold a third-order tensor. For unfolding along the first-mode, a tensor is unfolded into a matrix along the I_1 axis, and the matrix width direction is indexed by searching index I_2 and I_3 index iteratively. In the second-mode, the tensor is unfolded along the I_2 axis and the same trend afterwards.

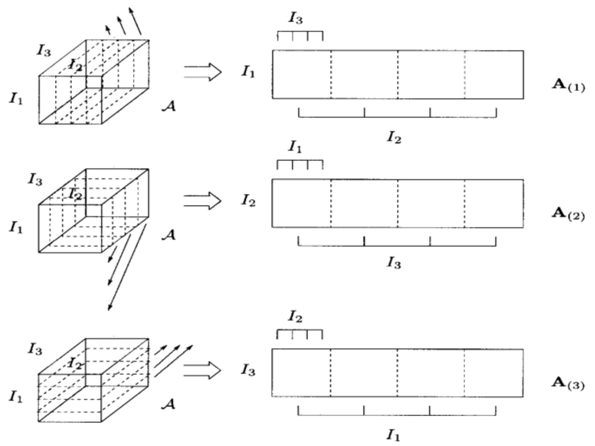


Fig. 1. Illustration of the *n-mode unfolding* of a third –tensor

Following standard multilinear algebra, tensor $\mathcal{A}$ can be expressed as the product

$$\mathcal{A} = \mathcal{S} \times_1 \mathbf{U}^{(1)} \times_2 \mathbf{U}^{(2)} \times \ldots \times_N \mathbf{U}^{(N)} \tag{1}$$

Where $\mathcal{S} = \mathcal{A} \times_1 \mathbf{U}^{(1)^{\mathrm{T}}} \times_2 \mathbf{U}^{(2)^{\mathrm{T}}} \times \ldots \times_N \mathbf{U}^{(N)^{\mathrm{T}}}$ and we call $\mathcal{S}$ core tensor that will be used for HOSVD and $\mathbf{U}^{(n)} = \left(\mathbf{u}_1^{(n)} \mathbf{u}_2^{(n)} \ldots \mathbf{u}_{I_n}^{(n)}\right)$ is an orthogonal $I_n \times I_n$ matrix.

Since $\mathbf{U}^{(n)}$ have orthonormal columns, $\|\mathcal{A}\|_F{}^2 = \|\mathcal{S}\|_F{}^2$ [7]. The relationship between unfolded tensor $\mathbf{A}_{(n)}$ and its decomposition core tensor $\mathcal{S}_{(n)}$ is

$$\mathbf{A}_{(n)} = \mathbf{U}^{(n)} . \mathcal{S}_{(n)} . (\mathbf{U}^{(n+1)} \otimes \mathbf{U}^{(n+2)} \otimes \ldots \otimes \mathbf{U}^{(N)} \otimes \mathbf{U}^{(1)} \otimes \mathbf{U}^{(2)} \otimes \ldots \otimes \mathbf{U}^{(n-1)})^{\mathrm{T}} \tag{2}$$

Where $\otimes$ means the Kronecker product [6].

The projection of an n-mode vector of $\mathcal{A}$ by $\mathbf{U}^{(n)^{\mathrm{T}}}$is computed as the inner product between the n-mode vector and the rows of $\mathbf{U}^{(n)^{\mathrm{T}}}$. For example in "Fig. 2", a third-order tensor $\mathcal{A} \in \mathbb{R}^{I_1 \times I_2 \times I_3}$ is projected in the 1-mode vector space by a projection matrix $\boldsymbol{\mathcal{B}}^{(1)^{\mathrm{T}}} \in \mathbb{R}^{m_1 \times I_1}$, the projected tensor is $\mathcal{A} \times_1 \boldsymbol{\mathcal{B}}^{(1)^{\mathrm{T}}} \in \mathbb{R}^{m_1 \times I_2 \times I_3}$. In the 1-mode projection, each 1-mode vector of length I_1 is projected by $\boldsymbol{\mathcal{B}}^{(1)^{\mathrm{T}}}$ to obtain a vector of length m_1.

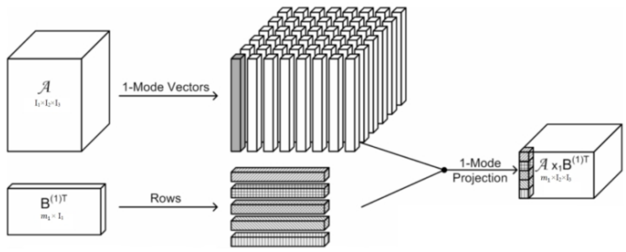
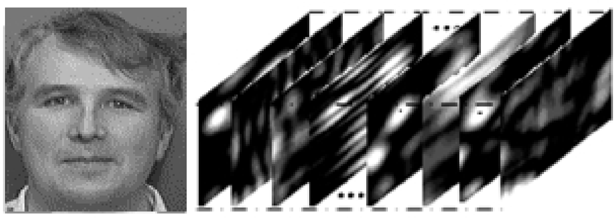


Fig. 2. Illustration of multilinear projection in the 1 – mode vector space

III. Multilinear Pricipal Component Analysis and Discriminant Analysis with Tensor Representation

Most of the previous approaches to subspace learning, such as the popular PCA and LDA turn the input data into a 1-D vector so the learning algorithms should apply on a very high dimension feature space. It makes these methods to suffer from the problem of *curse of dimensionality*. Most of the objects in computer vision are more naturally represented as second or third order tensors or higher. For example, the image matrix in "Fig. 3(a)" is a second-order tensor and the filtered Gabor image in "Fig. 3(b)" is a third-order tensor.



(a) 2nd-order Tensor (b) 3rd-order Gabor Tensor

Fig 3. Second- and third-order Tensor representations samples

In this section, first we see the MPCA solution for tensor objects and then we will see the DTC and *n-mode cluster-based discriminant analysis* that is used in DATER for tensor objects. M tensor objects $\{\mathcal{X}_1, \mathcal{X}_2, \ldots, \mathcal{X}_M\}$ is available for training. Each of them $\mathcal{X}_m \in \mathbb{R}^{I_1 \times I_2 \times \ldots \times I_N}$ assumes values in a tensor space $\mathbb{R}^{I_1} \otimes \mathbb{R}^{I_2} \ldots \otimes \mathbb{R}^{I_N}$, where I_n is the *n*-mode dimension of the tensor. The MPCA objective is the determination of the projection matrices $\{\mathbf{U}^{(n)} \in \mathbb{R}^{I_n \times P_n}, n = 1, \dots, N\}$ that maximize the total tensor scatter, Ψ_y,

$$\{\mathbf{U}^{(n)}, n = 1, \dots, N\} = \arg \max_{\mathbf{U}^{(1)}, \mathbf{U}^{(2)}, \dots, \mathbf{U}^{(N)}} \Psi_y \tag{3}$$

Where $\Psi_y = \sum_{m=1}^{M} \|\mathcal{A}_m - \bar{\mathcal{A}}\|_F^2$ $\bar{\mathcal{A}} = (1/m) \sum_{m=1}^{M} \mathcal{A}_m$. The dimensionality P_n for each mode $\mathbf{U}^{(n)}$, is assumed to be known or predetermined.

A. MPCA Algorithm

There is no optimal solution for optimizing the N projection matrices at the same time. An Nth-order tensor consists of N projections with N matrix, so N optimization subproblems should be solved by finding the $\mathbf{U}^{(n)}$ that maximizes the scatter in the n-mode vector subspace. If $\{\mathbf{U}^{(n)}, n = 1, \dots, N\}$ be the answer of (3) and $\mathbf{U}^{(1)}, \dots, \mathbf{U}^{(n-1)}, \mathbf{U}^{(n+1)}, \dots, \mathbf{U}^{(N)}$ be all the other known projection matrices, the matrix $\mathbf{U}^{(n)}$ consists of the P_n eigenvectors corresponding to the largest eigenvalues of the matrix $\Phi^{(n)}$

$$\Phi^{(n)} = \sum_{m=1}^{M} \left(\mathbf{X}_{m(n)} - \bar{\mathbf{X}}_{(n)}\right). \mathbf{U}_{\Phi^{(n)}}. \mathbf{U}_{\Phi^{(n)}}^T . \left(\mathbf{X}_{m(n)} - \bar{\mathbf{X}}_{(n)}\right)^T \tag{4}$$

Where $\mathbf{U}_{\Phi^{(n)}} = \mathbf{U}^{(n+1)} \otimes \mathbf{U}^{(n+2)} \otimes \dots \otimes \mathbf{U}^{(N)} \otimes \mathbf{U}^{(1)} \otimes \mathbf{U}^{(2)} \otimes \dots \mathbf{U}^{(n-1)}$. The proof of (4) is given in [9].

$\Phi^{(n)}$ depends on all the other projection matrices, so there is no closed-form solution to this maximization problem. Instead, reference [9] introduce an iterative procedure that can be utilized to solve (3). MPCA used full projection for initialization. The term full projection refers to the multilinear projection for MPCA with $P_n = I_n$ for $n = 1, \dots, N$. There is no dimensionality reduction with this full projection. The optimal is obtained without any iteration, and the total scatter in the original data is fully captured. if all eigenvalues are distinct per each mode, the full projection matrices are also distinct. Therefore, the full projection is unique [7]. This FPT initialization is equivalent to the HOSVD solution that is proofed in [9]. Although this FPT is not the optimal solution to (3), but it is a good starting point for the iterative procedure. After finding the projection matrices, $\mathbf{U}^{(n)}, n = 1, \dots, N$, we applied those matrices to the training set. At this point, we provide a set of tensors with the new dimension that would be the new training set for DATER algorithm.

B. Multilinear Discriminant Analysis

Here, the *DTC* is introduced which is used in DATER algorithm. The *DTC* is designed to provide multiple interrelated projection matrices, which maximize the interclass scatter and at the same time minimize the intraclass scatter. That is

$$\mathbf{U}^{(n)^*}\Big|_{n=1}^{N} = \arg \max_{\mathbf{U}^{(n)}|_{n=1}^{N}} \frac{\sum_c n_c \left\|\bar{\mathcal{X}}_c \times_1 \mathbf{U}^{(1)} \dots \times_n \mathbf{U}^{(n)} - \bar{\mathcal{X}} \times_1 \mathbf{U}^{(1)} \dots \times_n \mathbf{U}^{(n)}\right\|^2}{\sum_i \left\|\mathcal{X}_i \times_1 \mathbf{U}^{(1)} \dots \times_n \mathbf{U}^{(n)} - \bar{\mathcal{X}}_{c_i} \times_1 \mathbf{U}^{(1)} \dots \times_n \mathbf{U}^{(n)}\right\|^2}. \tag{5}$$

Where $\bar{\mathcal{X}}_c$ is the average tensor of a class c samples, $\bar{\mathcal{X}}$ is the total average tensor of all the samples, and n_c is sample number of class c. Equation (5) is equivalent to a higher order nonlinear optimization problem with a higher order nonlinear constraint; thus, it is difficult to find a closed-form solution. We could optimize that function by using *n-mode cluster-based discriminant analysis* approach from only one direction of the tensor, that is

$$\mathbf{U}^{(n)^*} = \arg \max_{\mathbf{U}^{(n)}} \frac{\sum_c n_c \left\|\bar{\mathcal{X}}_c \times_n \mathbf{U}^{(n)} - \bar{\mathcal{X}} \times_n \mathbf{U}^{(n)}\right\|^2}{\sum_i \left\|\mathcal{X}_i \times_n \mathbf{U}^{(n)} - \bar{\mathcal{X}}_{c_i} \times_n \mathbf{U}^{(n)}\right\|^2} \tag{6}$$

The optimization problem in (6) can be reformulated as a special discriminant analysis problem as follows:

$$\mathbf{U}^{(n)^*} = \arg \max_{\mathbf{U}^{(n)}} \frac{Tr(\mathbf{U}^{(n)^T} S_B \mathbf{U}^{(n)})}{Tr(\mathbf{U}^{(n)^T} S_W \mathbf{U}^{(n)})} \tag{7}$$

$$S_B = \sum_{j=1}^{\prod_{o \neq n} m_o} S_B^j, S_B^j = \sum_{c=1}^{N_c} n_c (\bar{\mathbf{X}}_{c(n)}^{j} - \bar{\mathbf{X}}_{(n)}^{j})(\bar{\mathbf{X}}_{c(n)}^{j} - \bar{\mathbf{X}}_{(n)}^{j})^T$$

$$S_W = \sum_{j=1}^{\prod_{o \neq n} m_o} S_W^j, S_W^j = \sum_{i=1}^{N_c} (\bar{\mathbf{X}}_{i(n)}^{j} - \bar{\mathbf{X}}_{c_i(n)}^{j})(\bar{\mathbf{X}}_{i(n)}^{j} - \bar{\mathbf{X}}_{c_i(n)}^{j})^T$$

C. Classification with Multilinear Discriminant Analysis

With the learned projection matrices $\mathbf{U}_k^*|_{k=1}^{n}$, the low-dimensional representation of the training sample $\mathcal{X}_i$, $i = 1, \dots N$, can be computed as $\mathcal{Z}_i = \mathcal{X}_i \times_1 \mathbf{U}_1 \times_2 \mathbf{U}_2 \times \dots \times_n \mathbf{U}_n$ .When a new data $\mathcal{X}$ comes, we first compute its low-dimensional representation as $\mathcal{Z} = \mathcal{X} \times_1 \mathbf{U}_1 \times_2 \mathbf{U}_2 \times \dots \times_n \mathbf{U}_n$,Then its class label is predicted to be that of the sample whose low-dimensional representation is nearest to $\mathcal{Z}$, that is

$$i^* = arg\min_i \|\mathcal{Z}_i - \mathcal{Z}\| \tag{8}$$

Then the sample $\mathcal{X}$ is classified to the class c_{i^*}. In this paper, we use this method for final classification in all the experiments due to its simplicity in computation.

IV. Experiments

In this section, two standard face databases, CMU PIE [12], FERET [13] were used to evaluate the effectiveness of our proposed algorithm, MPCA+DATER, in face recognition accuracy. These algorithms were compared with the popular Eigenface, Fisherface and DATER algorithms. In this work, we report the best result on different test. We use the nearest neighbor classifier for final classification. The performances on the cases with different number of training samples were also evaluated to illustrate their robustness in the small sample size problems.

A. Feret Database

This Experiment is conducted on seventy people of the FERET database with six different images for each person. The data set was randomly divided into gallery and probe sets; two of them were applied for training and the other four for testing. We extracted 40 Gabor features with five different scales and eight different directions in the down-sampled positions. We compared all the above mentioned algorithms with and without Gabor features on the FERET database. Table 1 shows the comparative face recognition accuracies. As we see, it shows that the Gabor features significantly improve the performance of all algorithms.

TABLE 1: RECOGNITION ACCURACY (%) COMPARISON OF MPCA+DATER, EIGENFACE, FISHERFACE, DATER/2-1, DATER/2-2 AND DATER/3-3 ON THE FERET DATABASE

Algorithms	Accuracy
Eigenface (Grey)	65.70
Eigenface (Gabor)	75.70
Fisherface (Grey)	74.28
Fisherface (Gabor)	76.07
DATER/2-1 (Grey)	73.50
DATER/2-2 (Grey)	80.7
DATER/3-3 (Gabor)	83.57
MPCA+DATER	86.78

B. CMU PIE

The CMU PIE database contains more than 40,000 facial images of 68 people. The images were obtained over different poses, under variable illumination conditions and with different facial expressions. In our experiment, two sub-databases were used to evaluate our methods. In the first sub-database, PIE-1, five near frontal poses (C27, C05, C29, C09 and C07) and illumination indexed as 08 and 11 were used. The data set was randomly divided into training and test sets; and two samples per person was used for training. We extracted 40 Gabor features. Table 2 shows the detailed face recognition accuracies. The results clearly demonstrate that MPCA+DATER is superior to all other algorithms.

Table 2: Recognition Accuracy (%) Comparison of Eigenface, Fisherface, DATER and MPCA+DATER with tensors of different orders on PIE-1 Database

Algorithms	Accuracy
Eigenface (Grey)	57.14
Eigenface (Gabor)	70.47
Fisherface (Grey)	67.85
Fisherface (Gabor)	75.95
DATER/2-1 (Grey)	72.85
DATER/2-2 (Grey)	80.35
DATER/3-3 (Gabor)	83.57
MPCA+DATER	87.14

Another sub-database PIE-2 consists of the same five poses as in PIE-1, but the illumination indexed as 10 and 13 were also used. Therefore, the PIE-2 database is more difficult for classification. We conducted three sets of experiments on this sub-database.as we can see in Table 3 In all the three experiments, MPCA+DATER performs the best.

TABLE 3: RECOGNITION ACCURACY (%) COMPARISON OF MPCA+DATER, EIGENFACE, FISHERFACE, DATER/2-1 AND DTAER/2-2 ON THE PIE-2 DATABASE

Algorithms	Test-Train		
	4-6	***3-7***	***2-8***
Eigenface	39.21	28.15	27.2
Fisherface	79.90	65.96	47.98
DATER/2-1	74.02	71.84	63.42
DATER/2-2	81.86	80.88	66.91
MPCA + DATER	84.06	82.56	68.56

V. CONCLUSION

In this paper, we apply MPCA algorithm and after that we use DATER to find the best subspaces. A MPCA framework is used for analysis of tensor objects and determines a multilinear projection onto a tensor subspace with lower dimensional that captures most of the covariance of the original tensorial objects. After that projection, a DATER has been applied on a new database for supervised dimensionality reduction. Compared with traditional algorithms, such as PCA and LDA, our proposed algorithm effectively avoids the curse of dimensionality dilemma and overcome the small sample size problem. Since the low requirement on samples and the higher performance in classification problem, our method should be a general alternative of LDA algorithm for problems encoding objects as tensors. We are eager to apply this algorithm for video-based (fourth order tensor) face recognition and we are going to explore this work in our future researches.

REFERENCES

[1] X. He, D. Cai, and P. Niyogi, "Tensor subspace analysis," *in Advances in Neural Information Processing Systems 18 (NIPS).* Cambridge, MA: MIT Press, 2005.

[2] G. Shakhnarovich and B. Moghaddam, "Face recognition in subspaces," *in Handbook of Face Recognition*, S. Z. Li and A. K. Jain, Eds. New York: Springer-Verlag, 2004, pp. 141–168.

[3] P. Belhumeur, J. Hespanha, and D. Kriegman, "Eigenfaces vs. Fisherfaces: Recognition using class specific linear projection," *IEEE Trans. Pattern Anal*. Mach. Intell., vol. 19, no. 7, pp. 711–720, Jul. 1997.

[4] M. Turk and Pentland, "Face recognition using eigenfaces," *presented at the IEEE Conf. Computer Vision and Pattern Recognition*, Maui, HI, 1991.

[5] J. Yang, D. Zhang, A. Frangi, and J. Yang, "Two-dimensional PCA: A new approach to appearance-based face representation and recognition," *IEEE Trans. Pattern Anal*. Mach. Intell., vol. 26, no. 1, pp. 131–137, Jan. 2004.

[6] L. D. Lathauwer, B. D. Moor, and J. Vandewalle, "A multilinear singular value decomposition," *SIAM J. Matrix Anal*. Appl., vol. 21, no. 4, pp. 1253–1278, 2000.

[7] T. Kolda, "Orthogonal tensor decompositions," *SIAM J. Matrix Anal. Appl.*, vol. 23, no. 1, pp. 243–255.2000.

[8] P. Kroonenberg and J. Leeuw, "Principal component analysis of three-mode data by means of alternating least squares algorithms," *Psychometrika*, vol. 45, no. 1, pp. 69–97, 1980.

[9] H. Lu, K. N. Plataniotis, and A. N. Venetsanopoulos, "MPCA: Multilinear Principal Component Analysis of Tensor Objects ," *IEEE Trans. Neural Networks*, no. 1, vol. 19, pp. 18-39 Aug. 2008.

[10] S. Yan, D. Xu, Q. Yang, L. Zhang, X. Tang, and H.-J. Zhang, "Discriminant analysis with tensor representation," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, Jun. 2005, vol. I, pp. 526–532.

[11] L. Haiping, K.N. Plataniotis, and A.N. Venetsanopoulos, "Uncorrelated Multilinear Discriminant Analysis With Regularization and Aggregation for Tensor Object Recognition" *IEEE Trans on Neural Networks,* Vol. 20, no. 1 , pp. 103-123, 2009.

[12] T. Sim, S. Baker, and M. Bsat. "The CMU Pose, Illumination, and Expression (PIE) Database", *Proceedings of the IEEE International Conference on Automatic Face and Gesture Recognition,* May, 2002.

[13] I. Philips, H.Wechsler, J. Huang, and P. Rauss, "The FERET database and evaluation procedure for face recognition algorithms," *Image Vis. Computer*, vol. 16, pp. 295–306, 1998.