A New Method Based on MDA to Enhance the Face Recognition Performance

Aref Shams Baboli Aref.shams2222@gmail.com
BSc *Student at Department of Electrical and computer Engineering, Noshirvani University of Technology, mazandaran, Iran*

Seyyedeh Maryam Hosseyni Nia maryamhosseini16@yahoo.com
MSc Student at Department of Biomedical Engineering and Medical Physics Shahid Beheshti University, Tehran, Iran

Ali Akbar Shams Baboli ali.shams2222@gmail.com
MSc Student at Department of Electrical Engineering, Iran University of Science and Technology, Tehran, Iran

Gholamali Rezai Rad rezai@iust.ac.ir
Associate professor at Department of Electrical Engineering, Iran University of Science and Technology, Tehran, Iran

Abstract

A novel tensor based method is prepared to solve the supervised dimensionality reduction problem. In this paper a multilinear principal component analysis (MPCA) is utilized to reduce the tensor object dimension then a multilinear discriminant analysis (MDA), is applied to find the best subspaces. Because the number of possible subspace dimensions for any kind of tensor objects is extremely high, so testing all of them for finding the best one is not feasible. So this paper also presented a method to solve that problem, the main criterion of algorithm is not similar to Sequential mode truncation (SMT) and full projection is used to initialize the iterative solution and find the best dimension for MDA. This paper is saving the extra times that we should spend to find the best dimension. So the execution time will be decreasing so much. It should be noted that both of the algorithms work with tensor objects with the same order so the structure of the objects has been never broken. Therefore the performance of this method is getting better. The advantage of these algorithms is avoiding the curse of dimensionality and having a better performance in the cases with small sample sizes. Finally, some experiments on ORL and CMPU-PIE databases are provided.

Keywords: Dimensionality Reduction, *HOSVD*, Subspace Learning, Multilinear Principal Component Analysis, Multilinear Discriminant Analysis.

1. INTRODUCTION

A typical tensor object in machine vision or pattern recognition applications is actually in a high-dimensional tensor space. In reality, the extracted features of an object often has some specific structures that are in the form of second or even higher order tensors [1]. Most previous works transform the input image data into a 1-D vector, which ignores the underlying data structure so these methods suffer from *curse of dimensionality* and *small sample size problem.* Subspace learning is one of the most important directions in computer vision research [2], [3]. Most traditional algorithms, such as LDA [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 object as a 2-D matrix rather than vectors 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. As we mentioned before, in this paper a method that utilized the MDA after MPCA algorithms has been proposed in which both of those algorithms work with tensor objects that give us the better results.

It should be noted that recently there are many developments in the analysis of higher order. Reference [6] used a MPCA method based on HOSVD [7]. There is also a recent work on multilinear discriminant analysis (MDA) in [8] which is used for maximizing a tensor-based discriminant criterion. Previously, we proposed MPCA+MDA [9] for face recognition. In that paper we use MPCA algorithm for tensor object feature extraction. MPCA is a multilinear algorithm reducing dimension in all tensor modes to find those bases in each mode that allows projected tensors to achieve most of the original tensors variation. Then these bases are applied on samples and a new data set with a new dimension will be generated. This new data set will be the inputs of our MDA algorithm. MDA uses a novel criterion for dimensionality reduction, *discriminant tensor criterion* (DTC), which maximizes the interclass scatter and simultaneously time minimizes the intraclass scatter. In that paper we should give the goal dimension for reduction manually. As we know, the number of possible subspace dimensions for tensor objects is extremely high, so testing all of them to find the best one is not feasible. To solve that problem, a method is used to find the best dimension that gives us the best accuracy. Our method is a little similar to SMT that is used in MPCA algorithm [5]. To start the algorithm like SMT we need to initialization the subspaces. So this paper is used full projection to initialize the iterative solution for MDA [6]. The main idea of this paper is saving the extra times that we should spend to find the best dimension and of course the final dimension in MDA that we find practically is not optimal. But with our improvement we are decreasing the execution time so much.

MPCA+Improved MDA can avoid the *curse of dimensionality dilemma* by using higher order tensor for objects and *n-mode optimization* approach. Due to using the MDA after applying the MPCA, this method is performed in a much lower-dimension feature space than MDA and the traditional vector-based methods, such as LDA and PCA do. Also because of the structure of MDA, it can overcome the *small sample size* problem. As we know, the available feature dimension of LDA is theoretically limited by the number of classes in the data set but in our algorithm it is not limited. So it would 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 PCA algorithms and more general than MDA 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 2 introduces basic multilinear algebra notations and concepts. In Section 3, the Initialization procedures of MPCA and introducing the DTC and *n-mode optimization* that are used in MDA is discussed after that we will see our proposed method for finding the best subspaces dimension. Then, in Section 4, we present the face recognition experiments by encoding the image objects as second or third-order tensors and compare them to traditional subspace learning algorithms and MDA algorithm. Finally, in Section 5, the major point of this paper and the future work is summarized.

2. MULTILINEAR NOTATIONS AND BASIC ALGEBRA

This section briefly will be reviewed some basic multilinear concepts used in our framework and see an example for *n–mode unfolding* of a tensor. Here, vectors are denoted by lowercase boldface letters, such as, **x**, **y**. The bold uppercase symbols are used for representing matrices, such as **U**, **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 $\mathbf{U}$, is

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

The scalar product of two tensors $\mathcal{A},\mathcal{B} \in R^{I_1 \times I_2 \times .. \times I_N}$ is defined as $\prec \mathcal{A},\mathcal{B} \succ = \sum_{i_1}\sum_{i_2}\cdots\sum_{i_N} \mathcal{A}(i_1,\ldots,i_N).\mathcal{B}(i_1,\ldots,i_N)$ and the Frobenius norm of $\mathcal{B}$ is defined as $\mathcal{B}_F = \sqrt{\prec \mathcal{B},\mathcal{B} \succ}$ [7].
Unfolding along the n-mode is denoted as $\mathbf{A}_{(n)} \in R^{I_n \times (I_1 \times .. \times i_{n-1} \times i_{n+1} \times .. \times i_N)}$. The column vectors of $\mathbf{A}_{(n)}$ are the n-mode vectors of $\mathcal{A}$. 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.
Following standard multilinear algebra, tensor $\mathcal{A}$ can be expressed as the product $\mathcal{A} = \mathcal{S} \times_1 \mathbf{U}^{(1)} \times_2 \mathbf{U}^{(2)} \times \ldots \times_N \mathbf{U}^{(N)}$. 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. 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 [7].
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 $\mathcal{A} \in R^{I_1 \times I_2 \times I_3}$ is projected in the 1-mode vector space by a projection matrix $\mathbf{B}^{(1)\mathrm{T}} \in R^{m_1 \times I_1}$, the projected tensor is $\mathcal{A} \times_1 \mathbf{B}^{(1)\mathrm{T}} \in R^{m_1 \times I_2 \times I_3}$. In the 1- mode projection, each 1-mode vector of length I_1 is projected by $\mathbf{B}^{(1)\mathrm{T}}$ to obtain a vector of length m_1.

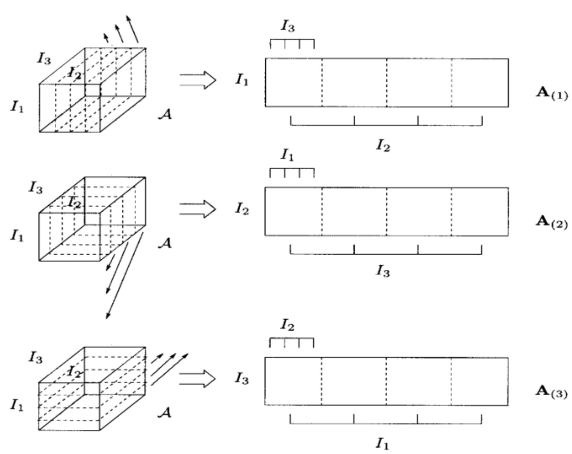

FIGURE 1: Illustration of the *n-mode unfolding* of a third–order tensor.

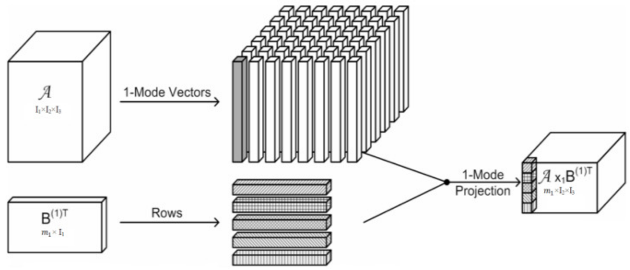

FIGURE 2: Illustration of multilinear projection in the mode 1

3. MULTILINEAR PRINCIPAL COMPONENT ANALYSIS & MULTILINEAR DISCRIMINANT ANALYSIS

Some previous approaches to subspace learning, such as PCA and LDA, consider an object as a 1-D vector so the learning algorithms should be applied on a very high dimension feature space. So these methods suffer from the problem of *curse of dimensionality*. Most of the objects in computer vision are more naturally represented as second or higher order tensors. 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.

In this section, first we see, how the MPCA solution for tensor objects is working and then we will see the DTC and *n-mode optimization* that is used in MDA for tensor objects. A set of M tensor objects $\{X_1, X_2, ..., X_M\}$ is available for training. Each tensor object $X_m \in R^{I_1 \times I_2 \times ... \times I_N}$ assumes values in a tensor space $R^{I_1} \otimes R^{I_2} \ldots \otimes R^{I_N}$, where I_n is the *n*-mode dimension of the tensor. The MPCA defines a multilinear transformation that maps the original tensor space into a tensor subspace. In other words, the MPCA objective is the determination of the projection matrices $\left\{\mathbf{U}^{(n)} \in R^{I_n \times P_n}, n = 1, \ldots, N\right\}$ that maximize the total tensor scatter, Ψ_y

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

Where $\Psi_y = \sum_{m=1}^{M} \left\| A_m - \overline{A} \right\|_F^2, \overline{A} = (1/m) \sum_{m=1}^{M} A_m$.

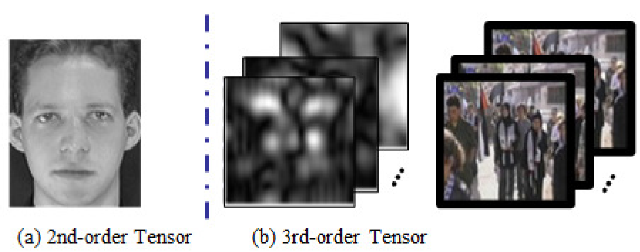

FIGURE 3: Second- and third-order Tensor representations samples

3.1 MPCA Algorithm

There is no optimal solution for optimizing the *N* projection matrices simultaneously. An *N*th-order tensor consists of N projections with *N* matrix, so *N* optimization subproblems can be solved by finding the $\mathbf{U}^{(n)}$ that maximizes the scatter in the *n*-mode vector subspace. If $\left\{\mathbf{U}^{(n)}, n = 1, \ldots, N\right\}$ be the answer of (3) and $\mathbf{U}^{(1)}, \ldots, \mathbf{U}^{(n-1)}, \mathbf{U}^{(n+1)}, \ldots, \mathbf{U}^{(N)}$ 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 $\Phi^{(n)}$

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

Where $\mathbf{U}_{\Phi^{(n)}} = \mathbf{U}^{(n+1)} \otimes \mathbf{U}^{(n+2)} \otimes \ldots \otimes \mathbf{U}^{(N)} \otimes \mathbf{U}^{(1)} \otimes \mathbf{U}^{(2)} \otimes \ldots \mathbf{U}^{(n-1)}$.

The proof of (4) is given in [6].

Since $\Phi^{(n)}$ depends on all the other projection matrices, there is no closed-form solution for this maximization problem. Instead, reference [6] introduce an iterative procedure that can be utilized to solve (4). For initialization, MPCA used full projection. The term full projection refers to the multilinear projection for MPCA with $P_n=I_n$ for n= 1, …, N. There is no dimensionality reduction through this full projection. The optimal is obtained without any iteration, and the total scatter in the original data is fully captured. After finding the projection matrices, $\mathbf{U}^{(n)}, n = 1, \ldots, N,$ 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 MDA algorithm.

3.2 Multilinear Discriminant Analysis

Here, the DTC is introduced which is used in MDA 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

$$\mathbf{U}^{(n)*}\Big|_{n=1}^{N} = \arg\max_{\mathbf{U}^{(n)}\big|_{n=1}^{N}} \frac{\sum_c n_c \overline{\mathrm{X}}_c \times_1 \mathbf{U}^{(1)} \ldots \times_n \mathbf{U}^{(n)} - \overline{\mathrm{X}} \times_1 \mathbf{U}^{(1)} \ldots \times_n \mathbf{U}^{(n)2}}{\sum_i \mathrm{X}_i \times_1 \mathbf{U}^{(1)} \ldots \times_n \mathbf{U}^{(n)} - \overline{\mathrm{X}}_{c_i} \times_1 \mathbf{U}^{(1)} \ldots \times_n \mathbf{U}^{(n)2}} . \tag{5}$$

Where $\overline{\mathrm{X}}_c$ is the average tensor of class c samples, $\overline{\mathrm{X}}$ is the total average tensor of all the samples, and n_c is sample number of class c. We could optimize that function by using *n-mode optimization* approach that is proved in [8]. The optimization problem can be reformulated as follows:

$$\mathbf{U}^{(n)*} = \arg\max_{\mathbf{U}^{(n)}} \frac{Tr(\mathbf{U}^{(n)T} S_B \mathbf{U}^{(n)})}{Tr(\mathbf{U}^{(n)T} S_W \mathbf{U}^{(n)})} \tag{6}$$

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

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

Where, $\mathbf{X}_{i(n)}^{j}$ is the jth column vector of matrix $\mathbf{X}_{i(n)}^{j}$ which is the n-mode unfolded matrix from sample tensor X_i . $\mathbf{X}_{c(n)}^{j}$ and $\overline{\mathbf{X}}_{(n)}^{j}$ are defined in the same way as $\mathbf{X}_{i(n)}$ with respect to tensors $\overline{\mathrm{X}}_c$ and $\overline{\mathrm{X}}$ and the proofs are given in [8]. To utilizing *n-mode optimization*, first the input tensors (that are the outputs of MPCA) should be projected with all the other modes matrices and then all the new tensors are unfolded into a matrix along the *n*th-mode. Therefore, the optimization problem in (5) can be reformulated as a special discriminant analysis problem, and it can be solved in the same way for the traditional LDA algorithm [8]. Since DTC has no closed form the projection matrices can be iteratively optimized.

3.3 Determination of the Tensor Subspace Dimensionality

The target dimensionality P_n has to be determined. So the objective MPCA function should be revised to include a constraint on the favorite dimensionality reduction. The revised objective function is as follows [6]:

$$\left\{\mathbf{U}^{(n)}, P_n, n=1,\ldots,N\right\} = \arg \max_{\mathbf{U}^{(1)},\mathbf{U}^{(2)},\ldots,\mathbf{U}^{(N)},P_1,P_2,\ldots,P_N} \frac{Tr(\mathbf{U}^{(n)T} S_B \mathbf{U}^{(n)})}{Tr(\mathbf{U}^{(n)T} S_W \mathbf{U}^{(n)})} \quad \text{subject to} \frac{\prod_{n=1}^{N} P_n}{\prod_{n=1}^{N} I_n} < \Omega \tag{7}$$

Where the ratio between the reduced dimensionality and the original tensor space dimensionality is utilized to measure the amount of dimensionality reduction, and Ω is a threshold to be specified by user.

The proposed tensor subspace dimensionality determination solution is Starting with $P_n=I_n$ for all n at $t=0$, at each subsequent step $t=t+1$, this algorithm truncates, in a selected mode n, the P_n th n-mode eigenvector of the reconstructed input tensors. The truncation can be interpreted as the elimination of the corresponding P_n th n-mode slice of the total scatter tensor. For the specific mode selection, the scatter loss rate $\delta_t^{(n)}$ due to the truncation of its P_n th eigenvector is calculated for each mode. $\delta_t^{(n)}$ is defined as follows [6]:

$$\delta_t^{(n)} = \frac{\dfrac{Tr(\mathbf{U}^{(n)T} S_B \mathbf{U}^{(n)})}{Tr(\mathbf{U}^{(n)T} S_W \mathbf{U}^{(n)})}_{\Upsilon(t)} - \dfrac{Tr(\mathbf{U}^{(n)T} S_B \mathbf{U}^{(n)})}{Tr(\mathbf{U}^{(n)T} S_W \mathbf{U}^{(n)})}_{\Upsilon(t-1)}}{\left[P_n.\prod_{j=1, j\neq n}^{N} P_j\right] - \left[(P_n-1).\prod_{j=1, j\neq n}^{N} P_j\right]} = \frac{\lambda_{P_n}^{(n)}}{\prod_{j=1, j\neq n}^{N} P_j} \tag{8}$$

Where $Tr(\mathbf{U}^{(n)T} S_B \mathbf{U}^{(n)}) \big/ Tr(\mathbf{U}^{(n)T} S_W \mathbf{U}^{(n)})_{\Upsilon(t)}$ is maximizing the between class scatter and at the same time minimizing the within class scatter at step *t,* $\prod_{j=1, j\neq n}^{N} P_j$ is the amount of dimensionality reduction achieved, and $\lambda_{P_n}^{(n)}$, which is the corresponding P_n th n-mode eigenvalue, is the loss due to truncating the P_n th n-mode eigenvector. The mode with the smallest $\delta_t^{(n)}$ is selected for the step-t truncation. For the selected mode n, P_n is decreased by 1: $P_n = P_n - 1$ and $\prod_{n=1}^{N} P_n \big/ \prod_{n=1}^{N} I_n$ is tested. The truncation stops when $\prod_{n=1}^{N} P_n \big/ \prod_{n=1}^{N} I_n$ is satisfied. The term full projection refers to the multilinear projection for MDA with $P_n=I_n$ for $n=$ 1, … , N for starting the algorithm. There is no dimensionality reduction through this full projection [5]. The optimal is obtained without any iteration. As we know, if all eigenvalues (per mode) are distinct, the full projection matrices are also distinct. Therefore, the full projection is unique [6].

4. EXPERIMENTS

In this section, two standard face databases ORL [10], CMU PIE [11] were used to evaluate the effectiveness of our proposed algorithm, MPCA+Improved MDA, in face recognition accuracy. These algorithms were compared with the popular Eigenface, Fisherface and MDA/2-1, MDA/2-2, MDA/3-3 and the MPCA + MDA algorithms. In this work, we report the best result on different test and for the fisherface on different feature dimensions in the LDA step, in all the experiments, the training and test data were both transformed into lower dimensional tensors or vectors via the learned subspaces, and we use the nearest neighbour 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.

4.1 ORL Database

The ORL database includes 400 images of 40 persons. These images were captured at different times and have different expression such as open or closed eyes, smiling or nonsmiling and facial details like: glasses or no glasses. All images were in grayscale and centered with the resolution of 112*92 pixels. Ten sample images of one person in the ORL database are displayed in Figure

FIGURE 4: Ten samples of one person in the ORL face database

Four sets of experiments were managed to compare the performance of our algorithm with Eigenface, Fisherface, and MDA/2-1, MDA/2-2. In each experiment, the image set was partitioned into the test and train set with different numbers. Table 1 shows the best face recognition accuracies of all the algorithms in our experiments with different train and test set partitions. The results show that our algorithm outperforms Eigenface, Fisherface, MDA/2-1, MDA/2-2 and MPCA+MDA on all four sets of experiments, especially in the cases with a small number of training samples and also we can see the performance of MPCA+Improved MDA is the same as MPCA+MDA or even better than that. It means we provide the same performance without spending the spare time to find the best dimension. So we can say our very new proposed algorithm has the best performance and also save the spare times.

TABLE 1: Recognition Accuracy (%) Comparsion of MDA+MPCA, Eigenface, Fisherface, MDA/2-1, MDA/2-2, MPCA+MDA on ORL database

Algorithms	**Train-Test**			
	5-5	***4-6***	***3-7***	***2-8***
Eigenface	97.0	91.25	87.50	81.56
Fisherface	93.0	85.83	87.50	79.68
MDA/2-1	97.5	96.25	94.28	88.13
MDA/2-2	99.0	97.91	95.00	90.31
MPCA + MDA	99.0	98.75	96.43	91.56
MPCA + Improved MDA	99.0	98.75	96.78	91.87

4.2 CMU PIE Database

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 II shows the detailed face recognition accuracies. The results clearly demonstrate that MPCA+Improved MDA is superior to all other algorithms. As we knew, this database is really hard for algorithms and most of them had a problem with that. As we can see, our algorithms perform a really good job here and have the most accuracy and also work faster than the others, especially the Improved algorithm, MPCA+Improved MDA, because it eliminates the extra time that we should spend to find the best dimension.

TABLE 2: Recognition Accuracy (%) Comparison of Eigenface, Fisherface, MDA, MPCA+MDA, of MPCA+ Improved MDA with tensors of different orders on PIE-1 Database

Algorithms	Accuracy
Eigenface (Grey)	57.2
Eigenface (Gabor)	70.5
Fisherface (Grey)	67.9
Fisherface (Gabor)	76
MDA/2-1 (Grey)	72.9
MDA/2-2 (Grey)	80.4
MDA/3-3 (Gabor)	83.6
MPCA+MDA	87.2
MPCA + Improved MDA	87.5

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 + Improved MDA performs the best and the eigenface has the worst performance. Especially in the cases with a small number of training samples. Also for gaining that performance from our algorgorithm we don't have to spend much time that we use for MPCA+MDA and because of that privilege, our algorithm became a great algorithm to choose.

TABLE 3: Recognition Accuracy (%) Comparison of MPCA+ Improved MDA, MPCA+MDA, Eigenface, Fisherface, MDA/2-1 and MDA/2-2 on the PIE-2 Database

Algorithms	Test-Train		
	4-6	**3-7**	***4-6***
Eigenface	39.3	Eigenface	39.3
Fisherface	79.9	Fisherface	79.9
MDA/2-1	74.1	MDA/2-1	74.1
MDA/2-2	81.9	MDA/2-2	81.9
MPCA + MDA	84.1	MPCA + MDA	84.1
MPCA + Improved MDA	84.5	MPCA + Improved MDA	84.5

5. CONCLUSION

In this paper, we improve the performance of MPCA + MDA algorithm by optimizing the subspaces dimension and full projection. Full projection is utilized for initialization the changed SMT and the changed SMT is used to find the optimal subspaces dimension. After that, MDA has been applied 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 and the advantage of this work is finding the subspaces dimension Because in MDA algorithm the number of possible subspace dimensions for tensor objects is extremely high, comprehensive testing for determination of parameters is not feasible so with this work we save that amount of time. We are eager to apply this algorithm for video-based (fourth order tensor) face recognition and we want to explore this work in our future researches.

REFERENCES

[1]. M. Vasilescu and D. Terzopoulos, “Multilinear subspace analysis for image ensembles,” in Proc. Computer Vision and Pattern Recognition, Madison, WI, Jun. 2003, vol. 2, pp. 93–99.

[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]. K. Fukunaga, Statistical Pattern Recognition. New York: Academic, 1990.

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

[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]. 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.

[7]. 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.

[8]. S. Yan, D. Xu, Q. Yang, L. Zhang, X. Tang, and H.-J. Zhang, “Multilinear discriminant analysis for face recognition,” IEEE Trans. Image Process., vol. 16, no. 1, pp. 212–220, Jan. 2007.

[9]. A. Shams baboli and G. Rezai-rad, “MPCA+MDA: A Novel approach for Face Recognition Based on Tensor,” presented at the 5th Int. Symposium on Telecommunication, Tehan, Iran, 2010.

[10]. The Olivetti & Oracle Research Laboratory Face Database of Faces, 2002.

[11]. 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.