Pattern Recognition Letters 31 (2010) 171–178

Contents lists available at ScienceDirect

Pattern Recognition Letters

journal homepage: www.elsevier.com/locate/patrec

Independent components extraction from image matrix

Quanxue Gao [a,b,*], Lei Zhang [b], David Zhang [b], Hui Xu [a]

[a] State Key Laboratory on Integrated Serves Networks, XIDIAN University, 710071 XI'AN, China
[b] Biometric Research Center, The Hong Kong Polytechnic University, Hong Kong, China

ARTICLE INFO

Article history:
Received 13 February 2007
Received in revised form 9 October 2009
Available online 23 October 2009

Communicated by A. Shokoufandeh

Keywords:
Independent component analysis
Directional image
Feature extraction
Face recognition

ABSTRACT

The key problem of extracting independent components (ICs) is to learn the demixing matrix from the known training images which can be unfolded to vectors in conventional independent component analysis (ICA). However, the unfolded vectors lead to the small sample size problem (SSS) and the curse of dimensionality. In this paper, a novel independent feature extraction method is proposed to solve these problems by encoding each input image as a matrix. In addition, the row and column directional images of the matrix are introduced to better exploit the spatial and structural information embedded in image during the training phase. Compared with the conventional ICA, the proposed method directly evaluates the two correlated demixing matrices from the image matrix without matrix-to-vector transformation, greatly alleviates the SSS and the curse of dimensionality, reduces the computational complexity, and simultaneously exploits the spatial and structural information embedded in image. Extensive experiments show that the proposed method is superior to the standard ICA method and some unsupervised methods.

1. Introduction

Dimensionality reduction (DR) is one of the fundamental problems in computer vision, machine learning and Biometric recognition. The goal of DR is to capture the meaningful low-dimensional structures embedded in high-dimensional data and obtain more useful representations of the data for subsequent analysis such as classification, visualization, clustering, or outlier detection (Lawrence and Sam, 2003; Koren and Carmel, 2004). Up to now, it assembles numerous methods, all striving to present high-dimensional data in a low -dimensional space, in a way that faithfully captures the meaningful structures and unexpected relationships embedded in images. In the absence of prior knowledge, such representations must be learned or discovered automatically. Automatic methods which discover hidden structure from the statistical regularities of large data sets can be studied in the general framework of unsupervised and supervised learning (Lawrence and Sam, 2003; Koren and Carmel, 2004).

Principal component analysis (PCA) and linear discriminant analysis (LDA) are the two classical dimensionality reduction techniques in unsupervised and supervised learning, respectively (Yan et al., 2007). In using them for face representation and recognition, Turk and Pentland (1991) and Belhumeur et al. (1997) developed the well-known Eigenfaces and Fisherfaces algorithms, respectively. Based on these contents, PCA and LDA have been successfully used and widely investigated in pattern recognition, machine learning, and computer vision (Koren and Carmel, 2004; Yan et al., 2007; Moghaddam, 2002; Yang et al., 2004; Ye, 2005). The two techniques, however, exploit the second-order statistics which capture the amplitude spectrum of the images but they ignore the higher-order statistical dependencies which may contain useful structural information of the 2D image for subsequent analysis (Yang et al., 2004; Ye, 2005; Bartlett et al., 2002).

Independent component analysis (ICA), as an extension of the PCA, extracts a set of statistically independent components via analyzing the higher-order statistics in the training dataset and has been widely used in blind source separation, signal processing, medical image analysis, pattern recognition, etc. (Hyvařinen, 1999). In using ICA for face representation and recognition, Bartlett et al. (2002) proposed two ICA architectures. Architecture I seeks to find a set of spatially independent basis vectors (images) and the coefficients are not necessarily independent, while architecture II seeks to find a set of basis vectors which make the projection coefficients as statistically independent as possible. Experimental results on FERET database illustrate that ICA is superior to PCA in face recognition and representation. Since then, ICA has been widely investigated in face representation, recognition, and image retrieval, and many algorithms have been developed to improve the classification accuracy up to now (Liu, 2004; Pong and Lai, 2002; Bressan and Vitrià, 2003; Vicente et al., 2007). What's more, Kim et al. (2005) and Li et al. (2005) found that ICA outperforms PCA and some unsupervised learning algorithms in estimating

* Corresponding author. Address: State Key Laboratory on Integrated Serves Networks, XIDIAN University, 710071 XI'AN, China. Fax: +86 29 88204752.
E-mail address: gaoqxnput@163.com (Q. Gao).

0167-8655/$ - see front matter © 2009 Elsevier B.V. All rights reserved.
doi:10.1016/j.patrec.2009.10.014

the pose and multi-view subspace and the case of partial occlusions and local distortions, respectively.

In above algorithms, linear algebra is used to extract the independent components (ICs). Thus they are hard to distinguish the statistic features arisen from different factors inherent to image formation, such as viewpoint, illumination, facial expression, etc. (Vasilescu and Terzopoulos, 2005). To address this problem, Vasilescu and Terzopoulos (2005) used multi-linear algebra to represent theses factors of datasets and extract ICs which contain the relationships between factors, and had achieved better recognition accuracy. But it is not convenient for image recognition because of the unknown different factors in training. What's more, each input image is still unfolded to vector as the same as the above mentioned algorithms. The unfolded vectors, however, lead to the small sample size (SSS) problem and the losing of useful local structure embedded in image.

Inspired by the successful application of face representation with matrix or higher-order tensor in PCA and FLD (Fishers Linear Discriminant) (Yang et al., 2004; Ye, 2005; Liu et al., 1993; Yan et al., 2005) and the work of Zhang et al. (2006), we propose a novel ICs extraction method by encoding each image as a matrix. The proposed method uses two interrelated demixing matrixes (low-dimensional spaces), which correspond to the row and column direction of image, respectively, rather than one demixing matrix in conventional ICA for ICs extraction. Different from the classical two-dimensional subspace analysis that directly learns the low-dimensional subspace from the row/column vectors of images (Gao, 2007; Gao et al., 2007), the proposed method introduces the row and column directional images of the image matrix to efficiently learn the demixing matrixes which efficiently preserve the local structure and spatial information embedded in images. Experimental results on two well-known dataset (Yale and AR) and one palmprint database verify that the proposed method is superior to the conventional ICA algorithm, PCA and 2DPCA, even supervised algorithms.

The rest of this paper is organized as follows: Section 2 reviews the ICA. A novel ICs extraction algorithm is presented in Section 3. Section 4 presents experiment results and analysis. Finally, we provide some concluding remarks and suggestions for future work in Section 5.

2. Independent component analysis (ICA)

As above mentioned, there are two architectures for pattern representation and recognition via ICA. Here, we use the architecture II. Denote by $\boldsymbol{x}$ an arbitrary p-dimensional image vector, the aim of ICA is to seek a sequence of projection vectors $\boldsymbol{w}_1, \boldsymbol{w}_2, \ldots, \boldsymbol{w}_q$ ($q \leqslant p$) to maximize the statistical independence of the projected data $\boldsymbol{s}$. It can be expressed as follows:

$$\boldsymbol{s} = \boldsymbol{W}^T \boldsymbol{x}, \tag{1}$$

where $\boldsymbol{s} = [\boldsymbol{s}_1, \boldsymbol{s}_2, \ldots, \boldsymbol{s}_q]$ denotes the ICs of $\boldsymbol{x}$ and $\boldsymbol{W} = [\boldsymbol{w}_1, \ldots, \boldsymbol{w}_q]$ is called the demixing matrix, i.e. projection matrix.

To efficiently evaluate the demixing matrix $\boldsymbol{W}$, many algorithms based on three criterions, maximum non-Gaussian, minimization of mutual information and maximum likelihood, have been proposed (Hyvařinen, 1999). Among them, the FastICA algorithm (Hyvařinen, 1999), which is based on maximum non-Gaussian, has been dominantly used. In order to reduce complex computation and the number of ICs, PCA is usually implemented to whiten the data and reduce the dimensionality before applying ICA (Bartlett et al., 2002; Liu, 2004; Pong and Lai, 2002; Bressan and Vitrià, 2003).

In order to improve the classification accuracy and further reduce the dimensionality of features, it usually selects a sub-matrix of $\boldsymbol{W}$ via the discriminability of each column of $\boldsymbol{W}$, which can be defined as (Bartlett et al., 2002)

$$r = \frac{\sigma_b}{\sigma_w}, \tag{2}$$

where $\sigma_b = \sum_j (\bar{\boldsymbol{c}}_j - \bar{\boldsymbol{c}})^T (\bar{\boldsymbol{c}}_j - \bar{\boldsymbol{c}})$ and $\sigma_w = \sum_j \sum_{\boldsymbol{s}_k \in c_j} (\boldsymbol{s}_k - \bar{\boldsymbol{c}}_j)^T (\boldsymbol{s}_k - \bar{\boldsymbol{c}}_j)$ represent the between-class and the within-class variability, respectively, of the projected coefficients of the training images. $\bar{\boldsymbol{c}}$ denotes the global mean of the projected coefficients, $\bar{\boldsymbol{c}}_j$ denotes the mean of the projected coefficients for the jth class. $\boldsymbol{s}_k$ denotes the kth ICs in $\boldsymbol{s}$. Based on the magnitude of r, an optimal sub-demixing matrix can be easily obtained by selecting the columns of $\boldsymbol{W}$ corresponding to the several biggest magnitudes. It is an open difficult problem to select the specific number.

In the conventional ICA algorithms, each input image is unfolded to vector $\boldsymbol{x}$ in learning the demixing matrix $\boldsymbol{W}$. Natural images, however, are usually represented in the form of matrices (second-order tensor) or higher-order tensors. Therefore it is not well suited to represent natural images using one-dimensional vectors. In addition, the unfolded vectors lead to the SSS problem and the curse of dimensionality. To address these problems, we will propose a novel ICs extraction scheme by using bilinear algebra in Section 3.

3. Methodology

3.1. Idea and model of extraction ICs

As mentioned in Sections 1 and 2, most existing algorithms stretch the input image to a vector for estimating the demixing matrix. The unfolded vectors lead to the SSS which makes the statistics estimation difficult and inaccurate. Inspired by the scheme of 2DPCA that dilutes the SSS and the successful application of it, we study how to perform ICA with the matrix representation for each input image.

Recently, some researches have revealed that 2DPCA is equivalent to perform PCA on all rows/columns of image matrices under the assumption that all sample are centered (Gao, 2007; Gao et al., 2007), and this scheme ignores the spatial relationship between different rows (columns) of image which is be related with image formation and useful for classification (Zhang et al., 2006, 2008). Inspired by the works of Gao (2007), Gao et al. (2007) and Zhang et al. (2006), we propose the following model to extract ICs with matrix representation for each input image

$$\boldsymbol{S} = \boldsymbol{U}^T \times \boldsymbol{A} \times \boldsymbol{V}, \tag{3}$$

where $\boldsymbol{A} \in \mathbb{R}^{m \times n}$ denotes an arbitrary image matrix, matrix $\boldsymbol{U} \in \mathbb{R}^{m \times d_r} (d_r \leqslant m)$, which is called the demixing or projection matrix, contains the orthogonal vectors spanning the column space of $\boldsymbol{A}$, and $\boldsymbol{V} \in \mathbb{R}^{n \times d_c} (d_c \leqslant n)$, called the demixing matrix, contains the orthogonal vectors spanning the row space of $\boldsymbol{A}$. $\boldsymbol{S}$, called the ICs of $\boldsymbol{A}$, governs the interaction between the matrix $\boldsymbol{U}$ and $\boldsymbol{V}$, and is used for subsequent analysis, such as classification.

Our goal is to find two demixing matrices $\boldsymbol{U}$ and $\boldsymbol{V}$ such that the elements of $\boldsymbol{S}$ are as independent as possible and simultaneously reflect the local structure information embedded in input images. Unfortunately, most existing algorithms use the iteration step or sequential row–column scheme to solve the matrix $\boldsymbol{U}$ and $\boldsymbol{V}$ (Vasilescu and Terzopoulos, 2005; Ye, 2005). The former needs larger computation, while the latter does not ensure to achieve the optimal projection matrices and features extracted by it may lose useful spatial and local structure information embedded in the image. Here, we introduce the row and column directional images of training images, and then estimate the demixing matrices via all the row/column vectors of directional images to solve this problem. This algorithm will be described in the following sub-sections.

Note that, the tensor form of Eq. (3) was published in CVPR'2008 (Zhang et al., 2008). For a high-order tensor, the dimensionality of column vector of a directional image is usually smaller than that of row vector of it. However, for an arbitrary second-order tensor $A \in R^{m\times n}$ (without loss of generality, suppose that $m \geqslant n$), it is not always true. The dimensionality of column vector is usually larger than that of row vector in column directional image of A. The proposed method is developed to address this problem via different directional image under this case.

3.2. Directional images and analysis

In the conventional two-dimensional technique, the projection matrices can be directly calculated from the row and column vectors of training images, respectively (Gao, 2007; Gao et al., 2007). So $\boldsymbol{U}$ and $\boldsymbol{V}$ only reflect the structural information of column vectors and row vectors of images, respectively. In other words, the spatial relationship embedded in different rows/columns in the original image, which is determined by the image formation and may be useful for classification, is lost. To address this problem and improve the classification accuracy, we propose to use row and column directional images to estimate the demixing matrices $\boldsymbol{U}$ and $\boldsymbol{V}$ in Eq. (3).

Given an arbitrary image matrix $\boldsymbol{A} \in \mathbb{R}^{m\times n}$ (without loss of generality, suppose $m \geqslant n$), the column and row directional images can be defined as follows:

Column directional image: Re-sample and re-arrange the matrix $\boldsymbol{A}$ along the row direction to generate the column directional image $\boldsymbol{B}_{(c)}$, as shown in Fig. 1a. Where $h = n/2$.

Row directional image: Re-sample and re-arrange the matrix $\boldsymbol{A}$ along the column direction to generate the row directional image $\boldsymbol{B}_{(r)}$, as shown in Fig. 1b. Where $h = n/2$.

Fig. 2 shows the row and column directional images of the original image, respectively. From the Figs. 2 and 1, it is easy to find that arbitrary column/row of column/row directional images contains pixels which come from the multi-columns/rows rather than one column/row of original image. Thus, the spatial relationship embedded in different columns/rows of the original image can be preserved in column/row directional image. If column/row vectors of column/row directional image are used to learn the demixing

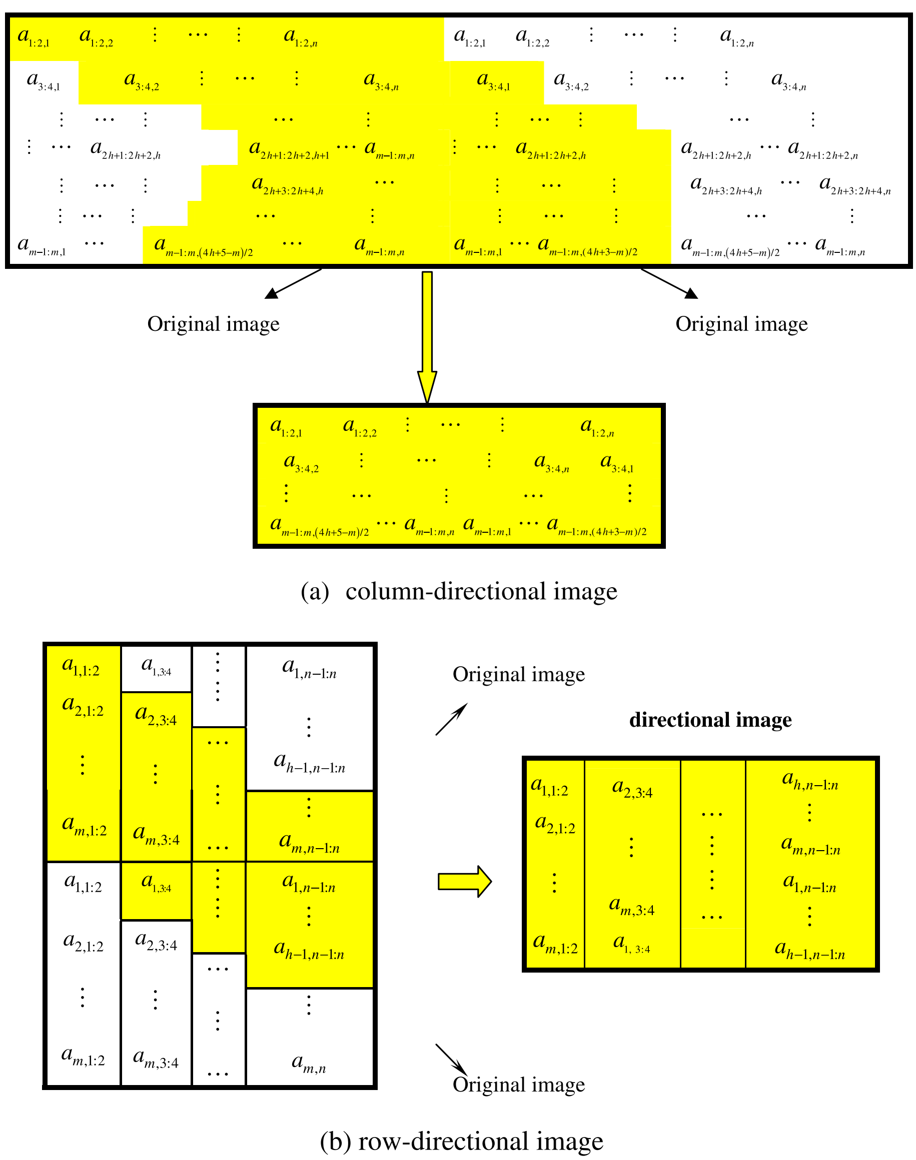

Fig. 1. Row and column directional images formation.

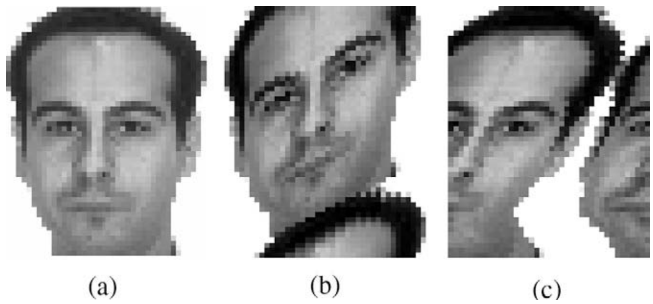

Fig. 2. Original image and directional images. (a) original image; (b) column directional image and (c) row directional image.

matrices $\boldsymbol{U}$ and $\boldsymbol{V}$, respectively, then the spatial structure embedded in original images can be preserved in $\boldsymbol{U}$ and $\boldsymbol{V}$. By projecting the original image onto the two demixing matrices, the corresponding ICs can contain the spatial structure embedded in original images because the column/row vectors of column/row directional images are different from the column/row vectors of original images. Next in Section 3.3 we will introduce how to estimate the two demixing matrices $\boldsymbol{U}$ and $\boldsymbol{V}$, and then in Section 3.4 we will present the ICs extraction and classification of the proposed algorithm.

3.3. Evaluate the demixing matrices

Different from the conventional two-dimensional subspace algorithms of solving low-dimensional spaces, the proposed algorithm views column/row vectors of column/row directional images as training samples to learn the demixing matrices $\boldsymbol{U}$ and $\boldsymbol{V}$ via the FastICA algorithm without the iteration step, and similarly preserves the local structure embedded in different columns/rows of the original image. Algorithm of calculating the demixing matrices can be summarized as follows:

Step 1: Get the training samples $\boldsymbol{A}_j \in \mathbb{R}^{m\times n}, j = 1,2,\ldots,N$, where N denotes the total number of training samples.
Step 2: Obtain the column directional images $\boldsymbol{B}^j_{(c)}, j = 1,2,\ldots,N$. $\boldsymbol{B}^j_{(c)}$ denotes the column directional image of original image $\boldsymbol{A}_j$ obtained via the column directional image definition.
Step 3: Calculate the demixing matrix $\boldsymbol{U}_c$. $\boldsymbol{U}_c$ can be learned from the column vectors of $\boldsymbol{B}^j_{(c)} (j = 1,2,\ldots,N)$ by using the FastICA algorithm.
Step 4: Extract the demixing matrix $\boldsymbol{U} \in \mathbb{R}^{m\times d_r} (d \leqslant m)$ of Eq. (3). $\boldsymbol{U}$ is a sub-matrix of $\boldsymbol{U}_c$ and can be selected via Eq. (2).
Step 5: Obtain the row directional images $\boldsymbol{B}^j_{(r)}, j = 1,2,\ldots,N$. $\boldsymbol{B}^j_{(r)}$ denotes the row directional image of original image $\boldsymbol{A}_j$ obtained via the row directional image definition in Section 3.2.
Step 6: Calculate the demixing matrix $\boldsymbol{V}_r$. $\boldsymbol{V}_r$ can be learned from the row vectors of $\boldsymbol{B}^j_{(r)} (j = 1,2,\ldots,N)$ by using the FastICA algorithm.
Step 7: Obtain the demixing matrix $\boldsymbol{V} \in \mathbb{R}^{n\times d_r}$ of Eq. (3). $\boldsymbol{V}$ is a sub-matrix of $\boldsymbol{V}_r$ and can be selected via Eq. (2).

3.4. Feature selection and classification

After obtaining the demixing matrices $\boldsymbol{U}$ and $\boldsymbol{V}$, the ICs, $\boldsymbol{S}_j$, of image matrix $\boldsymbol{A}_j$ can be extracted by simultaneously projecting $\boldsymbol{A}_j$ onto the spaces spanned by $\boldsymbol{U}$ and $\boldsymbol{V}$

$$\boldsymbol{S}_j = \boldsymbol{U}^T \times \boldsymbol{A}_j \times \boldsymbol{V}, \quad j = 1,2,\ldots,N. \tag{4}$$

For an arbitrary probe image, A^*, the ICs of probe image A^* is

$$\boldsymbol{S}^* = \boldsymbol{U}^T \times A^* \times V. \tag{5}$$

For classification, it needs to calculate the similarity between $\boldsymbol{S}_j$ and $\boldsymbol{S}^*$. Here, we use Euclidean distance to measure their similarity. The distance between $\boldsymbol{S}_j$ and $\boldsymbol{S}^*$ can be defined as:

$$d(\boldsymbol{S}^*, \boldsymbol{S}_j) = \sum_i \|\boldsymbol{S}^*(:,i) - \boldsymbol{S}_j(:,i)\|_2, \tag{6}$$

where $\boldsymbol{S}^*(:,i)$ and $\boldsymbol{S}_j(:,i)$ denote the ith column of $\boldsymbol{S}^*$ and $\boldsymbol{S}_j$, respectively.

If $d(\boldsymbol{S}^*,\boldsymbol{S}_k) = \arg\min_j(d(\boldsymbol{S}^*,\boldsymbol{S}_j))$, then the probe image and the $\boldsymbol{A}_k$ belong to the same class.

4. Experimental results

The performance of the proposed algorithm is evaluated on the Yale and AR face database, and PolyU Palmprint database and compared with the classical unsupervised method, including PCA (Eigenfaces) (Turk and Pentland, 1991), ICA (Bartlett et al., 2002), 2DPCA (Yang et al., 2004), and Tensor-PCA (Ye, 2005), as well as the classical supervised methods, such as FLD (Fisherfaces) (Belhumeur et al., 1997), 2DFLD (Liu et al., 1993) and Tensor-FLD (Yan et al., 2005). In all experiments, the nearest neighbor classifier with Euclidean distance is employed for classification.

4.1. Face recognition

4.1.1. Experiment using the Yale database

The Yale face database (http://cvc.yale.edu/projects/yalefaces/yalefaces.html) was taken from the Yale Center for Computational Vision and Control. It consists of images from 15 different people, using 11 images from each person, and has 165 images in total. The images contain variations with the following total expressions or configurations: center-light, with glasses, happy, left-light,

Fig. 3. Some preprocessed images of Yale database.

Table 1
Top recognition accuracies (%) and the associated dimensionalities on the Yale database by different schemes.

Methods	PCA	ICA	FLD	2DPCA	Tensor-PCA	2DFLD	Tensor-FLD	Proposed method
Accuracy	88.00	90.67	96.00	88.00	89.33	92.00	90.67	97.33
Dimension	8	20	13	504	32	336	198	36

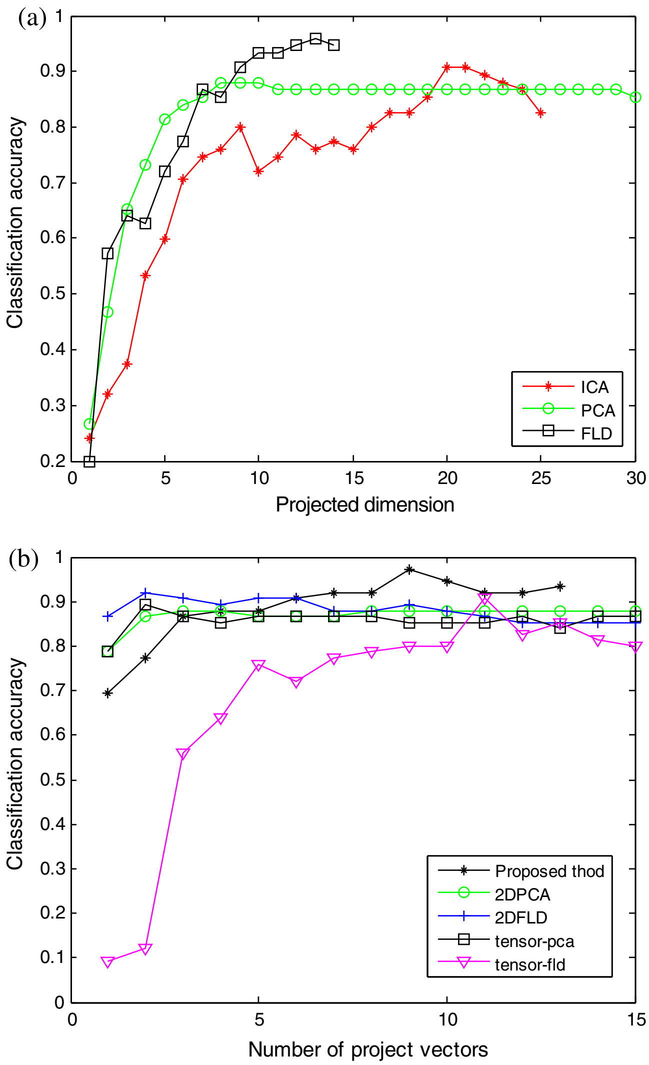

Fig. 4. The recognition accuracy of PCA, ICA, FLD, 2DPCA, 2DFLD, Tensor-PCA, and Tensor-FLD on the Yale database. (a) ICA, PCA, and FLD; (b) 2DPCA, 2DFLD, tensor-pca, tensor-fld, and the proposed method.

without glasses, normal, right-light, sad, sleepy, surprised, and wink. In experiments, the facial portion of each image is manually cropped and then normalized to the size of 168×120. Fig. 3 shows some face images of one subject.

In experiments, we select the first 6 samples of each person for training images, and the remaining images for the probe set. Thus, the total number of training samples and test images are 90 and 75, respectively. PCA, ICA, FLD, 2DPCA, 2DFLD, Tensor-PCA, Tensor-FLD and the proposed method are used to extract features, respectively. Table 1 shows the top recognition accuracy of different schemes with the corresponding dimensionality of features. The curve of recognition accuracy versus the different dimensions or projection vectors of features is illustrated in Fig. 4. Note that LDA, ICA, and the proposed method all involve a PCA phase. In this phase, we keep nearly 85%, 90%, 95% and 98% image energy, respectively, and select the number of principal components which corresponds the best recognition accuracy for each method. In the following experiments, we always select the principal components in PCA phase via this way for these methods.

It can be seen that the proposed method is obviously superior to all the other unsupervised algorithms (PCA, 2DPCA, Tensor-PCA and ICA) and even slightly better than the supervised methods (FLD, 2DFLD and tensor-FLD) in the recognition accuracy. The main reason may be that the proposed method uses higher-order statistics between variable and also exploits the structure information embedded in different rows/columns of original images by introducing column and row directional images. However, the proposed method may need more features than the conventional ICA.

4.1.2. Experiment using the AR face database

AR database (Martinez and Benavente, 2003) contains over 4000 color face images from 126 people (70 men and 56 women), including frontal views of faces with different facial expression, lighting conditions and occlusions. The pictures of most persons were taken in two sessions, separated by two weeks. Each session contains 13 color images per person and 120 individuals (65 men and 55 women) participated in both sessions. In our experiments, the facial portion of each image is manually cropped and then normalized to a size of 50×40. The images from the first session with (a) neutral expression, (b) smile, (c) anger, (d) scream, (e) left-light on, (f) right-light on, and (g) both side light on were selected for gallery. Thus we have 840 images from 120 individuals. Fig. 5 shows some sample images of one subject.

In the experiment, the four sample images per person with (a) neutral expression, (b) smile, (c) anger and (d) scream in the first session are selected for training, and the other three images for testing. The second experiment exchanges the training and testing images. Table 2 lists the top classification accuracies of different algorithms and the corresponding number of features. Fig. 6 plots the curve of the recognition accuracy versus the different dimensions or projection vectors of features in the second experiment. From them, we can easily find that the proposed method has the best recognition accuracy with the smaller number of features. It is consistent with the experimental results in Yale database.

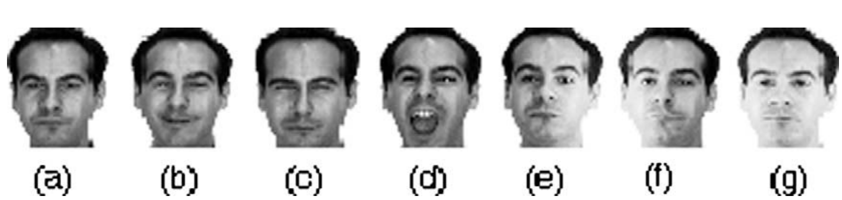

Fig. 5. Some sample images of one subject in the AR database.

Table 2
The recognition accuracies (%) of different schemes on the AR database. The values in parentheses are the corresponding number of features.

Methods	PCA	ICA	2DPCA	Tensor-PCA	Tensor-FLD	FLD	2DFLD	Proposed method
First experiment	68.89 (418)	97.22 (99)	98.61 (1250)	97.78 (500)	98.61 (270)	92.78 (113)	98.06 (1150)	100.00 (126)
Second experiment	79.79 (110)	95.00 (96)	92.92 (500)	92.71 (143)	95.21 (60)	95.83 (79)	95.63 (250)	96.46 (90)

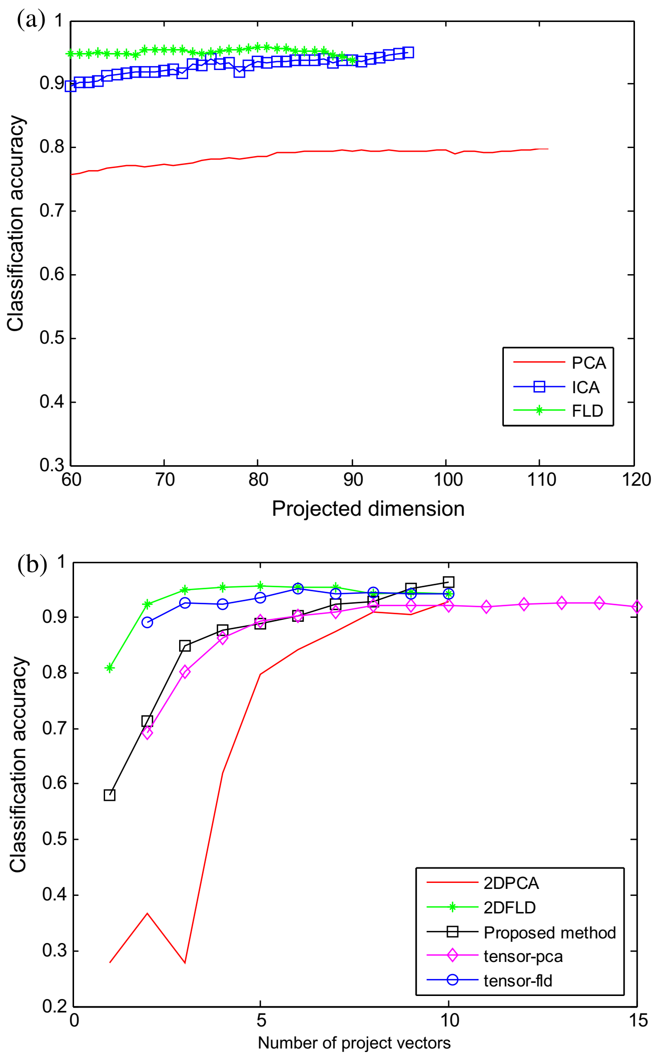

Fig. 6. The recognition accuracy of different vector-based and matrix-based representation algorithms on the AR database. (a) vector-based algorithms, i.e. ICA, PCA, FLD; (b) matrix-based algorithms, i.e. 2DPCA, 2DFLD, Tensor-FLD, Tensor-PCA and the proposed method.

4.1.3. Experiment using the palmprint database

The PolyU palmprint database (http://www.comp.polyu.edu.hk/~biometrics/) was collected from 50 people at different times. The palmprints from right-hand and left-hand of the same person are treated as palmprints from different people. Thus we view the palmprint images from 100 different palms. The resolution of the original palmprint images is 384×284. After preprocessing by using the algorithm mentioned in (Zhang, 2004), the central part of the image is cropped for feature extraction and matching and normalized to a size of 128×128. Fig. 7 shows some palmprint images after preprocessing.

In the experiment, each one of the 100 different palms has 6 samples taken in two sessions, where the first three are captured in the first session and the other three in the second session. The average interval between the first and the second session is two months. The samples from the first session are used for training, and the samples from the second session for testing. Thus, the total number of training samples and test images are both 300. Table 3 shows the top recognition accuracy of different schemes with the associated number of features. The curve of the recognition accuracy versus the different dimensions or projection vectors of features is illustrated in Fig. 8. It can be seen that the proposed method is obviously superior to all the other unsupervised algorithms (PCA, 2DPCA, Tensor-PCA and ICA) and even slightly better than the supervised methods (FLD, 2DFLD and Tensor-FLD) in the recognition accuracy. But, the proposed method may need more features than the conventional ICA. It is consistent with the conclusions in face database.

4.2. Computational complexity and dimensionality dilemma

Given an arbitrary training image $A \in R^{m \times n}$, in the conventional ICA, the size of the covariance matrix in the whitening stage is $L \times L$, where $L = m \times n$. Usually the training sample size N is much smaller than L, i.e., $N \ll L$, in most practical applications. It is hard to calculate accurately and robustly the statistics of the vector variable because the training sample size is much smaller than the dimensionality of the vector variable. In the proposed method, however, the size of the step-wise covariance matrix is only $n \times n$ or $m \times m$, which is much smaller than L in the conventional ICA. So the computational complexity is obviously reduced in our proposed method. On the other hand, the training samples are the column/row directional vectors of directional images and the number of them is $N \times m(N \times n)$, which is much greater than N. Therefore, the dimensionality dilemma is significantly alleviated.

4.3. Discussion

Comparing experiments in Section 4.1, it is easy to find that the proposed method efficiently improves the recognition accuracy, and is superior to some state-of-the-art methods. However, several questions remain to be investigated in our future work. First, the

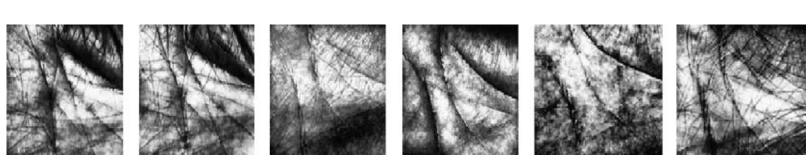

Fig. 7. Some preprocessed images in the palmprint database.

Table 3
Top recognition accuracies (%) and the associated dimensionalities on the palmprint database by different schemes.

Methods	PCA	ICA	FLD	2DPCA	Tensor-PCA	2DFLD	Tensor-FLD	Proposed method
Accuracy	88.00	92.00	93.00	94.00	94.33	95.00	99.00	99.33
Dimension	105	39	84	2432	253	2176	224	112

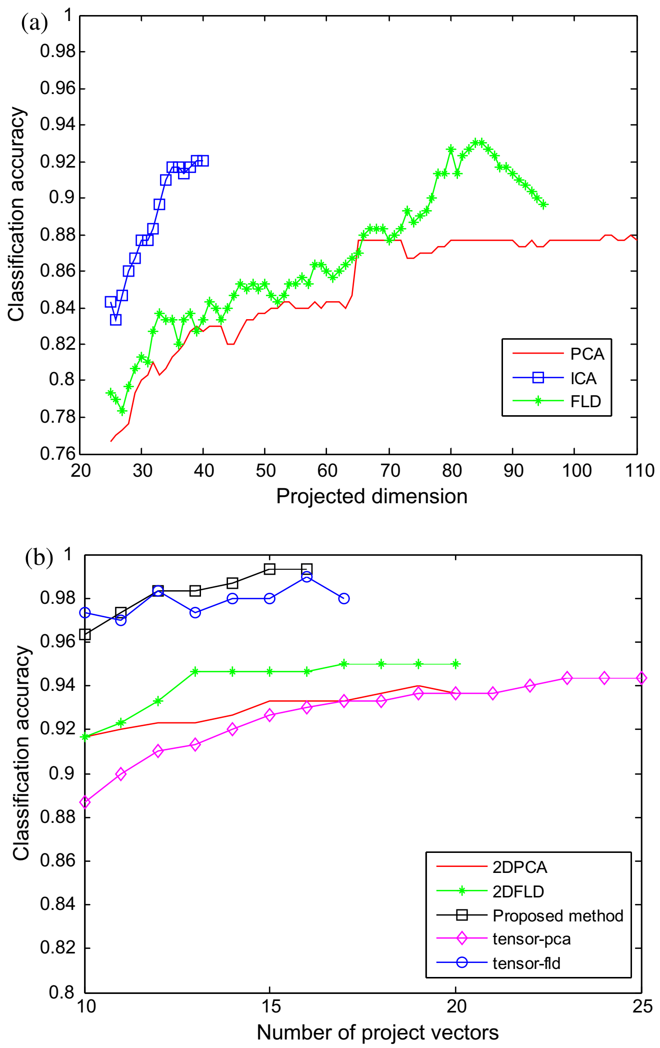

Fig. 8. The recognition accuracy of different vector-based and matrix-based representation algorithms on the palmprint database. (a) Vector-based algorithms, i.e. ICA, PCA, FLD; (b) matrix-based algorithms, i.e. 2DPCA, 2DFLD, Tensor-PCA, Tensor-FLD, and the proposed method.

optimal demixing matrices $\boldsymbol{U}$ and $\boldsymbol{V}$ in Eq. (3) are usually estimated using the iteration step, however, they are calculated via using the directional images separately without iteration step in our proposed method. It is unclear whether the proposed method gives an approximated optimal solution to Eq. (3) in theory. Second, as aforementioned analysis in Section 3.2, directional encoding plays an important role for improving the performance of proposed method. However, it is unclear whether the proposed directional encoding is optimal under arbitrary conditions in theory.

5. Conclusions

A novel algorithm is proposed to extract the independent components by encoding each input image as a matrix. A novel column/row directional image is used in training to learn the two demixing matrixes which better exploit the relationship between the different columns/rows of original image. Compared with traditional ICs extraction algorithms, the proposed method has the following advantages: (1) it helps to alleviate the small sample size problem and avoid the curse of dimensionality; (2) it reduces the computational complexity because of the reduced dimensionality of training samples; (3) it uses two interrelated demixing matrices which can better preserve the relationship embedded in rows/columns of original image. Experimental results show the efficiency of the proposed method. However, some questions (see Section 4.3) remain to be investigated in our future work.

Acknowledgements

The authors thank all the guest editors and anonymous reviewers for their constructive advices. We also thank Dr. Yang Jian for providing the cropped images of AR and Dr. Zhao Qijun for providing the cropped images of Yale database. This work is supported by the National Science Foundation of China under Grants No. 60802075, the Key Laboratory on Integrated Serves Networks of China under Grants No. ISN090403, and the 111 Project of China (B08038).

References

Bartlett, M.S., Movellan, J.R., Sejnowski, T.J., 2002. Face recognition by independent component analysis. IEEE Trans. Neural Networks 13 (6), 1450–1464.
Belhumeur, P.N., Hespanha, J.P., Kriegman, D.J., 1997. Eigenfaces vs. fisherfaces: Recognition using class specific linear projection. IEEE Trans. Pattern Anal. Machine Intell. 19 (7), 711–720.
Bressan, M., Vitrià, J., 2003. On the selection and classification of independent features. IEEE Trans. Pattern Anal. Machine Intell. 25 (10), 1312–1317.
Gao, Q., 2007. Is two-dimensional PCA equivalent to a special case of modular PCA? Pattern Recognition Lett. 28 (10), 1250–1251.
Gao, Q., Zhang, L., Zhang, D., Yang, J., 2007. Comment on 'on image matrix based feature extraction algorithms'. IEEE Trans. Systems Man Cybernet., Part B 37 (5), 1373–1374.
Hyvarinen, A., 1999. Survey on independent component analysis. Neural Comput. Surveys 2, 94–128.
Kim, J., Choi, J.M., Yi, J., Turk, M., 2005. Effective representation using ICA for face recognition robust to local distortion and partial occlusion. IEEE Trans. Pattern Anal. Machine Intell. 27 (12), 1977–1981.
Koren, Y., Carmel, L., 2004. Robust linear dimensionality reduction. IEEE Trans. Vis. Comput. Graphics 10 (4), 459–470.
Lawrence, K.S., Sam, T.R., 2003. Think globally, fit locally: Unsupervised learning of low dimensional manifolds. J. Machine Learn. Res. 4, 119–155.
Li, S.Z., Lu, X., Hou, X., Peng, X., Chen, Q., 2005. Learning multiview face subspaces and facial pose estimation using independent component analysis. IEEE Trans. Image Process. 14 (6), 705–712.
Liu, C., 2004. Enhanced independent component analysis and its application to content based face image retrieval. IEEE Trans. Systems Man Cybernet., Part B 34 (2), 1117–1127.
Liu, K., Cheng, Y., Yang, J., 1993. Algebraic feature extraction for image recognition based on an optimal discriminant criterion. Pattern Recognition 26 (6), 903–911.
Martinez, A.M., Benavente, R., 2003. The AR Face Database. http://rvl1.ecn.purdue.edu/~aleix/aleix_face_DB.html.
Moghaddam, B., 2002. Principal manifolds and probabilistic subspaces for visual recognition. IEEE Trans. Pattern Anal. Machine Intell. 24 (6), 780–788.
Pong, C.Y., Lai, J.H., 2002. Face representation using independent component analysis. Pattern Recognition 35, 1247–1257.
Turk, M., Pentland, A., 1991. Eigenfaces for recognition. J. Cognitive Neurosci. 3 (1), 72–86.
Vasilescu, M.A.O., Terzopoulos, D., 2005. Multilinear independent component analysis. In: Proc. IEEE Conf. on Computer Vision and Pattern Recognition.
Vicente, M.A., Hoyer, P.O., Hyvärinen, A., 2007. Equivalence of some common linear feature extraction techniques for appearance-based object recognition tasks. IEEE Trans. Pattern Anal. Machine Intell. 29 (5), 896–900.

Yan, S., Xu, D., Yang, Q., Zhang, L., Tang, X., Zhang, H., 2005. Discriminant analysis with tensor representation. In: Proc. IEEE Conf. on Computer Vision and Pattern Recognition (CVPR).

Yan, S., Xu, D., Zhang, B., Zhang, H., Yang, Q., Lin, S., 2007. Graph embedding and extensions: A general framework for dimensionality reduction. IEEE Trans. Pattern Anal. Machine Intell. 29 (1), 40–51.

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

Ye, J., 2005. Generalized low rank approximations of matrices. Machine Learn. 61, 167–191.

Zhang, D., 2004. Palmprint Authentication. Kluwer Academic.

Zhang, D., Zhou, Z., Chen, S., 2006. Diagonal principal component analysis for face recognition. Pattern Recognition 39, 140–142.

Zhang, L., Gao, Q., Zhang, D., 2008. Directional independent component analysis with tensor representation. In: Proc. Computer Vision and Pattern Recognition.