

High Performance and Efficient Facial Recognition Using Norm of ICA/Multiwavelet Features

Ahmed Aldhahab^{1,2}(✉), George Atia¹, and Wasfy B. Mikhael¹

¹ Department of Electrical and Computer Engineering,
University of Central Florida, Orlando, FL, USA

² Department of Electrical Engineering, University of Babylon, Hilla, Babylon, Iraq
Aldhahab2012@Knights.ucf.edu, {George.Atia, Wasfy.Mikhael}@ucf.edu

Abstract. In this paper, a supervised facial recognition system is proposed. For feature extraction, a Two-Dimensional Discrete Multiwavelet Transform (2D DMWT) is applied to the training databases to compress the data and extract useful information from the face images. Then, a Two-Dimensional Fast Independent Component Analysis (2D FastICA) is applied to different combinations of poses corresponding to the subimages of the low-low frequency subband of the MWT, and the ℓ_2 -norm of the resulting features are computed to obtain discriminating and independent features, while achieving significant dimensionality reduction. The compact features are fed to a Neural Network (NNT) based classifier to identify the unknown images. The proposed techniques are evaluated using three different databases, namely, ORL, YALE, and FERET. The recognition rates are measured using K-fold Cross Validation. The proposed approach is shown to yield significant improvement in storage requirements, computational complexity, as well as recognition rates over existing approaches.

1 Introduction

Facial recognition is an important task of numerous applications in computer vision, pattern recognition, and image processing, which has received renewed attention in recent year due to its wide applicability in security, biomedical diagnosis, control, and personal identification. The efficiency and reliability of different recognition systems depends on several factors, including the computational complexity, storage requirements, and recognition rates. Therefore, extracting efficient and compact features is central to every reliable recognition system.

Numerous algorithms and techniques were proposed and used for facial recognition. Recognition based on Independent Component Analysis (ICA) of wavelet features was proposed in [1]. There, various wavelet functions were used to extract the features by decomposing the face image into 8-levels. Then, the extracted facial features were analysed using 2D ICA and the Euclidean distance is used for accuracy measurements. The algorithm was tested on the ORL

database and the recognition rate recorded was 91.5%. A 2D Principal Component Analysis (2D PCA) approach based on multiresolution analysis for facial recognition was proposed in [2]. Multi-levels of decompositions of 2D Discrete Wavelet Transform (DWT) are used to extract the facial features. The highest recognition rate achieved using the combination of 2-level 2D DWT/2D PCA was 94.5% while using 2D PCA alone achieved 90.5% for the ORL database. The authors in [3] proposed a 2D Discrete Cosine Transform (2D DCT) approach wherein the feature vectors are extracted using 2D DCT and a SVM is used for recognition. The system was tested on the ORL database and a recognition rate of 95% was achieved. PCA and Daubechies wavelet subbands were presented in [4]. First, PCA was applied into the four levels of Daubechies wavelet subbands to select the features. Then, the City Block distance and Euclidean distance measures were used for accuracy measurements. The system was tested on the ORL database and a recognition rate of 96.87% was reported. An approach that integrates 2D Stationary Multiwavelet Transform (MWT) and 2D PCA was proposed in [5]. The features were extracted using 2D SMWT, then 2D PCA and the Histogram-Based Method (HBM) were applied to the extracted features. The highest recognition rates achieved were 63%, 89%, and 94.5% for 2D PCA, 2D SMWT, and HBM, respectively. These algorithms were tested on the ORL database.

In this paper, a new approach based on the 2D DMWT and 2D FastICA is proposed. In contrast to prior work, the proposed approach has two key contributions. First, we apply 2D FastICA to 4 different combinations of 2D DMWT features obtained by considering every subimage of the low-low (LL) frequency subband of the MWT from all poses, thereby exploiting the redundancy in these subimages to further reduce the dimensionality. Second, for the combination corresponding to each subimage we incorporate information from different poses for each person, which leads to a rich set of features that account for different illumination, rotations and facial expressions. As such, the features used lead to a notable dimensionality reduction, reduction in computational complexity, and high recognition rates compared to the existing approaches [1, 2]. The proposed approach is evaluated using three different databases, namely, ORL, YALE, and FERET, which have different light conditions, angle rotations, and facial expressions. The classification rates of the system are measured using K-fold Cross Validation (CV).

The paper is organized as follows. In Sect. 2, some preliminary background is provided. In Sect. 3, the proposed system is presented. The experimental results and analysis of the proposed approach are presented in Sect. 4. Section 5 provides some concluding remarks.

2 Multiwavelet Transform and ICA

2.1 Discrete MultiWavelet Transform

The Wavelet transform is broadly used in signal processing, image processing, and pattern recognition for Multiresolution Analysis. The Multiwavelet Transform (MWT) naturally extends the traditional scalar case, where only one scaling and

wavelet function are used, to the general case of R scaling and wavelet functions. The scaling function $\Phi(t)$ and the wavelet function $\Psi(t)$ are associated with the low pass and high pass filter, respectively [6]. In Multiwavelets with multiplicity R , the multi-scaling function $\Phi(t)$ and wavelet function $\Psi(t)$ can be written in vector notation as

$$\Phi(t) \equiv [\phi_1(t), \phi_2(t), \dots, \phi_R(t)]^T \tag{1}$$

$$\Psi(t) \equiv [\psi_1(t), \psi_2(t), \dots, \psi_R(t)]^T . \tag{2}$$

The case where $R = 1$ corresponds to the traditional scalar wavelet transform. The two-scale equations of Multiwavelets resemble those of scalar wavelet, i.e.,

$$\Phi(t) = \sqrt{2} \sum_{k=-\infty}^{\infty} H_k \cdot \Phi(2t - k) \tag{3}$$

$$\Psi(t) = \sqrt{2} \sum_{k=-\infty}^{\infty} G_k \cdot \Phi(2t - k) . \tag{4}$$

where H_k and G_k are the filter matrices with dimension $R \times R$ for each integer k [7]. In contrast to the scalar wavelet, MWT can simultaneously achieve perfect reconstruction while preserving length (orthogonality), good performance at the boundaries (via linear phase symmetry), and high-order of approximation (vanishing moments). These favorable properties cannot be achieved by the scalar wavelet at the same time, wherefore Multiwavelets offer more degrees of freedom, and potentially superior performance, for image and signal processing applications compared with scalar wavelet [7].

The multiscaling and wavelet functions are often used with multiplicity $R = 2$, in which case they can be written as $\Phi(t) = [\phi_1(t) \ \phi_2(t)]^T$ and $\Psi(t) = [\psi_1(t) \ \psi_2(t)]^T$, respectively [7]. In this paper, we use the well-known GHM filter due to Geronimo, Hardian, and Massopust [8]. The multiscaling and wavelet functions for GHM system have four scaling and wavelet matrices [8]:

$$H_0 = \begin{bmatrix} \frac{3}{5\sqrt{2}} & \frac{2}{5} \\ \frac{-1}{20} & \frac{-3}{10\sqrt{2}} \end{bmatrix}, H_1 = \begin{bmatrix} \frac{3}{2\sqrt{2}} & 0 \\ \frac{9}{20} & \frac{1}{\sqrt{2}} \end{bmatrix}, H_2 = \begin{bmatrix} 0 & 0 \\ \frac{9}{20} & \frac{-3}{10\sqrt{2}} \end{bmatrix}, H_3 = \begin{bmatrix} 0 & 0 \\ \frac{-1}{20} & 0 \end{bmatrix}$$

$$G_0 = \begin{bmatrix} \frac{-1}{20} & \frac{-3}{10\sqrt{2}} \\ \frac{1}{10\sqrt{2}} & \frac{3}{10} \end{bmatrix}, G_1 = \begin{bmatrix} \frac{9}{20} & \frac{-1}{\sqrt{2}} \\ \frac{9}{20} & \frac{1}{\sqrt{2}} \end{bmatrix}, G_2 = \begin{bmatrix} \frac{9}{10\sqrt{2}} & \frac{-3}{10\sqrt{2}} \\ \frac{9}{10\sqrt{2}} & \frac{-3}{10} \end{bmatrix}, G_3 = \begin{bmatrix} \frac{-1}{20} & 0 \\ \frac{1}{10\sqrt{2}} & 0 \end{bmatrix}$$

Unlike scalar wavelet, in MWT the low pass H_k and high pass G_k filters are $R \times R$ matrices, which calls for the use of prefiltering. There are several ways to do prefiltering. In this paper, the critically-sampled scheme is used [9].

2.2 Independent Component Analysis (ICA)

ICA has been primarily used for blind source separation (BSS) [10, 11]. ICA has also been successfully used for feature extraction in Image Processing, where it

can extract independent image bases that are not necessarily orthogonal. While PCA only considers second-order statistics, it is well-known that ICA exploits high order statistics in the data [1]. Moreover, the bases vectors of ICA capture local image characteristics. These are particularly useful properties for facial recognition as important information may be found in the high order relationships between the pixels, and also due to the fact that the local features of human faces are quite robust to facial expressions, illumination and occlusions owing to their non-rigid nature [12].

The idea of ICA is to represent a set of random variables (RVs) using bases functions such that the coefficients of the expansion are statistically independent or nearly independent. Consider observing M random variables X_1, X_2, \dots, X_M with zero mean, which are assumed to be linear combinations of N mutually independent components V_1, V_2, \dots, V_N . Let the vector $X = [X_1, X_2, \dots, X_M]^T$ denote an $M \times 1$ vector of observed variables and $V = [V_1, V_2, \dots, V_N]^T$ an $N \times 1$ vector of mutually independent components. The relation between X and V is expressed as

$$X = AV. \quad (5)$$

where A is a full rank $M \times N$ unknown matrix, called the mixing matrix or the feature matrix [13]. In this paper, X corresponds to the intensities of the pixels of the LL subband of the facial images.

ICA has two main drawbacks. First, it typically requires complex matrix operations [14]. Second, it has slow convergence [10]. To alleviate these two drawbacks, [11] introduced a new technique called FastICA, which is used in this paper. FastICA is computationally more efficient for estimating the ICA components and has a faster convergence rate by using a fixed point iteration algorithm [11]. According to [11, 13, 15], the mixing matrix A (feature matrix) can be considered as another representation of the face images.

3 Proposed Approach

In this section, the proposed system is presented. The supervised facial recognition system consists of three phases, namely, Preprocessing, Feature Extraction, and Recognition.

3.1 Preprocessing: *This Phase Consists of Two Steps:*

1. The first step aims to convert all dimensions of different image databases used in this paper (see Table 1) to an appropriate one. This dimension is chosen since the algorithm used requires dimensions that are power of two.
2. This step aims to convert all the images into $(128 \times 128 \text{ double})$ instead of using *uint8* extension, which is not suitable for the transform used in this paper.

Table 1. The dimensions of the databases

Databases	ORL	YALE	FERET	Proposed
Size	112×92	243×320	384×256	128×128

3.2 Feature Extraction

Feature extraction is an essential component of facial recognition systems. A face image may contain information that is not essential for the recognition purpose. Hence, the following effective tools are applied to get an efficient representation of the images by extracting discriminating and independent features:

1. The 2D DMWT based on MRA is used for:
 - (a) Dimensionality reduction.
 - (b) Noise reduction.
 - (c) Localizing all the useful information in one single band.
2. The FastICA is used for:
 - (a) Decorrrelating the high order statistics since most of the significant information is contained in the high order statistics of the image [1, 2, 11–15].
 - (b) The ICA features are independent leading to a better representation and hence better identification and recognition rates.
 - (c) The ICA features are less sensitive to the facial variations arising from different facial expressions and different poses [1, 2, 11, 13, 15].
 - (d) Reducing the computational complexity and improving the convergence rate.

First, the 2D DMWT is applied to the different databases. Figure 1 shows an example of applying 2D DMWT to different face images. As shown in Fig. 1, the images are divided into four main subbands with 64×64 dimension and each one is further divided into four 32×32 subimages. From Fig. 1a, b, c, it is easily seen that most of the useful information is localized in the upper left band, which is related to the low-low (LL) frequency band of the DMWT. Thus the LL subband is retained, and all the remaining subbands are eliminated. Therefore, the resulting image matrix is 64×64 . The following procedures are applied to the four subimages of the LL subband of DMWT in order to get an efficient representation for the input images.

1. Convert each 32×32 subimage into a 1024×1 vector.
2. Repeat (1) for all subimages of each pose. The resulting features for each pose have dimension 1024×4 .
3. For each of the four subimages of the LL subband, combine the 1024×1 subimage vectors of the different poses of each person as shown in Fig. 2-B.

Thus, the combinations shown in Fig. 2-B are defined as:

$$\text{Combination}_i = [S_{pi}] \quad (6)$$

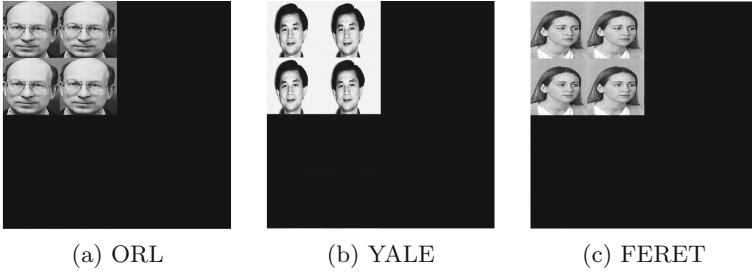


Fig. 1. Resultant images after applying 2-D DMWT to the different databases.

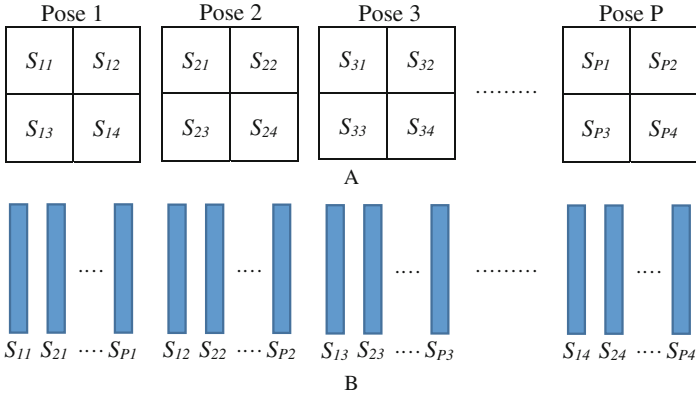


Fig. 2. Figure 2-A represents the subimages of the LL subband of 2D DMWT for P training poses from each person. Figure 2-B is rearranging Fig. 2-A in four combinations.

where $p \in \{1, 2, 3, \dots, P\}$ and P is the number of poses used in the training mode, $i \in \{1, 2, 3, 4\}$ indexes the four subimages of the LL subband.

It is very useful to extract new features from the original features to reduce the dimensionality of the feature space. To this end, we apply the following four techniques to the *Combinations* above:

A. *Technique 1*

- i. Apply 2D FastICA on each combination as shown in Fig. 3.
- ii. Choose the IF_p column that has the maximum ℓ_2 -norm to represent the features of each combination. In this case, the resultant features of each person have dimensions 1024×4 regardless of the number of poses used in both the training and the testing modes.

B. *Technique 2* (shown in Fig. 4)

- i. Apply 2D FastICA on each combination.
- ii. Form a 1024×1 column vector whose entries consist of the ℓ_2 -Norm for each row of the resultant features to reduce the dimensionality. The resultant features of each person have dimension 1024×4 regardless of the number of poses used in training mode.

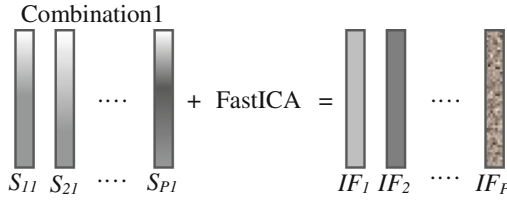


Fig. 3. An example of applying 2D FastICA to the first combination. IF_p represents the resultant Independent features where $p \in \{1, 2, 3, \dots, P\}$.

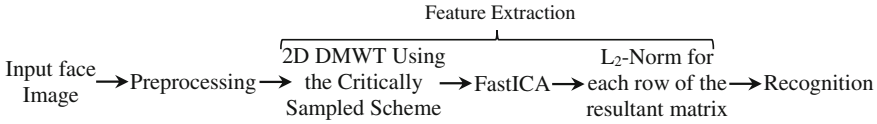


Fig. 4. Proposed Technique 2.

- C. *Technique 3*: This is similar to Tech.2, except that we use the mixing matrix A shown in (5) instead of using the resultant FastICA matrix. This is an alternative approach for representing the face images [11, 15].
 - D. *Technique 4*: Same as Tech.2 without applying step B-i.
4. Repeat [1-3] for all poses of the three different databases.
 5. Repeat Tech.1-Tech.4 for all persons in the three different databases. For all techniques, there are 1024×4 features corresponding to each person regardless of the number of poses used in training mode.

3.3 Recognition

A Neural Network (NNT) based on the Back Propagation Training Algorithm (BPTA) is used during the training and the testing phases. Since BPTA is a supervised learning algorithm, it is necessary to choose a desired output for each database used. In this paper, the system is evaluated using three different databases, each with a different number of persons. Therefore, there are 40, 15, 200 desired outputs for the ORL, YALE, and FERET databases, respectively, corresponding to the number of persons in each database. The configuration of the NNT has three layers, namely, an input layer, a hidden layer, and an output layer.

In the training mode, the classifier is configured using the described training features. For testing, the recognition rate is measured as $\frac{D}{L} \times 100\%$, where D is the total number of images correctly matched and L is the total number of images in the database.

In the testing mode, we follow the same steps of the training phase. First, the 2D DMWT is applied to the test image. Second, only the LL subband is

retained. Third, each subimage is converted to 1D form. Then, Tech.1-Tech.4 are applied to the 1024×4 resultant features corresponding to the four subimages of the *LL* subband. For each technique, there is a 1024×1 feature vector for each test image. This vector is fed into the NNT for classification.

4 Experimental Results and Discussion

The experimental results of the proposed techniques are presented in this section. We also compare the results of the proposed approach to some of the state-of-the-art methods. The techniques are tested using three different databases, ORL, YALE, and FERET, which have different facial expressions, light conditions, and rotations. K-fold Cross Validation [16] is used to evaluate our proposed techniques. Different values of K are used in this paper as show in Tables 2, 3, and 4.

4.1 Experimental Results for the ORL Database

The ORL database consists of 40 persons, each with 10 different poses. Therefore, P different poses are used in the training mode, and $10 - P$ poses are used in the testing mode. Table 2 summarizes the results for all different techniques. As shown, even when a small number of poses are used in the training mode, the algorithm achieves high recognition accuracy. As the number of poses increased, the performance is improved.

4.2 Experimental Results for the YALE Database

This database consists of 15 persons, each with 11 different poses. Table 3 summarizes the results of the proposed techniques. As before, our proposed techniques show superior performance.

Table 2. Experimental Results for the ORL Database

K Fold	Training Rates	Recognition Rates for Testing Phase				2D DWT and	2D DWT and
		Tech.1	Tech.2	Tech.3	Tech.4	2D ICA [1]	2D PCA [2]
K=2	100 %	96.75 %	98.25%	97.25 %	96.25 %	93.75 %	92.5 %
K=3	100 %	97.5 %	98.75%	97.75 %	96.75 %	95.25 %	94.75 %
K=5	100 %	98.25 %	99.25%	98.75 %	97.75 %	96.75 %	96.5 %

Table 3. Experimental results for the YALE database

K Fold	Training Rates	Recognition Rates for Testing Phase				2D DWT and	2D DWT and
		Tech.1	Tech.2	Tech.3	Tech.4	2D ICA [1]	2D PCA [2]
K=2	100 %	96.97 %	98.18%	96.97 %	95.15 %	92.73 %	92.12 %
K=3	100 %	97.58 %	98.79%	97.58 %	96.36 %	94.55 %	93.94 %
K=5	100 %	98.18 %	99.39%	98.78 %	97.58 %	96.36 %	95.15 %

Table 4. Experimental results for the FERET database

K Fold	Training Rates	Recognition Rates for Testing Phase				2D DWT and	2D DWT and
		Tech.1	Tech.2	Tech.3	Tech.4	2D ICA [1]	2D PCA [2]
K=2	100 %	96.68 %	98.32%	97.36 %	95.86 %	92.87 %	92.5 %
K=3	100 %	97.6 %	98.6%	98.1 %	96.9 %	95.14 %	94.86 %
K=5	100 %	97.78 %	99%	98.46 %	96.96 %	96.14 %	95.68 %

4.3 Experimental Results for the FERET Database

There are 200 persons in this database, each with 11 different poses. Table 4 summarizes the results for the proposed techniques. The results exhibit the same behavior as in the previous databases.

The goal of using FastICA in this paper is to find a set of statistically independent basis images that preserve the structural information of the face images present in the facial expressions, illumination, and rotations. ICA representations [15] are more robust to variations due to different light conditions, changes in hair location, make up, and facial expressions. This efficient representation leads to significant improvement in the recognition rates. As demonstrated in the previous results, Tech. 1, Tech. 2, Tech. 3 outperform Tech. 4, which does not use the FastICA step. Our proposed techniques are shown to outperform the existing method [1, 2] in the recognition rates, storage requirements, and computational complexity. The configuration of the Neural Network during the training mode can effect the overall accuracy of the system. Hence, choosing the number of hidden layers, number of neurons in each hidden layer, activation function, training algorithms, and the target performance play an important role in the overall system performance.

5 Conclusion

A facial recognition system based on the integration of Multiresolution Analysis and FastICA was proposed. The proposed approach leads to notable dimensionality reduction and subsequently less storage requirements, as well as higher recognition rates. These benefits are due to using combinations of different poses for the subimages of the LL subband of the 2D DMWT and the ℓ_2 Norm for 2D FastICA features in the feature extraction step. Therefore, we exploit the redundancy present in these subimages to reduce the dimensions of the used features, which also account for all the variations in one short feature vector per combination. Each person is represented using features of dimension 1024×4 regardless of the number of poses in the training phase. Compared with the existing methods, the proposed approach achieves higher recognition rates with less storage and computation requirements. The proposed techniques were evaluated using three different databases with different light conditions, rotations, and facial expressions. The recognition rates are analyzed using K-fold CV. For example, for K=2, the highest recognition rates achieved were 98.25 %, 98.18 %, and 98.32 % for the ORL, YALE, and FERET databases, respectively.

Acknowledgement. This work was supported in part by NSF grant (CCF - 1320547) and by the Iraqi government scholarship (HCED).

References

1. Kinage, K., Bhirud, S.: Face recognition based on independent component analysis on wavelet subband. In: 3rd IEEE International Conference on Computer Science and Information Technology (ICCSIT). vol. 9, pp. 436–440 (2010)
2. AlEnzi, V., Alfiras, M., Alsaqre, F.: Face recognition algorithm using two dimensional principal component analysis based on discrete wavelet transform. In: Snasel, V., Platos, J., El-Qawasmeh, E. (eds.) ICDIPC 2011, Part I. CCIS, vol. 188, pp. 426–438. Springer, Heidelberg (2011)
3. Hongtao, Y., Jiaqing, Q., Ping, F.: Face recognition with discrete cosine transform. In: Second International Conference on Instrumentation, Measurement, Computer, Communication and Control (IMCCC), pp. 802–805 (2012)
4. Satone, M., Kharate, G.: Face recognition based on pca on wavelet subband. In: IEEE Students' Conference on Electrical, Electronics and Computer Science (SCEECS), pp. 1–4 (2012)
5. Ismaeel, T.Z., Kamil, A.A., Naji, A.K.: Article: human face recognition using stationary multiwavelet transform. *Int. J. Comput. Appl. (IJCA)* **72**, 23–32 (2013)
6. Zhihua, X., Guodong, L.: Weighted infrared face recognition in multiwavelet domain. In: IEEE International Conference on Imaging Systems and Techniques (IST), pp. 70–74 (2013)
7. Reddy, P.V.N., Prasad, K.: Article: Multiwavelet based texture features for content based image retrieval. *Int. J. Comput. Appl.* **17**, 39–44 (2011)
8. Geronimo, J.S., Hardin, D.P., Massopust, P.R.: Fractal functions and wavelet expansions based on several scaling functions. *J. Approx. Theor.* **78**, 373–401 (1994)
9. Strela, V., Heller, P., Strang, G., Topiwala, P., Heil, C.: The application of multiwavelet filterbanks to image processing. *IEEE Trans. Image Process.* **8**, 548–563 (1999)
10. Bell, A., Sejnowski, T.: An information-maximization approach to blind separation and blind deconvolution. *Neural Comput.* **7**, 1129–1159 (1995)
11. Hyvarinen, A.: Fast and robust fixed-point algorithms for independent component analysis. *IEEE Trans. Neural Netw.* **10**, 626–634 (1999)
12. Lihong, Z., Ye, W., Hongfeng, T.: Face recognition based on independent component analysis. In: Control and Decision Conference (CCDC), Chinese, pp. 426–429 (2011)
13. Yuen, P.C., Lai, J.H.: Face representation using independent component analysis. *Pattern Recogn.* **35**, 1247–1257 (2002)
14. Comon, P.: Independent component analysis, a new concept? *Sig. Process.* **36**, 287–314 (1994)
15. Bartlett, M., Movellan, J.R., Sejnowski, T.: Face recognition by independent component analysis. *IEEE Trans. Neural Netw.* **13**, 1450–1464 (2002)
16. Arlot, S., Celisse, A., et al.: A survey of cross-validation procedures for model selection. *Stat. Surv.* **4**, 40–79 (2010)