Computers and Electrical Engineering 46 (2015) 217–230

Contents lists available at ScienceDirect

Computers and Electrical Engineering

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

# A novel algorithm using affine-invariant features for pose-variant face recognition ☆

Youen Zhao *, Li Li, Zhaoguang Liu

*School of Computer Science and Technology, Shandong University of Finance and Economics, Jinan, China*
*Provincial Key Laboratory of Digital Media Technology, Shandong University of Finance and Economics, Jinan, China*

ARTICLE INFO

*Article history:*
Received 8 January 2014
Received in revised form 14 December 2014
Accepted 16 December 2014
Available online 3 January 2015

*Keywords:*
Face recognition
Pose variance
Affine-invariant feature
Multi-scale autoconvolution
Principal component analysis
Illumination change

ABSTRACT

Pose variation is the major factor significantly affecting recognition efficiency in the field of face recognition. In this paper, we propose a novel algorithm for pose-variant face recognition. We first use the affine-invariant multi-scale autoconvolution (MSA) transformation to extract pose-invariant features. Following this, we use principal component analysis (PCA) to decorrelate the feature sets and reduce the number of required MSA dimensions. The PCA components with sufficiently large corresponding eigenvalues are then passed to the *k* nearest classifier. Compared with other pose-variant face recognition techniques, our proposed algorithm exhibits a superior ability to pose changes, illumination changes, and different face databases, moreover records a higher recognition rate.

© 2014 Elsevier Ltd. All rights reserved.

## 1. Introduction

Human face recognition continues to be one of the most active areas in the fields of computer vision and pattern recognition because of its wide range of applications in traditional security as well as numerous scenarios involving image tagging and searching. As one of the most important biometric techniques, face recognition has a clear advantage of being natural and passive over other techniques such as fingerprint recognition and iris recognition, which require cooperative subjects. In order to benefit from the unintrusive nature of this technique, a face recognition system must possess the capability to identify an uncooperative face in arbitrary poses, e.g., a forward-looking face, an overlooking face, a side-looking face, and an upward-looking face, etc., as shown in Fig. 1.

Owing to the complex three-dimensional (3D) structure of a human face, pose variations present a serious challenge to current face recognition systems. In particular, the innate facial characteristics distinguishing one face from another do not vary significantly among people. However, the extent of image variation caused by pose variations is often greater than that of dissimilarities due to the innate characteristics. Pose-invariant face recognition algorithms thus need to address the task of extracting innate characteristics from a variety of poses [1].

Considerable algorithmic progress has been made in face recognition involving forward-looking face images [2–4]. This progress can be attributed to greater interest among the researchers in pose-variant face recognition following its identification as an important problem. Consequently, a few promising methods have been proposed to solve the problem

☆ Reviews processed and recommended for publication to the Editor-in-Chief by Associate Editor Dr. Eduardo Cabal-Yepez.

* Corresponding author at: School of Computer Science and Technology, Shandong University of Finance and Economics, Jinan, China.
*E-mail address:* zhaoyouen@sohu.com (Y. Zhao).

http://dx.doi.org/10.1016/j.compeleceng.2014.12.010
0045-7906/© 2014 Elsevier Ltd. All rights reserved.

of pose-variant face recognition, such as Tied Factor Analysis (TFA) [5], 3D Modular Discriminant analysis (3DMDA) [6], and Eigen Light-Field (ELF) [7]. However, none of these methods has succeeded in solving the pose problem in face recognition. Hence, further research is needed to attain pose-variant face recognition.

In general, face recognition techniques involving poses fall into two categories: methods based on 3D models, and methods based on two-dimensional (2D) techniques [8]. The use of 3D models to tackle pose variations has been successful because human heads are 3D objects. Moreover, 3D facial data is considered invariant against changes of pose and illumination. The advantages of using 3D models in face recognition have been highlighted in [9]. However, 3D face recognition algorithms do not go beyond 2D methods in terms of performance. This is compounded by the weakness of these methods with regard to data collection and complexity. Therefore, 3D face recognition technology is not widely used [8]. 2D techniques include pose transformation in image space [7,10,11] and pose-tolerant feature extraction [5,12–14]. The former involves synthesizing virtual views into possible poses from a limited number of real views (often from a single view) by using reference faces as prior knowledge. These techniques can effectively handle pose variations within small-to-medium rotations usually limited to 45°. However, large pose variations can cause image discontinuities in 2D image space to the extent that they cannot be reliably handled. Pose-tolerant feature extraction attempts to transform the image space into a feature space where pose variation can be tolerated to a greater degree. In this respect, various pose-invariant transformation methods have been proposed [15]. However, these transformation methods are only invariant to zoom, rotation, and scale transformation, or a combination of these. A single instance of these transformations or their combination is not sufficient to express the actual changes of pose in real life, which can be satisfactorily approximated by an affine transformation. Rahtu et al. [16] have proposed an affine-invariant feature extraction method called "multi-scale autoconvolution" (MSA), which can attain more accurate identification results than other affine-invariant methods and cross-weighted moments. Several researchers have applied MSA features to recognition applications. However, only a few researchers have used MSA for face recognition.

Zhao et al. [17] were the first to use MSA features in pose-variant face recognition. However, they found that MSA features are sensitive to illumination changes and adapt poorly to different face databases. Hence, based on the past research, we propose a new face recognition algorithm in this paper. Our algorithm uses histogram equalization to mitigate the effects of illumination changes, and utilizes PCA to reduce the number of MSA dimensions and render MSA features more adaptable.

Our contributions in this paper are as follows:

1. We propose a new MSA + PCA pose-variant face recognition algorithm and verify its effectiveness using four common multi-pose face databases: the Olivetti Research Laboratory (ORL) database maintained by Cambridge University, the Audio Visual Technologies Group (GTAV) database at the Computer Vision Center of the Universitat Politécnica de Catalunya, the Pose, Illumination, and Expression (PIE) database of the Robotics Institute at Carnegie-Mellon University, and the datasets available at Pointing'04 Workshop at Cambridge University.
2. We analyze the applicability of MSA features to multi-pose face recognition tasks.
3. We study the influence of illumination changes on MSA features and offer solutions.
4. We analyze the influence of $(\alpha, \beta)$ pairs on MSA features as well as a range of suitable values, while pointing out the relationship between the number of $(\alpha, \beta)$ pairs and the recognition rate.
5. We verify the stability of MSA features against noise (of varying density and type).

The remainder of this paper is structured as follows: in Section 2, we present our pose- variant face recognition algorithm, and assess its effectiveness in Section 3. Section 4 is devoted to exhibiting the ability of our proposed algorithm through a series of experiments, whereas we offer our conclusions as well as directions for future research in Section 5.

## 2. Proposed algorithm

In this section, we introduce the MSA transform, test the classification performance of MSA features, and propose our face recognition algorithm.

### 2.1. MSA transform

We briefly summarize the MSA transform. Detailed derivation can refer to [16].

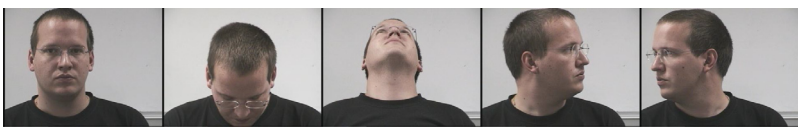



**Fig. 1.** Different real-life poses. Note that the pose is determined by two angles $(v, h)$, i.e., (vertical, horizontal), which varies from $-90°$ to $+90°$.

Let $f(x)$: $IR^2 \to IR$ be an image intensity function in $L^1(IR^2) \cap L^2(IR^2)$ and let $p(x) = \frac{1}{\|f\|_{L^1}} f(x)$ be the normalized version of $f$, so that $\int p(x)dx = 1$. Then, $p(x)$ is a probability density function, and we may take $X_0, X_1, X_2$ to be independent random variables with values in $IR^2$, so that $P(X_j = x_j) = \frac{1}{\|f\|_{L^1}} f(x_j)$.

For $\alpha, \beta \in R$, we define a random variable $U_{\alpha,\beta} = \alpha(X_1 - X_0) + \beta(X_2 - X_0) + X_0$.

Therefore, we obtain $U_{\alpha,\beta} = \alpha X_1 + \beta X_2 + \gamma X_0$, where $\gamma = 1 - \alpha - \beta$. Now, we can easily show that $U_{\alpha,\beta}$ has a probability density function

$$P(U_{\alpha,\beta} = u) = \frac{1}{\|f\|_{L^1}^3} (f_\alpha * f_\beta * f_\gamma)(u), \tag{1}$$

where $f_\alpha(x) = \alpha^{-2} f(x/\alpha)$ for $\alpha \neq 0$ and $f_\alpha = \|f\|_{L^1} \delta_0$ for $\alpha = 0$.

For $\alpha, \beta \in R$, we define the MSA transform of $f$ by $F(\alpha, \beta) = E\{f(U_{\alpha,\beta})\}$.

Writing this out in terms of the probability density function gives

$$F(\alpha, \beta) = \int f(u) P(U_{\alpha,\beta} = u) du = \frac{1}{\|f\|_{L^1}^3} \int f(u)(f_\alpha * f_\beta * f_\gamma)(u) du.$$

Taking the Fourier transform and using the convolution and correlation theorems, we obtain:

$$F(\alpha, \beta) = \frac{1}{\hat{f}(0)^3} \int_{IR^2} \hat{f}(-\xi)\hat{f}(\alpha\xi)\hat{f}(\beta\xi)\hat{f}(\gamma\xi) d\xi, \tag{2}$$

which holds for all $(\alpha, \beta)$.

In this paper, the $(\alpha, \beta)$ values refer to the experimental values of Rahtu et al., i.e., in the set $\{-1, -0.75, -0.5, -0.25, 0, 0.25, 0.5, 0.75, 1\}$, we select any possible combinations of two elements to form 29 $(\alpha, \beta)$ pairs.

### 2.2. Classification performance of MSA features

Given a gray-scale face image function $f$ and $(\alpha, \beta)$ pairs, we can obtain its MSA feature values using Eq. (2). In this subsection, we test the classification performance of MSA features using an experiment, the flow of which is shown in Fig. 2.

The dataset we adopted for the experiment was the Pointing'04 multi-pose face database [18]. The images in Figs. 1 and 3 were part of the database.

*K*-nearest classifiers, support vector machines (SVMs), and neural networks (NNs) are the typical classifiers used for face recognition. Therefore, we used them in our experiment to assess their influence on recognition rates.

The face images in Pointing'04—taken at $(v, h)$ angles $(90, 0)$, $(60, 30)$, $(30, -15)$, $(-15, -60)$, $(-15, 90)$, $(0, 45)$, $(15, 0)$, $(30, -45)$, $(60, -90)$, and $(60, 60)$—were used as the training image sets to train the classifiers. All the other 83 head-pose images were used to assess the performance of the classifiers. Table 1 shows the recognition rate of each.

The SVMs and NNs classifiers are more complex than the *k*-nearest classifier and usually deliver good recognition results. However, in our experiment, as shown in Table 1, different types of classifiers had little effect on the recognition results, which showed that MSA features are only slightly dependent on the types of classifiers.

The low recognition rates of the three classifiers (40.7%, 42.9%, and 43.7%) are unsatisfactory with regard to our requirements on MSA features. Therefore, we carried out another experiment to investigate the reasons for the poor performance.

We chose five sets of face images from the Pointing'04 database: person1, person2, person3, person4, and person5. Each set contained five different poses. The $(v, h)$ angles for these were $(0, 0)$ $(-90, 0)$ $(90, 0)$ $(0, -90)$ $(0, 90)$. Partial poses of the five subjects are shown in Fig. 3. We extracted the MSA features of each subject in the five different poses. The MSA features of different faces showed great similarity, as shown in Fig. 4.

From Fig. 4, we can see that the MSA values for different subjects at different poses were almost identical because of the great similarity of human faces.

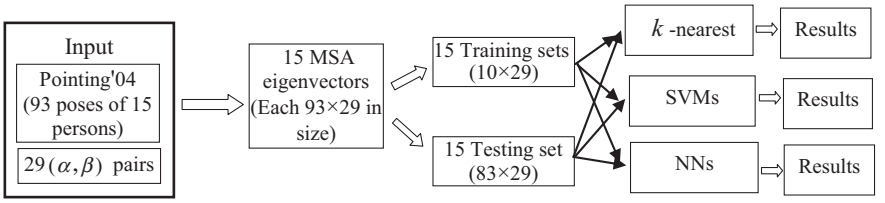
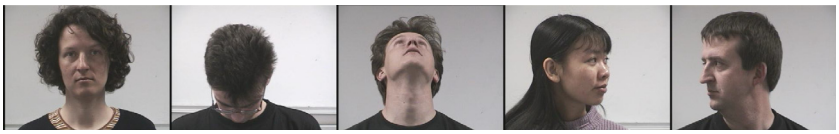



**Fig. 2.** Experimental flow.



person1 (0, 0) person2 (-90, 0) person3 (90, 0) person4 (0,-90) person5 (0, 90)

**Fig. 3.** Different subjects at different poses.

**Table 1**
Recognition rates of three classifiers.

| Classifier | Recognition rate (%) | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Person Id | 01 | 02 | 03 | 04 | 05 | 06 | 07 | 08 |
| *k*-nearest | 44.4 | 55.6 | 44.4 | 55.6 | 33.3 | 55.6 | 33.3 | 33.3 |
| SVMs | 44.4 | 55.6 | 44.4 | 55.6 | 33.3 | 55.6 | 33.3 | 33.3 |
| NNs | 44.4 | 55.6 | 44.4 | 55.6 | 33.3 | 55.6 | 33.3 | 44.4 |
| Person Id | 09 | 10 | 11 | 12 | 13 | 14 | 15 | Total |
| *k*-nearest | 55.6 | 55.6 | 33.3 | 33.3 | 33.3 | 22.2 | 22.2 | 40.7 |
| SVMs | 55.6 | 55.6 | 33.3 | 33.3 | 33.3 | 33.3 | 44.4 | 42.9 |
| NNs | 55.6 | 55.6 | 33.3 | 33.3 | 33.3 | 33.3 | 44.4 | 43.7 |

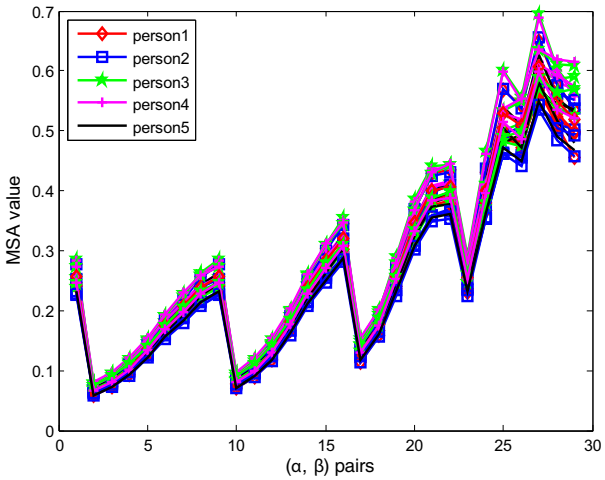



**Fig. 4.** MSA values before PCA.

### 2.3. Proposed MSA + PCA algorithm

Due to the great similarity of human faces, the classification performance of MSA features was poor (as shown in Table 1). Furthermore, we noticed the following in Fig. 4: the images of different subjects recorded similar MSA values, which renders classification difficult for any classifier. In pattern recognition field, the basic principle involves increasing the distinction among human faces on the premise of maintaining the invariance. Therefore, we adopt the classical PCA method for our algorithm. The function of PCA is to decorrelate the feature sets and reduce the number of required MSA dimensions. Dimension reduction is carried out such that only the PCA components with sufficiently large corresponding eigenvalues are passed to the classifiers.

Following PCA, the values in Fig. 5 that are close to zero present the non-principal components that contribute little to classification. The other values represent the principal components that are sufficiently distinct from one another for a classifier, all the while maintaining its invariance.

Our proposed MSA + PCA algorithm is as follows.

Algorithm: MSA + PCA

Step 1: Calculate MSA values

Input: $I$ is a gray-scale image matrix obtained histogram equalization, and *ind* is a column vector containing the $(\alpha, \beta)$ values.
1: Calculate the density function $p1, p2, p3$ of $I$.
2: Calculate the probability density function $P1, P2, P3$ by a double convolution.
3: Calculate the discrete Fourier transform pairs of $P1, P2, P3$ to obtain the resultant $F$.
Output: $F$ is a vector containing MSA values corresponding to the given coefficients $(\alpha, \beta)$ defined in the matrix *ind*.

Step 2: Calculate PCA components

Input: $F$.
1: Load $F$.
2: Construct the training datasets *RD* and the test datasets *ED*.
3: Calculate the covariance matrix $G$ of *RD*.
4: Calculate the eigenvalue *EV* and eigenvector *EC* of $G$.
5: Sort *EC*, which has the max *EV*; obtain the transformation matrix *TM*.
6: $RD \times TM$; obtain the primary components $R1$ for the training datasets.
7: $ED \times TM$; obtain the primary components $R2$ for the test datasets.
Output: PCA components $R1, R2$.

Step 3: Training and classification

Input: $R$.
1: Load $R$.
2: Train the $k$-nearest classifier with $R1$.
3: Test the classifier with $R2$.
4: Obtain the classification outputs $g$.
Output: Classification results $g$.

## 3. Analysis of the proposed algorithm

In this section, we theoretically analyze the performance of our proposed algorithm.

### 3.1. Applicability of MSA features

Many researchers have adopted MSA features for simple target recognition, such as letters, road-oriented markings, objects from the Columbia Object Image Library (COIL)-100 database, etc. Because the human face includes rich textures, facial expressions, accessories, etc., we analyze the applicability of MSA features to multi-pose face recognition.

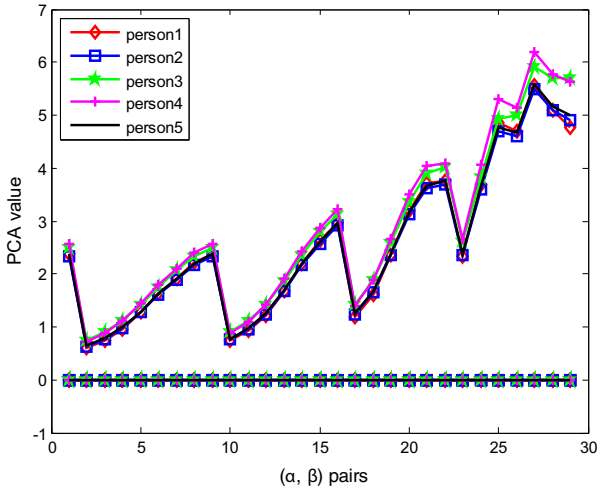



**Fig. 5.** MSA values after PCA.

Let us assume that a face image $f(x)$ undergoes an affine transformation such that $f'(x) = f(Tx + t)$, where $T$ is a nonsingular matrix and $t$ is a translation vector. Similarly, we define $X'_0, X'_1, X'_2$ in an affine space:

$$X'_0 = TX_0 + t, X'_1 = TX_1 + t, X'_2 = TX_2 + t, \quad U'_{\alpha,\beta} = \alpha(X'_1 - X'_0) + \beta(X'_2 - X'_0) + X'_0.$$

Hence, $U'_{\alpha,\beta} = TU_{\alpha,\beta} + t$, $F'(\alpha, \beta) = F(\alpha, \beta)$.

The MSA eigenvectors $F(\alpha, \beta)$ in the original image space and $F'(\alpha, \beta)$ in the affine image space are equivalent. Therefore, the transform coefficients $F(\alpha, \beta)$ are invariant against the affine transformation of the image coordinates. This provides a method to obtain affine-invariant features for an image function $f$. Since multi-pose faces in real life satisfy affine transformations, we can adopt the MSA transform in pose-variant face recognition tasks.

We extracted the MSA features of the subjects shown in Fig. 1 and found that the MSA eigenvectors extracted showed satisfactory pose-invariance and robustness; i.e., for images of the same subject in different poses, the MSA features remained stable, as shown in Fig. 6.

### 3.2. Influence of $(\alpha, \beta)$ pairs

Given the image of a human face, we can calculate the MSA values by Eq. (2). By changing the $(\alpha, \beta)$ pairs, we can obtain an infinite number of affine-invariant MSA features. However, this raises the question of the $(\alpha, \beta)$ value to select with respect to an infinite number of $(\alpha, \beta)$ pairs.

When projected to the $(\alpha, \beta)$ plane, the equality $F(\alpha, \beta) = F(\beta, \alpha)$ causes diagonal symmetry with respect to the line $\alpha - \beta = 0$, i.e., symmetry in the direction of the $\beta$-axis with respect to the line $\alpha + 2\beta = 1$, and that in the direction of the $\alpha$-axis with respect to the line $2\alpha + \beta = 1$. By using these symmetries, we can find small regions in the $(\alpha, \beta)$ plane, that essentially provide all MSA transform values. There are several possible values, one of which is shown in Fig. 7.

Fig. 7 shows the probable region in the $(\alpha, \beta)$ plane. However, there still remain an infinite number of possible choices. The convoluted nature of the MSA transform produces a smooth continuous surface. Therefore, $(\alpha, \beta)$ pairs that lie in close proximity to one another provide highly correlated transform values.

We adopt three sets of $(\alpha, \beta)$ pairs to observe their influence on MSA values. The first collection of sets, called pair1, is $(0, -0.5)$, $(0.4, 0.8)$, $(-0.2, 0.9)$, and $(0.3, -0.5)$; the second collection is $(-1, 1)$, $(-0.8, 0.8)$, $(-0.6, 0.6)$, $(-0.6, 0.8)$, $(-0.4, 0.6)$, $(-0.4, 0.4)$, $(-0.2, 0.2)$, $(-0.2, 0.4)$, $(-0.2, 0.6)$, $(0, 0.2)$, $(0, 0.4)$, $(0.2, 0.2)$, and $(0.2, 0.4)$, and is called pair2; the third collection consists of the 29 $(\alpha, \beta)$ pairs used by Rahtu et al., and is called pair3. The results in Fig. 8 show that different $(\alpha, \beta)$ pairs only affect the MSA values and the size of the feature vectors.

### 3.3. Influence of illumination changes

The Face Recognition Vendor Test (FRVT) 2002 [19] confirmed that illumination and pose variations are the two major problems plaguing existing face recognition systems. Thus, a pertinent question in the context of our research relates to the influence of illumination changes on MSA features.

Suppose $f(x)$: $IR^2 \to IR, f \geqslant 0$ is an image intensity function in $IR^2$. Let $x_0, x_1, x_2$ be three random points in $f$, then, $u_{\alpha,\beta} = \alpha(x_1 - x_0) + \beta(x_2 - x_0) + x_0$, where $(\alpha, \beta)$ are the coordinates of $u$ in the space spanned by the vectors $x_1 - x_0$ and $x_2 - x_0$ with origin at $x_0$.

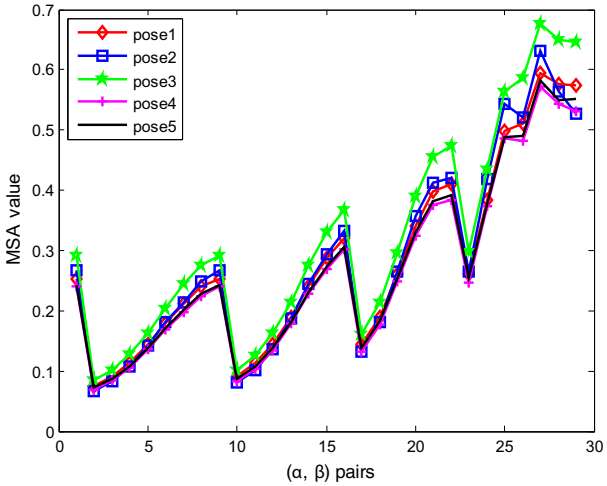



**Fig. 6.** MSA values of the same subject in five different poses.

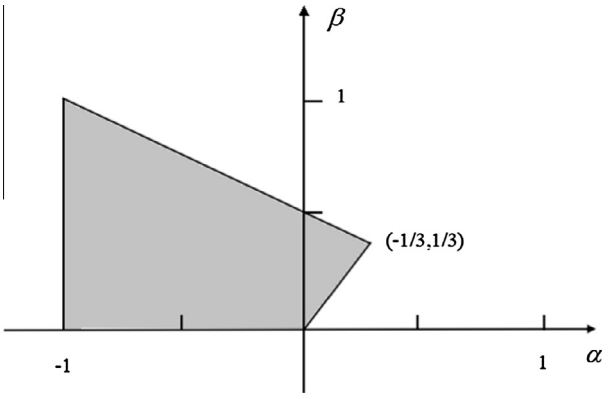



**Fig. 7.** The $(\alpha, \beta)$ plane region.

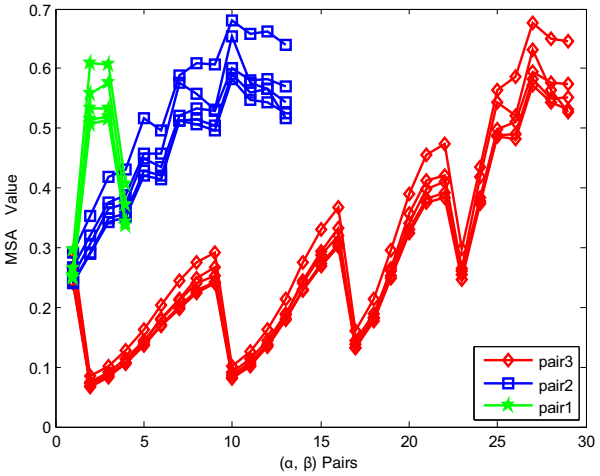



**Fig. 8.** Effect of different $(\alpha, \beta)$ pairs.

Now let $A\{T, t\}$ be an affine transformation. If we take $A$ to the three points $x_0$, $x_1$, $x_2$, then $x'_0 = Tx_0 + t$, $x'_1 = Tx_1 + t$, $x'_2 = Tx_2 + t$, and $u'_{\alpha,\beta} = \alpha(x'_1 - x'_0) + \beta(x'_2 - x'_0) + x'_0 = \alpha(Tx_1 - Tx_0) + \beta(Tx_2 - Tx_0) + Tx_0 + t$.

Thus, $u'_{\alpha,\beta} = Tu_{\alpha,\beta} + t$.

Let $x_0, x_1, x_2$ be samples of the random variables $X_0, X_1, X_2$, and define a new random variable $U_{\alpha,\beta} = \alpha(X_1 - X_0) + \beta(X_2 - X_0) + X_0$.

Similarly,

$U'_{\alpha,\beta} = \alpha(X'_1 - X'_0) + \beta(X'_2 - X'_0) + X'_0$, where $X'_0 = TX_0 + t$, $X'_1 = TX_1 + t$, $X'_2 = TX_2 + t$.

Hence $U'_{\alpha,\beta} = TU_{\alpha,\beta} + t$. Substituting $x = U'_{\alpha,\beta}$, we obtain

$$f'(U'_{\alpha,\beta}) = f(T^{-1}U'_{\alpha,\beta} - T^{-1}t) = f(T^{-1}(TU_{\alpha,\beta} + T) - T^{-1}t) = f(U_{\alpha,\beta}).$$

The MSA transformation treats an image as a probability density function, and the probability distribution remains stable when the image undergoes an affine transformation $A\{T, t\}$. Thus, the MSA transform is sensitive to illumination changes caused by affine transformation.

We chose five images from the PIE database [20], which reflects extreme changes in illumination, called illumination1, illumination2, illumination3, illumination4, and illumination5, as shown in Fig. 9. We then extracted the MSA features of these images. The results are shown in Fig. 10. The illumination changes showed a significant effect on MSA values. The blue[1] curve represents the lightest face, whereas the pink curve represents the darkest one.

We can typically reduce the illumination effect according to different lighting models. However, in real-world conditions, it is difficult to obtain a model for it. Therefore, we reduced the illumination effect by equalizing the histogram. The illumination effect can thus be correspondingly reduced, as shown in Fig. 11.

[1] For interpretation of color in Fig. 10, the reader is referred to the web version of this article.

**Fig. 9.** Five illumination changes in images from the PIE database.

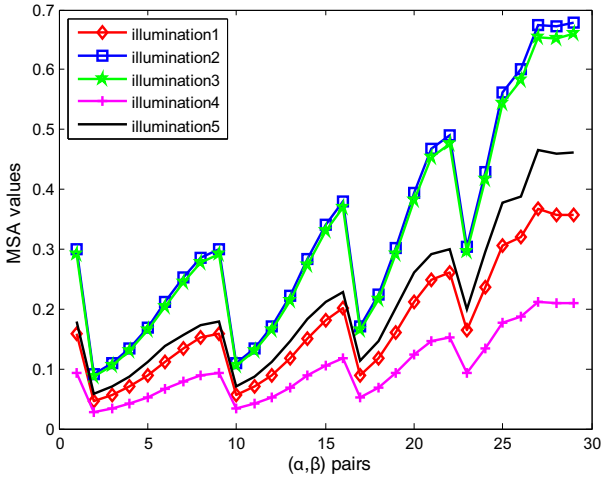



**Fig. 10.** Illumination effect before equalizing histogram.

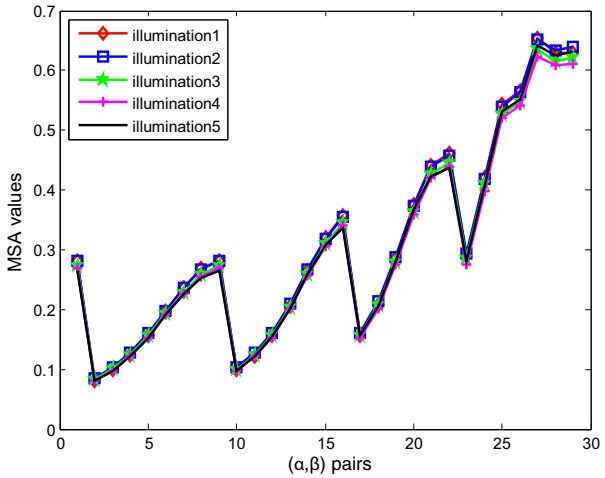



**Fig. 11.** Illumination effect after equalizing histogram.

### *3.4. Influence of noise*

The images in face recognition tasks are typically noisy [21]. Therefore, in this section, we study the influence of noise on MSA features.

We first added some Gaussian noise (with 0 mean and 0.01variance) to the face in Fig. 1 using five different density values (0.01, 0.03, 0.05, 0.07, and 0.09), called noise1, noise2, noise3, noise4, and noise5, respectively. The noise density used was the probability of a pixel of being disturbed by noise.

Following this, we repeated the same experiment with different types of noise—Gaussian noise, Salt-and-Pepper noise, Poisson noise, and Speckle noise—with a density of 0.05. The results in Figs. 12 and 13 show that the MSA features remain invariant to noise of different densities and types.

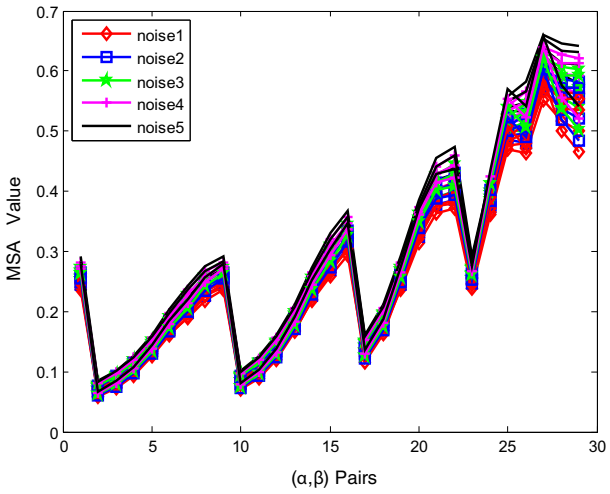



**Fig. 12.** Effects of Gaussian noise.

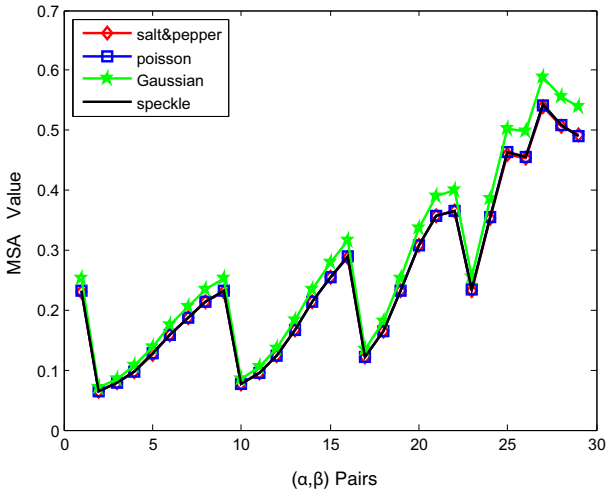



**Fig. 13.** Effects of different types of noise.

## 4. Experiments

### *4.1. Datasets used in our experiments*

In order to verify the potential discriminative power of our proposed MSA + PCA algorithm, we tested our algorithm using the Pointing'04, PIE, ORL, and GTAV databases.

The Pointing'04 database consists of 15 sets of face images in various skin colors. Each set contains 93 images of the same subject at different poses, and wearing glasses or not. The pose or head orientation is determined by two angles $(v, h)$, which vary from $-90°$ to $+90°$. The Pointing'04 dataset contains complete pose changes, and each pose corresponds to a $(v, h)$ parameter. Therefore, the dataset is often used for pose-variant face recognition or face orientation estimation.

The PIE database contains 41,368 facial images of 68 subjects, where each person has been photographed in 13 different poses, under 43 illumination conditions, and assuming 4 different expressions. Capturing images of every subject under every possible combination of pose, illumination, and expression is not practical because of the massive amount of storage space required. The PIE database, therefore, consists of two major partitions: one with pose and illumination variations, and the other with pose and expression variations. We used the former for this experiment, which consists of 68 subjects each of 170 images at different poses and under varying illumination. The subjects shown in Fig. 9 are part of this database.

The ORL database [22] contains 400 images of 40 subjects in varying poses and expressions. The GTAV database [23] comprises 44 subjects each of 27 images corresponding to different poses, views (0°, ±30°, ±45°, ±60°, and ±90°), and under 3 illumination conditions (natural light, a strong light source from a 45° angle, and an almost frontal mid-strong light source).

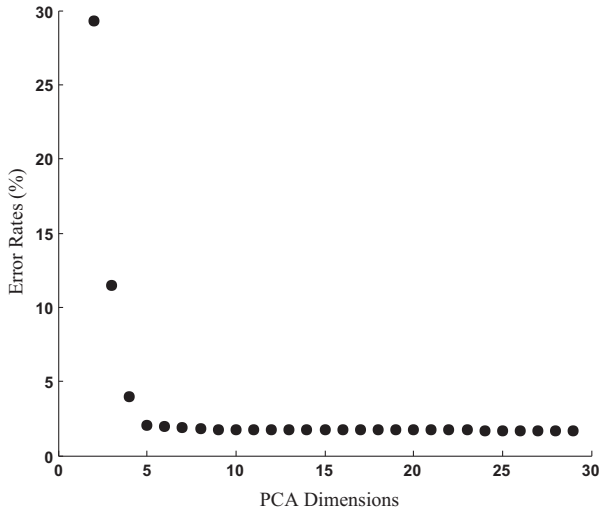



**Fig. 14.** Relationship between recognition rate and the number of PCA dimensions.

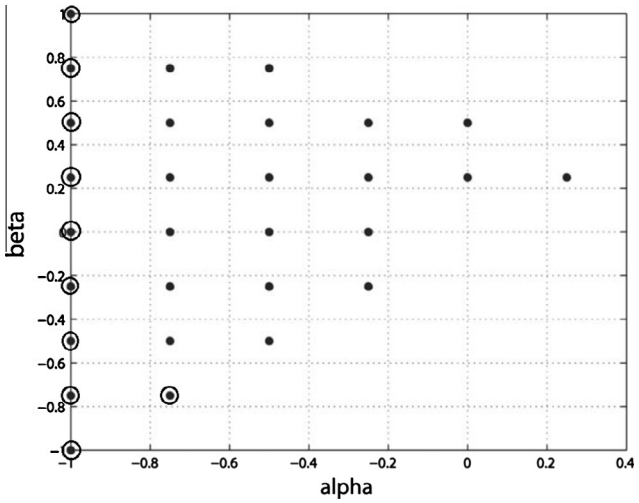



**Fig. 15.** $(\alpha, \beta)$ space after PCA.

In Section 2.2, we confirmed that MSA features exhibit excellent performance that is independent of different classifiers. Therefore, we adopted the simplest *k*-nearest classifier in the following experiments.

### *4.2. Classification performance on Pointing'04 database*

In the first case, we used the Pointing'04 multi-pose human face database to test our algorithm. The Pointing'04 database was introduced in Section 2.2, and the training and testing sets used for these experiments were identical to those described in Section 2.2.

The Pointing'04 database consists of 15 sets of subjects, each of which contains 93 images of the same subjects at different poses. Therefore, the size of each feature matrix was $93 \times 29$, where 29 is the number of $(\alpha, \beta)$ pairs, as shown in Fig. 15 (black points). Following PCA, the size of the matrix was reduced to $29 \times 10$. The corresponding ten $(\alpha, \beta)$ pairs are shown in Fig. 15 (black points with circles), where the value for $\alpha = -1$ corresponding to nine pairs and the value for $\alpha = -0.75$ corresponding to one pair represent the first 10 $(\alpha, \beta)$ pairs when we calculated the original MSA features.

The relationship between the error rates and PCA dimensions are shown in Fig. 14. When the number of PCA dimensions equals 2, the error rate is close to 40%. When the number of PCA dimensions increases to 10, the error rates reduce to 3.33% and remain invariant. Thus, in the following experiments, we maintained the PCA feature size at 10.

We wanted to investigate the relationship between the number of $(\alpha, \beta)$ pairs and error rates. The $(\alpha, \beta)$ pairs used were the same as shown in Fig. 8: pair1, pair2, and pair3. The results in Fig. 16 show that a larger number of $(\alpha, \beta)$ pairs resulted in

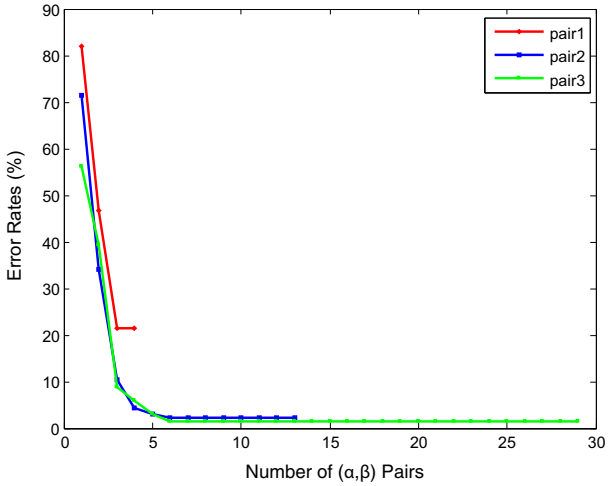



**Fig. 16.** Relationship between error rates and number of $(\alpha, \beta)$ pairs.

higher recognition rates. However, a larger number of $(\alpha, \beta)$ pairs implies higher computational complexity, as shown in Table 2. The elapsed time is the computing time for the images of each subject in Poingting'04 during the MSA feature extraction period with standard personal computer configurations (Windows 7 OS, Intel(R) Core(TM) processor, 2 GB random-access memory (RAM), MATLAB 7.8). Therefore, we needed to balance the number of $(\alpha, \beta)$ pairs and recognition rates.

### 4.3. Classification performance on PIE, ORL, and GTAV databases

We partitioned the dataset by randomly choosing 15 images per subject as training and the rest as testing datasets on PIE database.

**Table 2**
The comparison of elapsed time of different pairs.

| $(\alpha, \beta)$ pairs | Elapsed time |
|---|---|
| Pair1: 4 pairs of (0, −0.5), (0.4, 0.8), (−0.2, 0.9), (0.3, −0.5) | 36 s |
| Pair2: 13 pairs of (−1, 1), (−0.8, 0.8), (−0.6, 0.6), (−0.6, 0.8), (−0.4, 0.6), (−0.4, 0.4), (−0.2, 0.2), (−0.2, 0.4), (0.2, 0.6), (0, 0.2), (0, 0.4), (0.2, 0.2), (0.2, 0.4) | 132 s |
| Pair3: 29 pairs of [16] | 427 s |

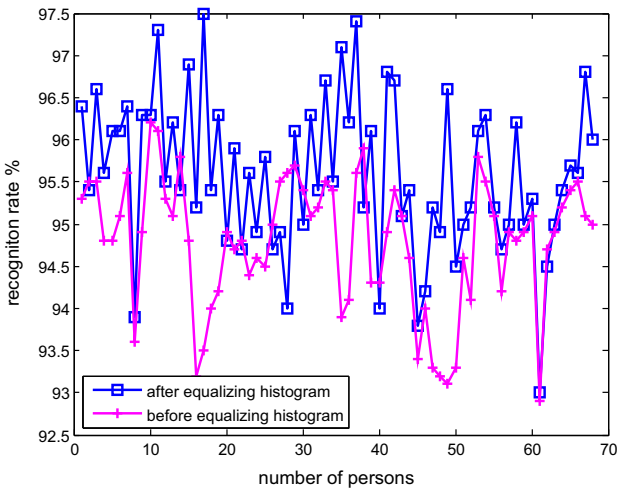



**Fig. 17.** Recognition rate on PIE database.

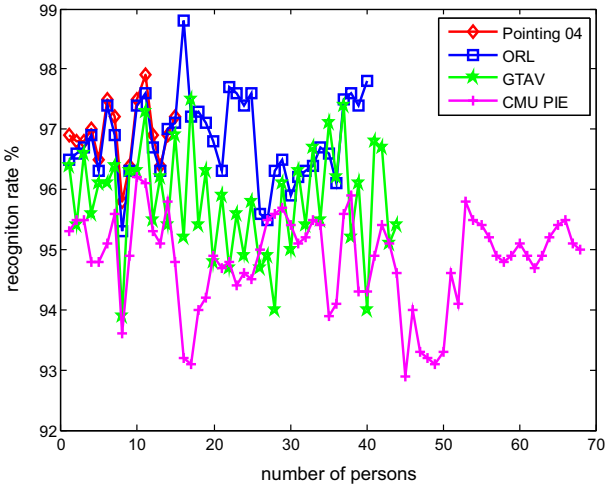



**Fig. 18.** Recognition rates on four databases.

ORL database.

GTAV database.

**Fig. 19.** Partial face images of ORL and GTAV databases.

Fig. 17 compares the recognition rates before and after equalizing the histogram. Equalization can reduce the influence of illumination changes.

We partitioned the ORL dataset by randomly choosing 5 images per subject as training and the rest as testing datasets. For the GTAV dataset, we used a random subset of 8 images per subject as training and the rest as testing datasets.

Fig. 18 shows the recognition rates of our algorithm for four databases. Its average recognition rate on Pointing'04, ORL, GTAV, and PIE was 96.7%, 96.83%, 95.79%, and 94.82%, respectively. The number of subjects tested in the four databases was 15, 40, 44, 68, respectively (see Fig. 19).

### *4.4. Classification performance comparison with other algorithms*

We compared the recognition rate of our proposed algorithm with other pose-variant face recognition methods. The results are shown in Table 3.

**Table 3**
Comparison of recognition rates of different algorithms.

| Database | Recognition rate (%) Literature 1 | Recognition rate (%) Literature 2 | Recognition rate (%) PCA | Recognition rate (%) LDA | Recognition rate (%) Our work |
|---|---|---|---|---|---|
| ORL | 97.25 | 96.5 | 88.1 | 93.9 | 96.7 |
| GTAV | 64.9 | Not mentioned | Not mentioned | Not mentioned | 95.79 |
| PIE | Not mentioned | 91.3 | 62.1 | 89.1 | 92.82 |
| Pointing'04 | Not mentioned | Not mentioned | Not mentioned | Not mentioned | 96.7 |

Note: Literature 1 refers to [24]; Literature 2 refers to [25].

## 5. Conclusions

In this paper, we proposed a new pose-variant face recognition algorithm based on MSA features and the PCA dimension-reduction method. The proposed algorithm was analyzed theoretically and assessed experimentally using four common multi-pose face databases.

In the theoretical part of this paper, we proposed an MSA + PCA algorithm and analyzed its applicability to pose-variant face recognition tasks. We also evaluated the influence of $(\alpha, \beta)$ pairs, illumination changes, and noise on MSA features. The results of experiments using the ORL, GTAV, PIE, and Pointing'04 multi-pose face databases showed that the proposed algorithm attains higher recognition rates than prevalent pose-variant face recognition methods, and has a clear advantage of weakly dependence on the database and classifiers.

In pattern recognition field, a balance is often required between performance and computational complexity. Higher performance usually implies higher computational complexity. The limitation of our proposed algorithm is that it is computationally expensive during MSA feature extraction because it involves more $(\alpha, \beta)$ pairs. During our experiment using MATLAB 7.8, the computation time for images of each subject in the Pointing'04 database (with a resolution of $384 \times 288$) was 427.486505 s, given the 29 $(\alpha, \beta)$ pairs and standard personal computer configurations. The computation time for the PCA and subsequent classification was 5.014438 s. The reduction of computational complexity will involve reducing the number of $(\alpha, \beta)$ pairs. However, this will result in lower recognition rates.

In the future, we intend to investigate methods to mitigate the complexity of our proposed algorithm. Furthermore, we will apply the MSA + PCA algorithm to real-time multi-pose face recognition tasks by extracting MSA features in offline mode, and conducting training and classification in online mode.

## Acknowledgments

The work was supported by the National Nature Science Foundation of China (Grant No. 61202151), the Project of Shandong Province Higher Educational Science and Technology Program (No. J12LN34, No. J13LG86), and the Program for Scientific Research Innovation Team in Colleges and Universities of Shandong Province. We would like to thank anonymous reviewers for their valuable comments and suggestions, which were used to significantly improve this paper. We would like to extend our appreciation and gratitude to the owners of the Pointing'04, ORL, GTAV, and PIE databases.

## References

[1] Zhang Xiaozheng, Gao Yongsheng. Face recognition across pose: a review. Pattern Recogn 2009;42:2876–96.
[2] Ahonen T, Hadid A, Pietiainen M. Face description with local binary patterns: application to face recognition. IEEE Trans Pattern Anal Mach Intell 2006;28(12):2037–41.
[3] Zhao H, Yuen PC. Incremental linear discriminant analysis for face recognition. IEEE Trans Syst Man Cybern Part B (Cybern) 2008;38(1):210–21.
[4] Meshgini Saeed, Aghagolzadeh Ali, Seyedarabi Hadi. Face recognition using Gabor-based direct linear discriminant analysis and support vector machine. Comput Electr Eng 2013;39(3):727–45.
[5] Prince SJD, Warrell J, Elder JH, Felisberti FM. Tied factor analysis for face recognition across large pose differences. IEEE Trans Pattern Anal Mach Intell 2008;30(6):970–84.
[6] Safayani Mehran, Manzuri Shalmani Mohammad Taghi. Three-dimensional modular discriminant analysis (3DMDA): a new feature extraction approach for face recognition. Comput Electr Eng 2011;37(5):811–23.
[7] Gross R, Matthews I, Baker S. Appearance-based face recognition and light-fields. IEEE Trans Pattern Anal Mach Intell 2004;26(4):449–65.
[8] Abate AF, Nappi M, Riccio D. 2D and 3D face recognition: a survey. Pattern Recogn Lett 2007;28(14):1885–906.
[9] Bowyer KW, Chang K, Flyun P. A survey of approaches and challenges in 3D and multi-model 2D + 3D face recognition. J Comput Vision Image Understand 2006:1–15.
[10] González-Jiménez D, Alba-Castro JL. Toward pose-invariant 2-D face recognition through point distribution models and facial symmetry. IEEE Trans Inf Forensic Secur 2007;2(3–1):413–29.
[11] Kahraman F, Kurt B, Gokmen M. Robust face alignment for illumination and pose invariant face recognition. In: Proceedings of the IEEE conference on CVPR; 2007. p. 1–7.
[12] Chai X, Shan S, Chen X, Gao W. Locally linear regression for pose-invariant face recognition. IEEE Trans Image Process 2007;16(7):1716–25.
[13] Xie X, Lam K-M. Gabor-based kernel PCA with doubly nonlinear mapping for face recognition with a single face image. IEEE Trans Image Process 2006;15(9):2481–92.
[14] Yang J, Frangi AF, Yang J, Zhang D, Jin Z. KPCA plus LDA: a complete kernel Fisher discriminant framework for feature extraction and recognition. IEEE Trans Pattern Anal Mach Intell 2005;27(2):230–44.
[15] Xiong HL, Zhang TX, Moon YS. A translation and scale invariant adaptive wavelet transform. IEEE Trans Image Process 2000;9(12):2100–8.
[16] Rahtu E, Salo M, Heikkil J. Affine invariant pattern recognition using multiscale autoconvolution. IEEE Trans Pattern Anal Mach Intell 2005;27(6):908–18.
[17] Zhao Youen, Li Li, Yan Hua. Multi-pose face recognition using the primary component of multi scale autoconvolution. ICIC Express Lett Part B Appl 2013;4(6):1719–25.
[18] http://www-prima.inrialpes.fr/Pointing04/data-face.html.
[19] Phillips PJ, Grother P, Micheals RJ, Blackburn DM, Tabassi E, Bone JM. Face Recognition Vendor Test (FRVT 2002). National Institute of Standards and Technology. Evaluation report IR 6965; 2003.
[20] http://mu.edu/projects/project_418.html.
[21] Shen XiaoHong, Wang Kai, Guo Qiang. Local thresholding with adaptive window shrinkage in the contourlet domain for image denoising. Sci China Inf Sci 2013;56(9). 092107:1–9.
[22] http://am.ac.uk/research/dtg/attarchive/fa6cedatabase.html.
[23] http://gps-tsc.upc.es/GTAV/ResearchAreas/UPCFaceDatabase/GTAVFaceDatabase.htm.
[24] Wijaya I Gede Pasek Suta, Uchimura Keiichi, Koutaki Gou. Multi-pose face recognition using fusion of scale invariant features. In: Proceedings of the 2011 2nd international congress on computer applications and computational science; 2011. p. 207–13.

[25] Wright John, Hua Gang. Implicit elastic matching with random projections for pose-variant face recognition. In: Proceedings of 2009 IEEE computer society conference on computer vision and pattern recognition workshops, CVPR workshops; 2009. p. 1502–9.

**Youen Zhao** received her B.E. and M.E. from Shandong Normal University in 1997 and in 2004, respectively. She is currently an associate professor at Shandong University of Finance and Economics. Her research interests include pattern recognition and image processing.

**Li Li** received her B.E. and M.E. from Shandong Technology University in 1998 and Shandong University in 2000, respectively, and her Ph.D. from Shandong University in 2011. She is currently an associate professor at Shandong University of Finance and Economics. Her research interests include computer vision, image and video processing.

**Zhaoguang Liu** received his B.E. and M.S. from Xi'an University of Technology in 1998 and 2001, respectively, and his Ph.D. in School of Information Science and Engineering from Shandong University in 2009. He is now an associate professor at Shandong University of Finance and Economics. His current research interests include video analysis and digital image processing.