Journal of
Electronic Imaging

SPIEDigitalLibrary.org/jei

Rotation and noise invariant near-infrared face recognition by means of Zernike moments and spectral regression discriminant analysis

Sajad Farokhi
Siti Mariyam Shamsuddin
Jan Flusser
Usman Ullah Sheikh
Mohammad Khansari
Kourosh Jafari-Khouzani

SPIE

IS&T
imaging.org

Rotation and noise invariant near-infrared face recognition by means of Zernike moments and spectral regression discriminant analysis

Sajad Farokhi
Siti Mariyam Shamsuddin
Universiti Teknologi Malaysia
Faculty of Computing
81310, Johor, Malaysia
E-mail: fsajad2@live.utm.my

Jan Flusser
Institute of Information Theory and Automation of the Academy of Sciences of the Czech Republic
182 08, Prague, Czech Republic

Usman Ullah Sheikh
Universiti Teknologi Malaysia
Faculty of Electrical Engineering
81310, Johor, Malaysia

Mohammad Khansari
University of Tehran
Faculty of New Sciences and Technologies
14399-55941, Tehran, Iran

Kourosh Jafari-Khouzani
Athinoula A. Martinos Center for Biomedical Imaging
Department of Radiology
Massachusetts General Hospital, Harvard Medical School
02114, Boston, Massachusetts

Abstract. *Face recognition is a rapidly growing research area, which is based heavily on the methods of machine learning, computer vision, and image processing. We propose a rotation and noise invariant near-infrared face-recognition system using an orthogonal invariant moment, namely, Zernike moments (ZMs) as a feature extractor in the near-infrared domain and spectral regression discriminant analysis (SRDA) as an efficient algorithm to decrease the computational complexity of the system, enhance the discrimination power of features, and solve the "small sample size" problem simultaneously. Experimental results based on the CASIA NIR database show the noise robustness and rotation invariance of the proposed approach. Further analysis shows that SRDA as a sophisticated technique, improves the accuracy and time complexity of the system compared with other data reduction methods such as linear discriminant analysis. © 2013 SPIE and IS&T [DOI: 10.1117/1.JEI.22.1.013030]*

Paper 12370 received Sep. 9, 2012; revised manuscript received Jan. 29, 2013; accepted for publication Feb. 14, 2013; published online Mar. 1, 2013.

0091-3286/2013/$25.00 © 2013 SPIE and IS&T

1 Introduction

Although face recognition seems to be an easy task for a human, automatic face recognition is still a challenging problem due to variations in illumination, facial expression, and head pose.[1,2] Techniques that can provide efficient feature extraction with high discrimination power and low computational complexity are crucial. Based on the way of considering spectrum information, face-recognition methods may be divided into two main categories: visible (VIS) and infrared (IR). During the last two decades, most studies in the field of face recognition have focused only on visible imagery due to availability of visible cameras and high quality of acquired images. However, the main problem with these approaches is a high dependency of system performance on external light, angle of light, and even skin color of people.[3–5] To date, different methods have been proposed to solve the illumination problem and improve the performance of algorithms.[6] Recently near-infrared (NIR) imagery has received much attention due to the high quality of acquired images as well as high performance of NIR cameras under illumination variations.[7] The volume of literature in the NIR domain is very limited in comparison with visible or thermal IR domain. Zou et al. offered an important discussion on proposing a new method in the NIR domain to solve the illumination problem.[8] The Fisherface technique is used as a feature-extraction step.[9] The authors were successful in proposing active NIR as the approach to address the illumination problem in face-recognition systems; however, they did not offer an adequate discussion on some common challenges, such as facial expression and noise. Moreover, the feature-extraction procedure requires a lot of images in the training set to represent good results, and Xiaoyang et al. showed that the performance of the principal component analysis (PCA) drops sharply with the decreasing number of training samples for each person.[10] As a result, the feature-extraction process is not only time-consuming but also needs a large memory to process the features. Hence it cannot be used for large-scale high-dimensional data.

Li et al. introduced a novel design of active NIR camera hardware and enhanced algorithm to propose an illumination invariant face-recognition system.[11] Local binary pattern (LBP) features are extracted to compensate monotonic transform and obtain illumination invariant face representation.[12] Finally two statistical learning algorithms are used to decrease the dimension of features and extract most discriminative features. However, a serious weakness of this approach is the high sensitivity of LBP features to noise, slight head pose, and alignment errors.[13,14] Another problem with this method is applying the statistical learning algorithms on high-dimensional LBP features (748,592-dimensional vector), which may lead to a complex and time-consuming learning process with large memory requirements.

A new method based on discrete wavelet transform and two-dimensional (2-D) principal component analysis (2DPCA) was introduced by He et al.[15] The proposed method highlighted the efficiency of wavelet transform in the NIR domain for the first time. Despite its strength, there are some issues, such as head pose and noise, which are not fully addressed. Moreover, the wavelet transform used in this approach is translation variant and a simple shift of the input signal may change the coefficients of wavelet transform considerably.[16,17]

The research to date has tended to focus on designing NIR hardware to solve the illumination problem, and there has been little agreement on proposing an accurate NIR face-recognition method with respect to most common challenges. Furthermore, collecting sufficient prototype images capable of covering all the challenges is truly impossible in real face-recognition systems. Hence, proposing a small-size training face-recognition system, which is robust to most of the variations, is still a challenging problem in the face-recognition area.

In this paper, we propose a global feature-based method based on Zernike moments (ZMs) and spectral regression discriminant analysis (SRDA) to deal with facial expressions, head pose, small number of images in the training set, image rotation, noise, and high time complexity of systems. Experimental results show that our proposed algorithm overcomes the shortcomings of other methods in the presence of tilt (image rotation) and noise. In addition, it enhances the feature-extraction procedure considerably.

The contribution of this paper consists of the following:

- Application of ZMs coupled with SRDA in NIR domain to generate salient features.
- Design of a face-recognition system in the NIR domain, which has good accuracy with small number of images in the training set.
- Design of a face-recognition system in the NIR domain, which is fast, accurate, and suitable for real-time face recognition.

The remainder of this paper is organized as follows: in Sec. 2, we provide a brief review of ZMs. In Sec. 3, SRDA is explained. The proposed system is expressed in Sec. 4. Experimental results and performance analysis are given in Sec. 5. Final conclusion and discussion are presented in Sec. 6.

2 Zernike Moment Invariants

2.1 *Introduction and Definition*

ZMs have been used widely as a powerful feature extractor in pattern-recognition systems with satisfactory results. Their history dates to 1980 when Teague introduced ZMs based on the theory of orthogonal polynomials in image analysis and constructed useful moments which were invariant regarding to rotation.[18,19] The application of ZMs as global invariants has generated a lot of interest in face-recognition algorithms due to their rotation invariance property and their high robustness to noise.[20–24] Furthermore, they have proven to be better than other orthogonal moments in terms of computation time, feature representation capability, and multilevel representation for describing the shapes of patterns.[25] The kernel of ZMs is a set of orthogonal Zernike polynomials defined on a unit circle in Cartesian coordinates. ZMs are a special case of general radial moments F_{pq}

$$F_{pq} = \iint f(r,\theta) g_{pq}(r) e^{\hat{j}q\theta} r \mathrm{d}r \mathrm{d}\theta, \qquad \hat{j} = \sqrt{-1}, \quad (1)$$

where g_{pq} is a function of radial variable and p, q are integer indices. Under image rotation by α, the moment F_{pq} changes due to the Fourier shift theorem as follows:

$$F'_{pq} = F_{pq} e^{\hat{j}q\alpha}. \quad (2)$$

Hence, the moment magnitude $|F_{pq}|$ is a rotation invariant and is shown in detail in Ref. 26. For a continuous image function $f(x,y)$, the ZM of order p with repetition q is defined by the following equation:

$$Z_{pq} \frac{p+1}{\pi} \int_{\theta=0}^{2\pi} \int_{r=0}^{1} V^{*}_{pq}(r,\theta) f(r,\theta) r \mathrm{d}r \mathrm{d}\theta, \qquad |r| \leq 1, \quad (3)$$

where the symbol $*$ is the sign of complex conjugate and V_{pq} denotes the Zernike polynomial of order $p = 0, 1, 2, \ldots$ and repetition $q = -p, -p+2, \ldots, p$, which is defined as follows:

$$V_{pq}(r,\theta) = R_{pq}(r) e^{\hat{j}q\theta}. \quad (4)$$

The real-valued radial polynomial R_{pq} is expressed as follows:

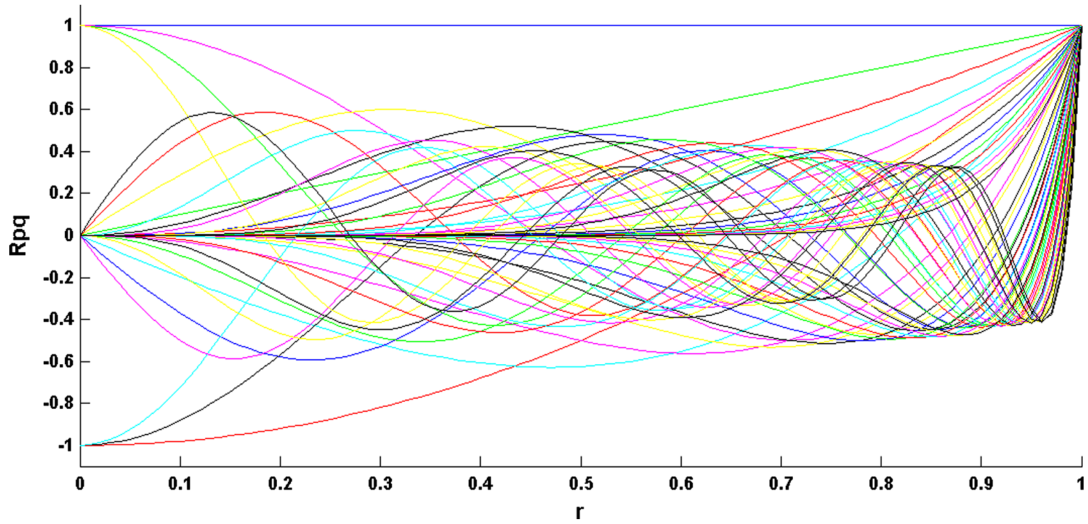

Fig. 1 Zernike polynomials up to the 13th degree.

$$R_{pq} = \sum_{k=0}^{\frac{p-|q|}{2}} (-1)^k \frac{(p-k)!}{k!\left(\frac{p+|q|}{2}-k\right)!\left(\frac{p-|q|}{2}-k\right)!} r^{p-2k}, \tag{5}$$

where $p \leq q$ and $p-|q|$ is even. The radial part of ZMs up to the 13th degree is plotted in Fig. 1.

2.2 *Discrete Approximation of Zernike Moments*

Since ZMs are defined over polar coordinates inside a circle, their calculation requires a transformation of the coordinates, which maps the image inside the unit disk. We use the transform that is given as follows:

$$x_i = \frac{\sqrt{2}}{N-1} i - \frac{1}{\sqrt{2}}, \quad y_j = \frac{\sqrt{2}}{N-1} j + \frac{1}{\sqrt{2}}, \tag{6}$$

$$r_{ij} = \sqrt{x_i^2 + y_j^2}, \qquad \theta_{ij} = \tan^{-1}\left(\frac{y_i}{x_i}\right). \tag{7}$$

As a result, the discrete approximation of continuous ZMs is defined by the following equation:

$$Z_{pq} = \lambda_Z(p, R, C) \sum_{i=0}^{R-1} \sum_{j=0}^{C-1} R_{pq}(r_{ij}) e^{-\hat{j}q\theta_{ij}} f(i,j), \tag{8}$$

where $N = \max(R, C)$ and

$$\lambda_Z(p, R, C) = \frac{2(p+1)}{\pi(R-1)(C-1)}. \tag{9}$$

3 Spectral Regression Discriminant Analysis

3.1 *Definition*

SRDA is an advanced data reduction technique introduced by Deng Cai as an efficient dimensionality reduction tool for large-scale discriminant analysis.[27,28] It is based on a combination of spectral graph analysis and regression, which decreases the time complexity of the system while enhancing its performance considerably. Theoretically, in this approach only a set of regularized least square problems is solved and no eigenvector computation is involved. As a result, both time and memory are saved significantly in comparison with other dimensionality reduction algorithms such as linear discriminant analysis (LDA). As shown in Table 1, SRDA decreases computational complexity from cubic time complexity to linear time complexity. Moreover it can cope with the "small sample size" problem, whereas LDA would fail to work in this case due to the singularity of the scatter matrix in LDA when the number of features is larger than the number of samples, a common issue in large-scale experiments.

3.2 *Algorithmic Procedure*

Suppose $x_1, \ldots, x_m \in R^n$ be a set of data points that belong to C classes and m_k be the number of samples in the k'th class ($\sum_{k=1}^{c} m_k = m$). The algorithmic procedure of SRDA can be summarized in three steps as follows:

3.2.1 *Response generation*

Suppose

$$y_k = [\underbrace{0,\ldots,0}_{\sum_{i=1}^{k-1} m_i}, \underbrace{1,\ldots,1}_{m_k}, \underbrace{0,\ldots,0}_{\sum_{i=k+1}^{c} m_i}]^T$$

and $y_0 = [1, 1, \ldots, 1]^T$ be a vector of ones. Now y_0 is taken as the first vector, and then the Gram-Schmidt process is

Table 1 Computational complexity of LDA and SRDA.[27]

Algorithm	Time complexity	Memory
LDA	$O(mnt + t^3)$	$O(mn + mt + nt)$
SRDA	$O(ms)$	$O(ms)$

Note: m, the number of samples; n, the number of features; t, $\min(m, n)$; s, the average number of nonzero features for one sample.

employed to orthogonize $\{y_k\}$. As a result $c-1$ vectors are derived as follows:

$$\{\bar{y}_k\}_{k=1}^{c-1}, \qquad (\bar{y}_i^T y_0 = 0, \bar{y}_i^T \bar{y}_j = 0, i \neq j). \tag{10}$$

3.2.2 *Regularized least squares*

In the second step, first a new element "1" is added to each x_i and is still denoted by x_i to avoid complication. Now $c-1$ vectors $\{a_k\}_{k=1}^{c-1}$, which are the basis vectors of SRDA, are calculated by the regularized least squares problem as follows:

$$a_k = \arg\min_a \left[\sum_{i=1}^{m} (a^T x_i - \bar{y}_i^k)^2 + \alpha \|\alpha\|^2\right], \tag{11}$$

where $\bar{y}_i^k$ denotes the i'th element of $\bar{y}^k$ and α is a regularization parameter that controls the smoothness of the estimator.

3.3 *Embedding to (c − 1) dimensional subspace*

As soon as $c-1$ vectors $\{a_k\}_{k=1}^{c-1}$ are derived, the samples are embedded into $(c-1)$-dimensional subspace as follows:

$$x \to z = A^T \begin{bmatrix} x \\ 1 \end{bmatrix}, \tag{12}$$

where $A = [a_1, \ldots, a_{c-1}]$ is a $(n+1) \times (c-1)$ transformation matrix.

4 Proposed System (ZMSRDA)

In this section, feature extraction, discrimination analysis, and classification are discussed in detail. The proposed face-recognition system is composed of different parts as shown in Fig. 2. Since all face images in our database are already detected, no face detection is performed in the proposed algorithm.

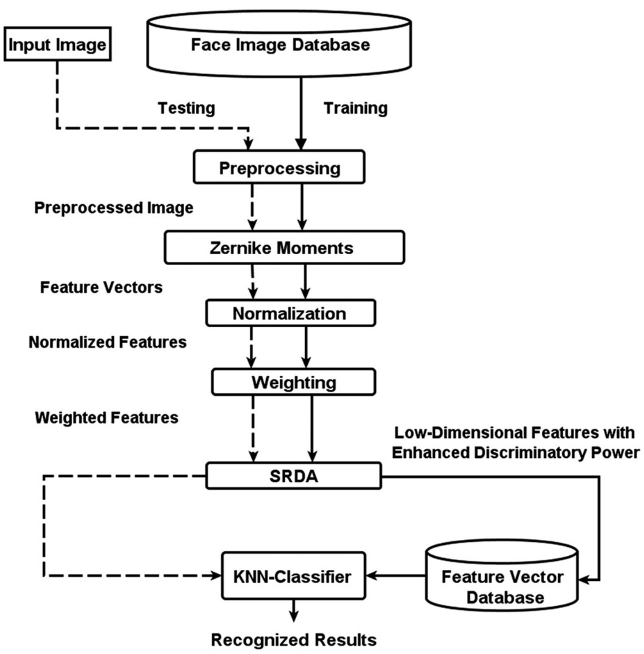

Fig. 2 The general procedure of the proposed face-recognition system.

4.1 *Feature Extraction*

Facial features in this approach are salient features derived by calculation of ZMs for different orders. Due to the orthogonal and global characteristics of ZMs, our system is robust to rotation and noise. In a nutshell the process of feature extraction can be summarized as follows:

- ZMs of an image are calculated. However, to improve the performance of the system, different orders are checked and the optimal is selected as the best order of ZMs (see Table 2).
- Since different features get a different range of possible values, the classification may be based primarily on the features with a wider range of values. As a result, in the second step the normalization process is performed with the features as follows:

$$\hat{a}_{i,j} = \frac{a_{i,j} - \mu_j}{\sigma_j}, \tag{13}$$

  where $a_{i,j}$ is the j'th feature of the i'th image and μ_j and σ_j are the mean and standard deviation of the features in the training set, respectively.
- In the third step, the normalized features are weighted by a number between 0 and 1, which are calculated by the correct classification rate of different orders of ZMs in the training set. As a result, the features of different orders with higher discrimination power obtain higher weights. This process not only enhances the discrimination power of the features but also improves the classification process of the system considerably.

4.2 *Discrimination Analysis*

As soon as weighted feature vectors are generated, discrimination analysis is applied to enhance the discrimination power of the system. For this purpose, SRDA is employed. The Tikhonov regularizer[29] is used to control the model complexity. Finally the optimum value of α is calculated and a model is generated (see Table 3).

4.3 *Classification*

To classify the probe image, $k-NN$ classifier $(k=1)$ with Euclidean distance criterion is used in the proposed algorithm due to its simplicity as well as speed. Finally, the recognition rate of the proposed system is calculated by the following equation and reported in our experiments:

$$\text{Recognition rate} = \frac{\text{Correctly classified samples}}{\text{Total number of samples}} \times 100. \tag{14}$$

5 Experimental Results and Performance Analysis

In this section, the performance of the proposed algorithm is tested and then compared with some existing NIR face-recognition schemes such as the Fisherface technique,[8] LBP + Fisherface (LBPF),[11] and wavelet transform +

2DPCA (W2DPCA).[15] In addition two more approaches based on ZM and ZM + Fisherface (ZMF) are evaluated to highlight the contribution of SRDA in our algorithm. In the first part of this section, the image database and the normalization process are described briefly. This is followed by the experiments carried out to evaluate the performance of the system. The following sets of experiments are carried out:

- Testing the performance of the proposed method with normal faces and defining the optimum order of ZMs and the optimum value of α in SRDA.
- Testing the performance of the system with different challenges such as facial expression and slight pose.
- Testing the rotation invariance of the algorithm.
- Testing the robustness of the system against salt-pepper noise.
- Measurement of computation time.

All of our experiments were conducted using MATLAB 2012a on a Pentium 4 2.50-GHz Windows 7 machine with 4 GBytes of memory.

5.1 *Database and Preprocessing*

In this paper, we use the CASIA NIR face database, which is composed of 3940 NIR face images of 197 people.[11] The images were taken with a home-brew device which produces images with 640×480 resolutions in bitmap format. Since some of the subjects do not exhibit special challenges, they were not considered, and only 100 subjects, which include variations in facial expressions and head pose, were selected. For every subject, there are 20 NIR images in the database. For our experiment, 10 images consisting of normal images, images with facial expressions, and images with head pose are selected for each subject. Hence the whole number of images in our experiment is 1000 images (10 images per subject). The tolerance for the head pose is maximum of 5 deg from the straight position; the images that include bigger head pose were not considered.

The normalization procedure consists of initial manual alignment of raw images based on eye coordinates of images. Consequently we crop and resize them to 64×64 to reduce computational complexity. Several samples of raw and normalized images are shown in Fig. 3.

5.2 *Experimental Results*

5.2.1 *Testing the performance of the proposed method for normal faces*

In this section, we assess the performance of the system for normal faces of 100 subjects to optimize the performance of different components in the proposed algorithm. The normal faces are frontal faces without facial expression and head pose. The number of images in training and testing set is three, and there is no overlap between training and testing set. The following sets of experiments are conducted in this section:

- Testing the performance of the system with different orders of ZMs and selecting the optimal to enhance the performance of the system.
- Testing the performance of the system with different values of α in SRDA.

Table 2 Percentage of classification for different orders.

Order	Dimensionality up to specified order	Mean recognition rate of ZM up to specified order (%)
0	2	1.4
1	8	19.1
2	18	58.8
3	32	82.3
4	50	85.2
5	72	83.8
6	98	88.2
7	128	94.1
8	162	95.6
9	200	95.6
10	**242**	**97.3**
11	288	96.4
12	338	95.6
13	392	96.4

Table 3 Effect of different values of α on the face-recognition rate.

α	Mean recognition rate (%) ZMSRDA
0.6	91.5
0.5	92.2
0.4	93.5
0.3	95.6
0.2	96.1
0.1	97.3
0.05	97.3
0.02	98.5
0.01	**100.0**
0.007	98.5
0.003	96.1

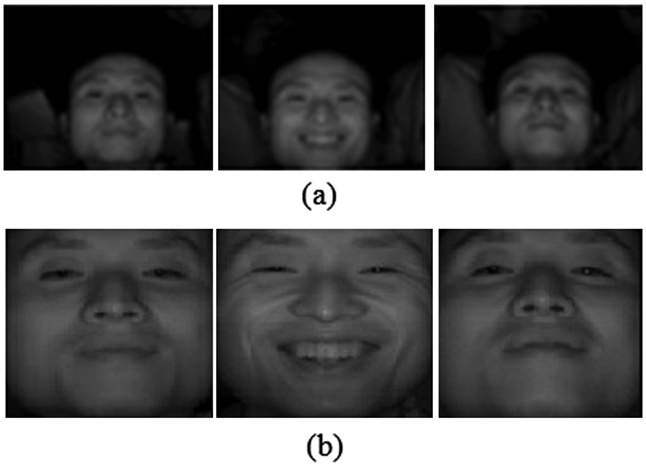

Fig. 3 (a) Samples of raw images; (b) samples of preprocessed images including normal image, image with facial expressions, and image with head pose, respectively.

Optimum order of Zernike moments. In the first experiment, different orders of ZMs are tested by empirical approach and the mean recognition rate is considered. Table 2 shows the mean recognition rate of ZMs up to a specified order. As can be seen the best recognition rate is obtained with order 10 and the recognition rate reduces when ZM order increases. As a result, order 10 is selected as an optimum order and the feature vectors of order 0 to 10 are used for our further experiments. The same results can be found in Refs. 20 to 22 and 24.

Parameter selection for SRDA. Regularization parameter α in Eq. (11) is an essential parameter that controls the smoothness of the estimator in the SRDA algorithm. In this experiment the impact of α on the performance of our algorithm is evaluated. Hence different values of α are tested empirically and the results are considered. As shown in Table 3, the best recognition rate is achieved with $\alpha = 0.01$, and the recognition rate is reduced when the value of this parameter is reduced. As a result $\alpha = 0.01$ is chosen as an optimum value of α in SRDA for our further experiments.

5.2.2 *Testing the performance of the system with different variations*

In this experiment, the performance of the algorithm for face images with different challenges including facial expressions and head pose is evaluated. A random subset with $l(= 2, 3, 4)$ images per person is taken with labels to form the training set, and the rest is considered as a testing set. One hundred subjects are used in this experiment. We employ a closed universe assumption (i.e., each test image will have a corresponding match in the training). For each l the average recognition rate with over 20 splits is calculated, and the result is considered. Finally a comparison study is performed to highlight the superiority of our method. W2DPCA, Fisherface, LBPF, ZM, and ZMF are used for comparison. The average recognition rates along with standard deviations are presented in Table 4. Standard deviation is the most useful criterion for evaluating whether different selections of training images affect the recognition performance. The lower the standard deviation, the less the effect of choosing different training image sets.[30] Recognition rates versus different number of training samples are plotted in Fig. 4. The following results can be concluded from Table 4 and Fig. 4:

Table 4 Performance comparison of various methods on CASIA database (mean ± SD percent)

Method	Number of images in the training set		
	2	3	4
W2DPCA[15]	86.4 ± 3.11	88.1 ± 2.27	92.4 ± 2.01
Fisherface[8]	90.3 ± 2.79	94.6 ± 2.01	97.2 ± 1.11
LBPF[11]	96.1 ± 1.46	97.9 ± 0.91	98.7 ± 0.74
ZM	89.8 ± 3.21	92.9 ± 2.65	95.3 ± 1.25
ZMF	93.8 ± 2.14	95.4 ± 1.82	97.8 ± 1.02
ZMSRDA	98.3 ± 1.29	99.1 ± 0.69	99.5 ± 0.25

- As shown in Table 4, ZMSRDA (the proposed algorithm) yields the maximum recognition rate with the minimum standard deviation, which shows a high discrimination power of extracting features that are as representative as possible. Moreover, ZMF and ZMSRDA performed better than ZM because of applying discriminant analysis, which significantly increases the discrimination power of the system. Of course, the performance of ZMSRDA is better than ZMF. It shows the superiority of regularized techniques such as SRDA in comparison with the classical Fisherface.
- From the graph below, we can see that the performance of the Fisherface drops sharply with the decreasing number of training samples for each person. It shows high dependency of the performance of the Fisherface technique on the size and representativeness of the training set. The underlying reason is that Fisherface perceives only the global Euclidean structure of image data. Hence it does not have sufficient capability to discover the intrinsic manifold of data when the number of images in the training set is low. This result is consistent with the result in Ref. 31, which shows the deficiency of the Fisherface method when the training set is small.
- The stability of our method with different number of images in the training set is considerable. It can be implied by the fluctuation of Fig. 4 based on ZMSRDA, which is more consistent than other methods. As a result, our method is irrelevant with respect to the number of images in the training set, which can be considered as a big advantage of our method in comparison with other methods.

5.2.3 *Testing the rotation invariance of the algorithm*

Rotation of a head approximately around the optical axis of the camera leads to an in-plane rotation of the face image. This unwanted variation must be overcome either by

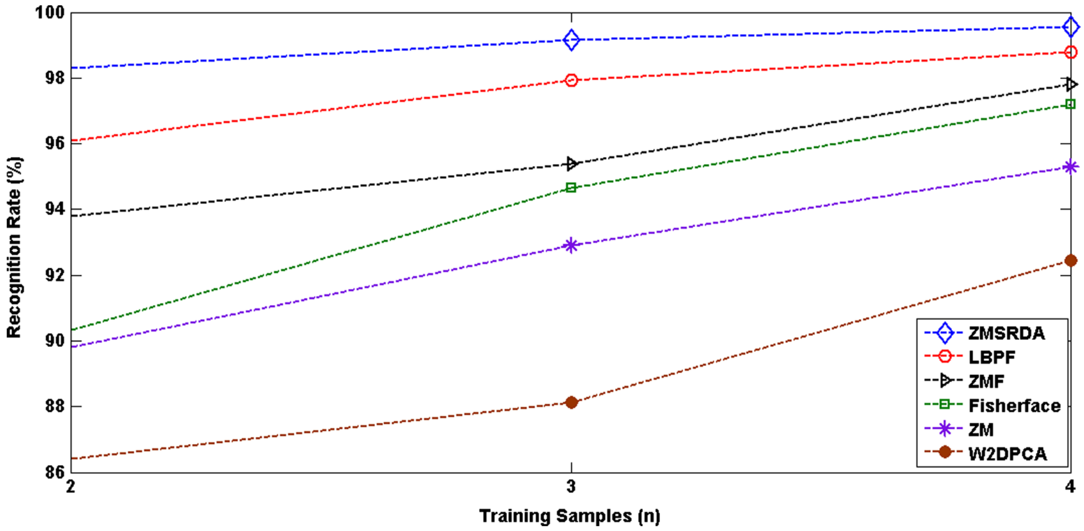

Fig. 4 Recognition rate versus different number of training samples.

transforming the face into a normalized position (using principal axes for instance), which is time-consuming, or by rotational invariance of the features used for recognition. The importance of the rotation invariance property in face recognition was highlighted in Refs. 20, 30, and 32.

In this section, the rotation invariance of our algorithm for 100 subjects is tested. In most implementations some part of the images are located outside the boundary of the images. Hence raw images are masked to remove the outside locations of images when they are rotated. Five images are used in training and the rest are rotated up to 150 deg to form a testing set. Some samples of rotated images are shown in Fig. 5. Rotation Zernike invariants up to order 10 are used in this experiment. Table 5 illustrates the mean recognition rate of the proposed algorithm for different angles. As can be seen in this table, features extracted using rotation Zernike invariants are invariant to rotation even with high angles, which shows the invariance property of ZMs. In other words, due to rotation, each ZM acquires a phase shift. Hence the magnitude of a rotated image remains identical to those before rotation.

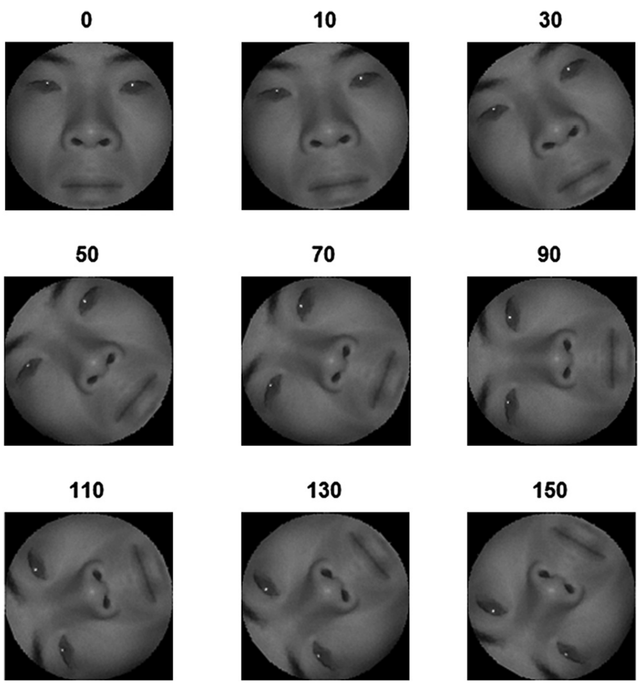

Fig. 5 Face images with different rotations in degrees.

5.2.4 *Testing the performance of the system against salt-pepper noise*

Image noise is a random, usually unwanted, variation in brightness or color information in an image.[33] Salt-pepper noise is one of the common types of noises that can be caused by dead pixels, analog to digital convertors, bit errors in transformation, etc. It has been given different appellation including "impulsive noise" and "spike noise." Actually an image containing salt-pepper noise will have dark pixels in bright regions and bright pixels in dark regions.[20] Some samples of facial images with different levels of salt-pepper noise are shown Fig. 6. In this section, experiments on CASIA

Table 5 Face-recognition results for different rotation for ZMSRDA.

Rotation (in deg)	Mean recognition rate (%)
0	99.1
10	96.1
30	94.8
50	96.1
70	93.7
90	98.6
110	94.4
130	96.1
150	93.5

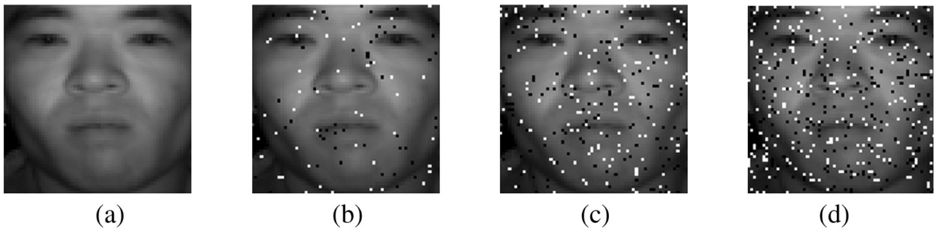

Fig. 6 (a) Typical original image from CASIA database; (b), (c), and (d) are the face images with noise density of 0.03, 0.06, and 0.1, respectively.

database for 100 subjects are performed to test the robustness of the proposed method to salt-pepper noise. Three images per subject are randomly selected for training, and the remaining seven images per subject are used for testing. While the training set includes image without noise, the testing set includes noisy images with different levels of noise densities. Ten realizations of noisy images are generated and their accuracies are averaged to get the "representative" value. The mean recognition rate of the algorithms based on salt-pepper noise with different densities are shown in Table 6 and plotted in Fig. 7, respectively. The following results can be concluded:

Table 6 Effect of salt-pepper noise on the performance of different face-recognition methods.

	Noise density			
Algorithms	**0**	**0.03**	**0.06**	**0.1**
W2DPCA	88.1	86.5	85.3	80.4
Fisherface	94.6	93.7	91.3	87.7
LBPF	97.9	95.1	88.9	72.3
ZM	92.9	92.2	91.2	90.3
ZMF	95.4	95.1	93.9	92.5
ZMSRDA	99.1	99.0	98.7	98.2

- The sensitivity of wavelet transform to heavy noise can be seen in this experiment. This is due to the down-sampling process which decreases the resolution of the images as well as the discrimination power of the system in presence of noise.
- The accuracy of the Fisherface approach in the presence of heavy noise decreases highly. It shows the sensitivity of appearance-based methods in the presence of heavy noise. A possible explanation for this is that the performance of appearance-based methods largely relies on representation of the training samples. Hence the recognition rates of face-recognition methods based on appearance-based methods are degraded sharply when the noisy images are used in the testing set. Our result bolsters the results presented in Refs. 32 and 34.
- In Fig. 7 there is a clear trend of decreasing recognition rate of LBPF, which shows high sensitivity and deficiency of LBPF to noise. The underlying reason is that LBP thresholds exactly at the value of the central pixel. Hence original LBP tends to be sensitive to noise. Our results prove the high sensitivity of LBP to noise which was already mentioned in Ref. 14.
- We can see in Fig. 7 that the method based on ZM invariants is more robust to noise than other methods, particularly than LBPF, which is the most vulnerable.

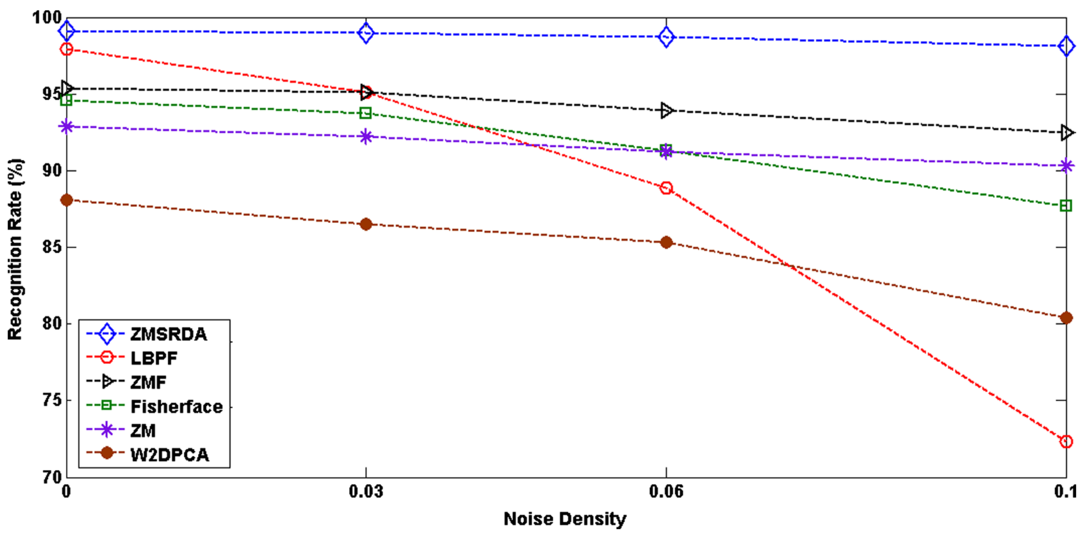

Fig. 7 Performance of different face recognition on noisy images.

Table 7 Comparison of running time for different algorithms.

Algorithms	Running time (400 images) Training time (s)	Running time (100 images) Testing time (s)	Mean recognition rate (%)
W2DPCA	7.59	3.82	88.1
Fisherface	5.91	0.96	95.6
LBPF	67.22	9.42	97.9
ZM	13.85	3.82	94.1
ZMF	16.17	4.93	95.4
ZMSRDA	13.93	3.86	99.4

This result is consistent with the result reported in Ref. 20, which proves high robustness of ZMs to noise.

5.2.5 *Measurement of computation time*

In this section the average time of feature extraction for each algorithm for training the 400 images and testing 100 images are calculated and shown in Table 7. As can be seen in Table 7, there is no significant difference between running time for testing of ZMSRDA and ZM, while the difference between recognition rates is very large. This observation proves the high performance of SRDA, yielding the proposed method fast and accurate. Moreover, the superiority of SRDA in comparison with Fisherface is more sensible. As can be seen in Table 7, ZMF takes somewhat more time for training and testing compared with ZMSRDA, which represents the effectiveness of SRDA in the proposed algorithm. Further analysis shows that the time required for training and testing of LBPF is very high compared with other methods. This result may be explained by two different factors. First, the process of generating LBP histogram features requires a lot of time and memory. Second, using PCA on high-dimensional LBP features as a preprocessing step increases time and memory costs. Thus it can be concluded that SRDA not only increases the discrimination power of our system but also accelerates the feature extraction in comparison with the Fisherface technique.

6 Discussions and Conclusion

In this paper we have presented a novel framework based on a combination of features extracted using ZMs and SRDA to compensate facial expression, slight pose, in-plane rotation, and noise for facial images. The method first used ZM invariants as a feature extractor and then normalized and weighted features to increase the accuracy of the system. Finally SRDA was implemented to reduce the dimension of features and to increase the discrimination power of the system. The results were low-dimensional salient features with high discrimination power. The paper has also compared the performance of the proposed method with other NIR algorithms in the presence of noise and image variations. The CASIA NIR database was used to test the feasibility of the algorithm. It was shown that ZMSRDA with small-size training samples performs better than the other algorithms such as W2DPCA, Fisherface, and LBPF. Moreover, the tests on noisy and rotated images proved the noise immunity and rotation invariance of the proposed approach. This study has found that LBP, which is one of the most popular feature extractors, has a deficiency in the presence of noise. Further analysis showed that our proposed method has a low time complexity as compared to LBP which has a high time complexity. Hence it can be concluded that our method offers a better potential for implementation in close-to-real-time or even real-time face-recognition systems.

The study also enhanced our understanding of the importance of ZMs in pattern recognition. Moreover it makes several noteworthy contributions to propose rotation and noise invariant face recognition in the NIR domain. The most important limitation lies in the fact that due ZMs are global features, any local change of the image affects all ZMs. Hence, the proposed method cannot properly handle the eyeglass versus no eyeglasses scenario (this drawback is, however, common in most face-recognition systems). The second important limitation of this study is related to interfacing with legacy systems that may only have visible imagery and may need matching visible imagery to NIR imagery. Since the raw intensities of both modalities are significantly different and ZMs are not designed for this purpose, such system should use some multimodal similarity measure such as mutual information instead.

We are currently exploring this problem in theory and practice and future works will focus on enhancing solutions to overcome these limitations.

Acknowledgments

The authors would like to thank Universiti Teknologi Malaysia for the support in Research and Development, and Soft Computing Research Group for the inspiration in making this study a success and Institute of Information Theory and Automation for providing MATLAB codes. This work is supported by The Ministry of Higher Education under Long Term Research Grant Scheme (LRGS/TD/2011/UTM/ICT/03-4L805) and by the Czech Science Foundation under Grant No. P103/11/1552.

References

1. W. Zhao et al., "Face recognition: a literature survey," *ACM Compt. Surveys* **35**(4), 399–458 (2003).
2. A. K. Jain, A. Ross, and S. Prabhakar, "An introduction to biometric recognition," *IEEE Trans. Circuits Systems for Video Techn.* **14**(1), 4–20 (2004).
3. S. Moritz, A. H. Jorgen, and G. E. Erik, "Physics-based modelling of human skin colour under mixed illuminants," *Robot. Autonom. Syst.* **35**(3), 131–142 (2001).
4. Y. Adini, Y. Moses, and S. Ullman, "Face recognition: the problem of compensating for changes in illumination direction," *IEEE Trans. Pattern Anal. Mach. Intell.* **19**(7), 721–732 (1997).
5. D. A. Socolinsky, A. Selinger, and J. D. Neuheisel, "Face recognition with visible and thermal infrared imagery," *Comput. Vis. Image Understand.* **91**(1), 72–114 (2003).
6. X. Zou, J. Kittler, and K. Messer, "Illumination invariant face recognition: a survey," in *Proc. First IEEE Int. Conf. on Biometrics: Theory, Applications, and Systems*, pp. 1–8, IEEE, Washington DC (2007).
7. S. Z. Li et al., "Highly accurate and fast face recognition using near infrared images," *in Advances in Biometrics*, D. Zhang and A. Jain, Eds., pp. 151–158, Springer, Hong Kong (2005).

8. X. Zou, J. Kittler, and K. Messer, "Face recognition using active near-IR illumination," in *Proc. British Machine Vision Conf*, pp. 1–11, Citeseer, Oxford, UK (2005).
9. P. N. Belhumeur, J. P. Hespanha, and D. J. Kriegman, "Eigenfaces vs. Fisherfaces: recognition using class specific linear projection," *IEEE Trans. Pattern Anal. Mach. Intell.* **19**(7), 711–720 (1997).
10. X. Tan et al., "Face recognition from a single image per person: a survey," *Pattern Recogn.* **39**(9), 1725–1745 (2006).
11. S. Z. Li et al., "Illumination invariant face recognition using near-infrared images," *IEEE Trans. Pattern Anal. Mach. Intell.* **29**(4), 627–639 (2007).
12. T. Ahonen et al., "Face description with local binary patterns: application to face recognition," *IEEE Trans. Pattern Anal. Mach. Intell.* **28**(12), 2037–2041 (2006).
13. D. Yi et al., "Outdoor face recognition using enhanced near infrared imaging," in *Proc. Advances in Biometrics*, pp. 1–9, Springer, Seoul, Korea (2007).
14. L. Nanni, A. Lumini, and S. Brahnam, "Local binary patterns variants as texture descriptors for medical image analysis," *Artif. Intell. Med.* **49**(2), 117–125 (2010).
15. Y. He et al., "Near infrared face recognition based on wavelet transform and 2DPCA," in *Proc. 2010 Int. Conf. on Intelligent Computing and Integrated Systems*, pp. 359–362, IEEE, Guilin, China (2010).
16. D.-Q. Dai, H. Yan, and S. Yat-Sen, *Wavelets and Face Recognition*, I-Tech, Austria (2007).
17. S. Mallat, "Wavelets for a vision," *Proc. IEEE* **84**(4), 604–614 (1996).
18. M. R. Teague, "Image analysis via the general theory of moments," *Opt. Soc. Am.* **70**(8), 920–930 (1980).
19. J. Flusser, T. Suk, and B. Zitova, *Moments and Moment Invariants in Pattern Recognition*, Wiley, Chichester (2009).
20. S. M. Lajevardi and Z. M. Hussain, "Higher order orthogonal moments for invariant facial expression recognition," *Digit. Signal Process.* **20**(6), 1771–1779 (2010).
21. A. Saradha and S. Annadurai, "A hybrid feature extraction approach for face recognition systems," *Int. J. Graph. Image Process.* **5**(5), 23–30 (2005).
22. S. M. Lajevardi and Z. M. Hussain, "Zernike moments for facial expression recognition," in *Proc. Int. Conf. on Communication, Computer and Power (ICCCP'09)*, pp. 378–381, ICCCP, Muscat, Oman (2009).
23. R. Zhi and Q. Ruan, "A comparative study on region-based moments for facial expression recognition," in *Proc. Congress on Image and Signal Processing*, pp. 600–604, IEEE, Sanya, Hainan, China (2008).
24. N. H. Foon et al., "An efficient method for human face recognition using wavelet transform and Zernike moments," in *Proc. Int. Conf. on Computer Graphics, Imaging and Visualization*, pp. 65–69, IEEE, Penang, Malaysia (2004).
25. C. H. Teh and R. T. Chin, "On image analysis by the method of moments," *IEEE Trans. Pattern Anal. Mach. Intell.* **10**(4), 496–513 (1998).
26. D. Shen and H. H. S. Ip, "Discriminative wavelet shape descriptors for recognition of 2-D patterns," *Pattern Recogn.* **32**(2), 151–165 (1999).
27. D. Cai, X. He, and J. Han, "SRDA: an efficient algorithm for large-scale discriminant analysis," *IEEE Trans. Knowledge Data Engin.* **20**(1), 1–12 (2008).
28. D. Cai, X. He, and J. Han, "Speed up kernel discriminant analysis," *VLDB* **20**(1), 1–13 (2011).
29. M. L. Whitney, "Theoretical and numerical study of Tikhonov's regularization and Morozov's discrepancy principle," MS Thesis, Georgia State University (2009).
30. D. V. Jadhav and R. S. Holambe, "Rotation, illumination invariant polynomial kernel Fisher discriminant analysis using Radon and discrete cosine transforms based features for face recognition," *Pattern Recogn. Lett.* **31**(9), 1002–1009 (2010).
31. D. Cai et al., "Orthogonal Laplacian faces for face recognition," *IEEE Trans. Image Process.* **15**(11), 3608–3614 (2006).
32. D. V. Jadhav and R. S. Holambe, "Radon and discrete cosine transforms based feature extraction and dimensionality reduction approach for face recognition," *Signal Process.* **88**(10), 2604–2609 (2008).
33. R. C. Gonzalez and R. E. Woods, *Digital Image Processing*, Addison-Wesley, USA (2008).
34. D. V. Jadhav and R. S. Holambe, "Feature extraction using Radon and wavelet transforms with application to face recognition," *Neurocomput* **72**(7), 1951–1959 (2009).

Sajad Farokhi completed his master's degree in pure mathematics in Iran. He is currently working toward a PhD in the Department of Computer Science, Universiti Teknologi Malaysia. His research areas include computer vision, machine learning, pattern recognition, biometrics, and image processing.

Siti Mariyam Shamsuddin received a BSc and MSc in mathematics from New Jersey and a PhD in pattern recognition and artificial intelligence from Universiti Putra Malaysia, Malaysia. Currently, she is head of Soft Computing Research Group and a full professor of computer science at Universiti Teknologi Malaysia, Johor, Malaysia. Her research interests include soft computing and its applications, pattern recognition, and geometric modeling.

Jan Flusser received an MSc in mathematical engineering from the Faculty of Nuclear Science and Physical Engineering, Czech Technical University, Prague, in 1985, a PhD in computer science from the Czechoslovak Academy of Sciences in 1990, and the DrSc degree in technical cybernetics from the Czech Technical University in 2001. He is the director of the Institute of Information Theory and Automation of the ASCR, Prague, Czech Republic, and a full professor of computer science at the Czech Technical University, Prague, and at the Charles University, Prague. He has authored and coauthored more than 200 research publications in these areas, including the book *Moments and Moment Invariants in Pattern Recognition* (Wiley, 2009), 50 papers in prestigious journals, and 20 tutorials at international conferences (ICIP'05, ICIP'07, EUSIPCO'07, CVPR'08, FUSION'08, SPPRA'09, SCIA'09, ICIP'09, SPPRA'10, and CGIM'10) and invited/keynote talks (ICCS'06, COMPSTAT'06, WIO'06, DICTA'07, CGIM'08). His publications have received about 3000 citations. His main areas are moments and moment invariants, image registration, multichannel blind deconvolution, and super-resolution imaging.

Usman Ullah Sheikh received his BEng (2003) in electrical and mechatronics engineering, MEng (2005) in telecommunications engineering, and PhD degree (2009) in image processing and computer vision from Universiti Teknologi Malaysia. His research work is mainly on computer vision and embedded systems design. He is currently a senior lecturer at Universiti Teknologi Malaysia, working on image processing for intelligent surveillance systems. He is a member of Institution of engineering and technology and IEEE and has published works in many IEEE conferences.

Mohammad Khansari received his BS, MS, and PhD degrees in computer engineering from the Sharif University of Technology, Tehran, Iran, in 1996, 1998, and 2008, respectively. He is one of the contributors of the Advanced Information and Communication Technology research Center (AICTC) of Sharif University, and the director of Iran Free/Open Source (FOSS) national project (formerly Persian Linux) in AICTC for five years of running the project. He had a short-time research fellowship from German Academic Exchange Service, Germany. He has given more than 50 invited talks on FOSS and Linux localization topics in Iran and international conferences such as United Nation Expert meeting 2004 (Geneva) and FOSS 2005 (Syria), United Nation World Information Society Summit 2005 (Tunisia) and FLOSSWorld 2005 (India), LinuxAsia 2007 and International Open Source Network summit (hosted by UNDP in India), and LinuxTag 2007 (Germany). He is the coauthor of four books with Professor Rabiee in free/open source software topics and has more than 40 papers in international conferences and journals. His main research interests are network science and complex networks, multimedia and healthcare sensor networks, multimedia delivery over peer-to-peer networks, free/open source software specially localization, customization, mining of software repositories and diffusion. He was the faculty member and assistant professor at School of Science and Engineering, Sharif International campus located in Kish Island and then head of Information Technology faculty and faculty member of Iran Telecommunication Research Center. Currently, he is the faculty member of Faculty of New Sciences and Technologies, University of Tehran, and the director of IT and Cyberspace Center of University of Tehran.

Kourosh Jafari-Khouzani received his PhD in computer science from Wayne State University in 2006. He is a research fellow in Athinoula A. Martinos Center for Biomedical Imaging at the Department of Radiology, Massachusetts General Hospital, Boston, Massachusetts. His main area of research interest is medical image analysis and neuroimaging.