Multimodal Biometric Person Authentication Using Fingerprint, Face Features

Tran Binh Long[1], Le Hoang Thai[2], and Tran Hanh[1]

[1] Department of Computer Science, University of Lac Hong 10 Huynh Van Nghe, DongNai 71000, Viet Nam
tblong@lhu.edu.vn
[2] Department of Computer Science, Ho Chi Minh City University of Science 227 Nguyen Van Cu, HoChiMinh 70000, Viet Nam
lhthai@fit.hcmus.edu.vn

Abstract. In this paper, the authors present a multimodal biometric system using face and fingerprint features with the incorporation of Zernike Moment (ZM) and Radial Basis Function (RBF) Neural Network for personal authentication. It has been proven that face authentication is fast but not reliable while fingerprint authentication is reliable but inefficient in database retrieval. With regard to this fact, our proposed system has been developed in such a way that it can overcome the limitations of those uni-modal biometric systems and can tolerate local variations in the face or fingerprint image of an individual. The experimental results demonstrate that our proposed method can assure a higher level of forge resistance in comparison to that of the systems with single biometric traits.

Keywords: Biometrics, Personal Authentication, Fingerprint, Face, Zernike Moment, Radial Basis Function.

1 Introduction

Biometrics refers to automatic identification of a person based on his physiological or behavioral characteristics [1],[2]. Thus, it is inherently more reliable and more capable in differentiating between an authorized person and a fraudulent imposter [3]. Biometric-based personal authentication systems have gained intensive research interest for the fact that compared to the traditional systems using passwords, pin numbers, key cards and smart cards [4] they are considered more secure and convenient since they can't be borrowed, stolen or even forgotten. Currently, there are different biometric techniques that are either widely-used or under development, including face, facial thermo-grams, fingerprint, hand geometry, hand vein, iris, retinal pattern, signature, and voice-print (figure.1) [3], [5]. Each of these biometric techniques has its own advantages and disadvantages and hence is admissible, depending on the application domain. However, a proper biometric system to be used in a particular application should possess the following distinguishing traits: uniqueness, stability, collectability, performance, acceptability and forge resistance [6].

P. Anthony, M. Ishizuka, and D. Lukose (Eds.): PRICAI 2012, LNAI 7458, pp. 613–624, 2012.

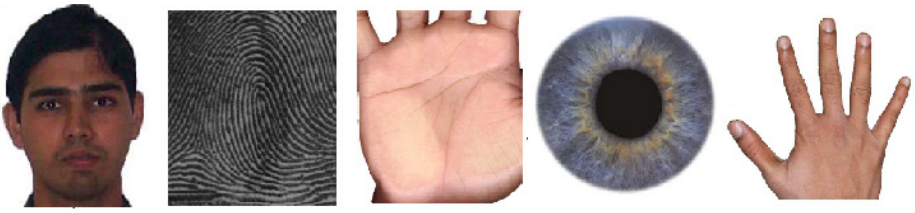

Fig. 1. Examples of biometric characteristic

Most biometric systems that are currently in use employ a single biometric trait; such systems are called uni-biometric systems. Despite their considerable advances in recent years, there are still challenges that negatively influence their resulting performance, such as noisy data, restricted degree of freedom, intra-class variability, non-universality, spoof attack and unacceptable error rates. Some of these restrictions can be lifted by multi-biometric systems [7] which utilize more than one physiological or behavioral characteristic for enrollment and verification/ identification, such as different sensors, multiple samples of the same biometrics, different feature representations, or multi-modalities. These systems can remove some of the drawbacks of the uni-biometric systems by grouping the multiple sources of information [8]. In this paper, multi-modalities are focused.

Multimodal biometrics systems are gaining acceptance among designers and practitioners due to (i) their performance superiority over uni-modal systems, and (ii) the admissible and satisfactory improvement of their system speed. Hence, it is hypothesized that our employment of multiple modalities (face and fingerprint) can conquer the limitations of the single modality- based techniques. Under some hypotheses, the combination scheme has proven to be superior in terms of accuracy; nevertheless, practically some precautions need to be taken as Ross and Jain [7] put that Multimodal Biometrics has various levels of fusion, namely sensor level, feature level, matching score level and decision level, among which [8] fusion at the feature level is usually difficult. One of the reasons for it is that different biometrics, especially in the multi-modality case, would have different feature representations and different similarity measures. In this paper, we proposed a method using face and fingerprint traits with feature level fusion. Our work aims at investigating how to combine the features extracted from different modalities, and constructing templates from the combined features. To achieve these aims, Zernike Moment (ZM)[9] was used to extract both face and fingerprint features. First, the basis functions of Zernike moment (ZM) were defined on a unit circle. Namely, the moment was computed in a circular domain. This moment is widely-used because its magnitudes are invariant to image rotation, scaling and noise, thus making the feature level fusion of face and fingerprints possible. Then, the authentication was carried out by Radial Basis Function (RBF) network, based on the fused features.

The remainder of the paper is organized as follows: section 2 describes the methodology; section 3 reports and discusses the experimental results, and section 4 presents the conclusion.

2 Methodology

Our face and fingerprint authentication system is composed of two phases which are enrollment and verification. Both phases consist of preprocessing for face and fingerprint images, extracting the feature vectors invariant with ZMI, fusing at feature level, and classifying with RBF. (Figure 2)

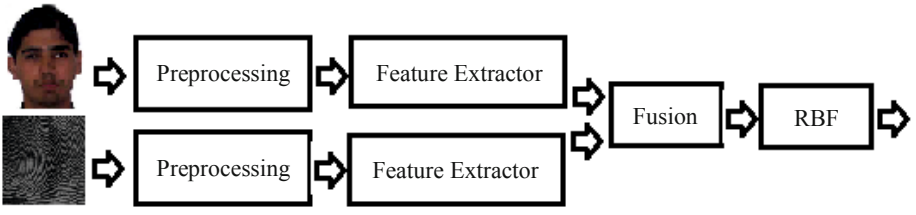

Fig. 2. The chart for face and fingerprint authentication system

2.1 Preprocessing

The purpose of the pre-processing module is to reduce or to eliminate some of the image variations for illumination. In this stage, the image had been preprocessed before the feature extraction. Our multimodal authentication system used histogram equalization, wavelet transform [10] to preprocess the image normalization, noise elimination, illumination normalization etc., and different features were extracted from the derived image normalization (feature domain) in parallel structure with the use of Zernike Moment (ZM).

Wavelet transform [10] is a representation of a signal in terms of a set of basic functions, obtained by dilation and translation of a basis wavelet. Since wavelets are short-time oscillatory functions with finite support length (limited duration both in time and frequency), they are localized in both time (spatial) and frequency domains. The joint spatial-frequency resolution obtained by wavelet transform makes it a good candidate for the extraction of details as well as approximations of images. In the two-band multi-resolution wavelet transform, signals can be expressed by wavelet and scaling basis functions at different scale, in a hierarchical manner. (Figure.3)

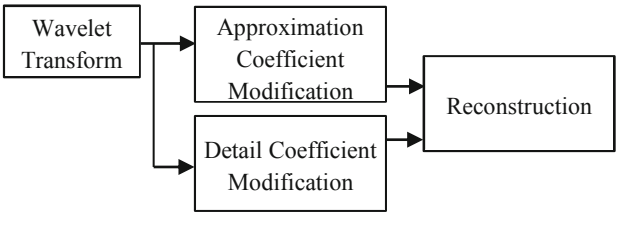

Fig. 3. Block diagram of normalization

$$f(x) = \sum_k a_{0,k}\emptyset_{0,k}(x) + \sum_j \sum_k d_{j,k}\psi_{j,k}(x) \quad (1)$$

$\emptyset_{j,k}$ are scaling functions at scale j and $\psi_{j,k}$ are wavelet functions at scale j. $a_{j,k}$, $d_{j,k}$ are scaling coefficients and wavelet coefficients.

After the application of wavelet transform, the derived image was decomposed into several frequency components in multi-resolution. Using different wavelet filter sets and/or different number of transform-levels will bring about different decomposition results. Since selecting wavelets is not the focus of this paper, we randomly chose 1-level db10 wavelets in our experiments. However, any wavelet-filters can be used in our proposed method.

2.2 Feature Extraction with Zernike Moment

The purpose of feature extraction is to extract the feature vectors or information which represents the image. To do it, Zernike Moment (ZM) was used. Zernike moment (ZM) used for face and fingerprint recognition in our work is based on the global information. This approach, also known as statistical method [11], or moment- and model- based approach [12][13], extracts the relevant information in an image. In order to design a good face and fingerprint authentication system, the choice of feature extractor is very crucial. The chosen feature vectors should contain the most pertinent information about the face and the fingerprint to be recognized. In our system, different feature domains were extracted from the derived images in parallel structure. In this way, more characteristics of face and fingerprint images for authentication were obtained. Among them, two different feature domains- ZM for Face and ZM for fingerprint - were selected.

Given a 2D image function f(x, y), it can be transformed from Cartesian coordinate to polar coordinate f(r, θ), where r and θ denote radius and azimuth respectively. The following formulae transform from Cartesian coordinate to polar coordinate,

$$r = \sqrt{x^2 + y^2}, \tag{2}$$

and

$$\theta = \arctan\left(\frac{y}{x}\right) \tag{3}$$

Image is defined on the unit circle that $r \leq 1$, and can be expanded with respect to the basic functions $V_{nm}(r, \theta)$.

For an image $f(x, y)$, it is first transformed into the polar coordinates and denoted by $f(r, \theta)$. The Zernike moment with order n and repetition m is defined as

$$M_{nm} = \frac{n+1}{\pi} \int_0^{2\pi} \int_0^1 [V_{nm}(r, \theta)]^* f(r, \theta) r dr d\theta \tag{4}$$

Where * denotes complex conjugate, n = 0, 1, 2. . . ∞, m is an integer subject to the constraint that n - |m| is nonnegative and even. $V_{nm}(r, \theta)$is the Zernike polynomial, and it is defined over the unit disk as follows

$$V_{nm}(r, \theta) = R_{nm}(r) e^{im\theta} \tag{5}$$

With the radial polynomial $R_{nm}(r)$ defined as

$$R_{nm}(r) = \sum_{s=0}^{\frac{(n-|m|)}{2}} \frac{(-1)^s (n-s)! r^{n-2s}}{s!\left(\frac{n+|m|}{2}-s\right)!\left(\frac{n-|m|}{2}-s\right)!} \tag{6}$$

The kernels of ZMs are orthogonal so that any image can be represented in terms of the complex ZMs. Given all ZMs of an image, it can be reconstructed as follows.

$$f(r,\theta) = \sum_n \sum_{(All\ m's)} M_{nm} V_{nm}(r,\theta) \tag{7}$$

The defined features of Zernike moments themselves are only invariant to rotation. To achieve scale and translation invariance, the image needs to be normalized first by using the regular Zernike moments.

The translation invariance is achieved by translating the original image $f(x,y)\ to\ f(x+\bar{x}, y+\bar{y})$, where $\bar{x} = m_{10}/m_{00}\ and\ \bar{y} = m_{01}/m_{00}$.

In other words, the original image's center is moved to the centroid before the Zernike moment's calculation. Scale invariance is achieved by enlarging or reducing each shape so that the image's 0th regular moment m'_{00} equals to a predetermined value β. For a binary image, m_{00} equals to the total number of shape pixels in the image, for a scaled image f(α x, α y), its regular moments $m'_{pq} = \alpha^{p+q+2} m_{pq}$, m_{pq} is the regular moments of f(x,y).

Since the objective is to make $m'_{00} = \beta$, we can let $= \sqrt{\beta/m_{00}}$. By substituting $= \sqrt{\beta/m_{00}}\ \ into\ m'_{00}$, we can obtain $m'_{00} = \alpha^2\ m_{00} = \beta$.

The fundamental feature of the Zernike moments is their rotational invariance. If f(x,y) is rotated by an angle α, then we can obtain that the Zernike moment Z_{nm} of the rotated image is given by

$$Z'_{nm} = Z_{nm} e(-jm\alpha) \tag{8}$$

Thus, the magnitudes of the Zernike moments can be used as rotationally invariant image features.

2.3 ZM-Based Features

It is known from the experiments that ZM performs better than other moments (e.g. Tchebichef moment [14], Krawtchouk moment [15]) do. In practice, when the orders of ZM exceed a certain value, the quality of the reconstructed image degrades quickly

Fig. 4. Example of ZM for feature extraction with face and fingerprint

Table 1. The first 10 order Zernike moments

Order	Dimensionality	Zernike moments
0	1	M_{00}
1	2	M_{11}
2	4	M_{20}, M_{22}
3	6	M_{31}, M_{33}
4	9	M_{40}, M_{42}, M_{44}
5	12	M_{51}, M_{53}, M_{55}
6	16	M_{60}, M_{62}, M_{64}, M_{66}
7	20	M_{71}, M_{73}, M_{75}, M_{77}
8	25	M_{80}, M_{82}, M_{84}, M_{86}, M_{88}
9	30	M_{91}, M_{93}, M_{95}, M_{97}, M_{99}
10	36	$M_{10\,0}$, $M_{10\,2}$, $M_{10\,4}$, $M_{10\,6}$, $M_{10\,8}$, $M_{10\,10}$

because of the numerical instability problem inherent with ZM. From the noted problem, we decided to choose the first 10 orders of ZM with 36 feature vector elements. In this way, ZM can perform better. (Table 1)

Fingerprint Feature Extraction

In the paper, the fingerprint image was first enhanced by means of histogram equalization, wavelet transform, and then features were extracted by Zernike Moments invariant (ZMI) that was used as feature descriptor so that each feature vector extracted from each image normalization can represent the fingerprint. And to obtain a feature vector, that is, $F^{(1)} = (z_1, \ldots, z_k)$, where z_k is feature vector elements $1 \le k \le 36$, let the feature for the i-th user be $F_i^{(1)} = (z_1, ..., z_k)$. (Figure.4)

Face Feature Extraction

To generate feature vector of size n, first the given face image was normalized by histogram equalization, wavelet transform, and then computed by the Zernike moment. Let the result be the vector $F^{(2)} = (v_1, \ldots v_n)$. Similar to the extraction of fingerprint features, where v_n is feature vector elements $1 \le n \le 36$, let the feature for the i-th user is $F_i^{(2)} = (v_1, ..., v_n)$.(Figure.4)

Feature Combination

After the generation of the features from both fingerprint and the face image of the same person (say, the i-th user), it is possible to combine the two vectors $F_i^{(1)}$ and $F_i^{(2)}$ into one, with the total number of n+k component. That is, the feature vector for the i-th user is $F_i = (u_1, \ldots, u_{n+k})$, where feature vector elements $1 \le n+k \le 72$ are combined.

2.4 Classification

In this paper, an RBF neural network was used as a classifier in a face and fingerprint recognition system in which the inputs to the neural network are the feature vectors derived from the proposed feature extraction technique described in the previous section.

RBF Neural Network Description

RBF neural network (RBFNN)[16][17] is a universal approximator that is of the best approximation property and has very fast learning speed thanks to locally- tuned neurons (Park and Wsandberg, 1991; Girosi and Poggio, 1990; Huang, 1999a; Huang, 1999b). Hence, RBFNNs have been widely used for function approximation and pattern recognition.

A RBFNN can be considered as a mapping: $\Re^r \rightarrow \Re^s$. Let $P \in \Re^r$be the input vector, and $C_i \in \Re^r$ $(1 \ll i \ll u)$ be the prototype of the input vectors, then the output of each RBF unit can be written as:

$$R_i(P) = R_i(\|P - C_i\|) \quad i = 1, \ldots, u \tag{9}$$

where $\| \cdot \|$ indicates the Euclidean norm on the input space. Usually, the Gaussian function is preferred among all possible radial basis function due to the fact that it is factorable. Hence,

$$R_i(P) = \exp\left(-\frac{\|P-C_i\|^2}{\sigma_i^2}\right) \tag{10}$$

where σ_i is the width of the ith RBF unit. The jth output $y_i(P)$ of a RBFNN is

$$y_i(P) = \sum_{i=1}^{u} R_i(P) \times w(j,i) \tag{11}$$

where w(j,i) is the weight of the jth receptive field to the jth output.

In our experiments, the weight w(j,i), the hidden center Ci and the shape parameter of Gaussian kernel function σ_i were all adjusted in accordance with a hybrid learning algorithm combining the gradient paradigm with the linear least square (LLS)[18] paradigm.

System Architecture of the Proposed RBFNN

In order to design a classifier based on RBF neural network, we set a fixed number of input nodes in the input layer of the network. This number is equal to that of the combined feature vector elements. Also, the number of nodes in the output layer was set to be equal to that of the image classes, equivalent to 8 combined fingerprint and facial images. The RBF units were selected equal to the set number of the input nodes in the input layer.

For a neural network: feature vector elements of ZM, equal to 72, correspond to 72 input nodes of input layer. Our chosen number of RBF units of hidden layer is 72, and the number of nodes in the output layer is 8.

3 Experimental Results

3.1 Database of the Experiment

Our experiment was conducted on the public domain fingerprint images dataset DB4 FVC2004 [19], ORL face database [20].

Fig. 5. Captured sample fingerprint images from FVC 2004 database

Fig. 6. Sample face images from ORL face database

In DB4 FVC2004 database, the size of each fingerprint image is 288x384 pixels, and its resolution is 500 dpi. FVC2004 DB4 has 800 fingerprints of 100 fingers (8 images of each finger). Some sample fingerprint images used in the experimentation were depicted by Figure.5.

ORL face database is comprised of 400 images of 40 persons with variations in facial expressions (e.g. open/close eyes, smiling/non-smiling), and facial details (e.g. with wearing glasses/without wearing glasses). All the images were taken on a dark background with a 92 x 112 pixels resolution. Figure.6 shows an individual's sample images from the ORL database.

With the assumption that certain face images in ORL and fingerprint images in FVC belong to an individual, in our experiment, we used 320 face images (8 images from each of 40 individuals) in ORL face database, and 320 fingerprint images (8 images from each of 40 individuals) in FVC fingerprint database. Combining those images in pairs, we had our own database of 320 double images from 40 different individual, 8 images from each one that we named ORL-FVC database.

3.2 Evaluation

In this section, the capabilities of the proposed ZM-RBFN approach in multimodal authentication are demonstrated. A sample of the proposed system with two different

feature domains and of the RBF neural network was developed. In this example, concerning the ZM, all moments from the first 10 orders were considered as feature vectors, and the number of combined feature vector elements for these domains is 72. The proposed method was evaluated in terms of its recognition performance with the use of ORL-FVC database. Five images of each of 40 individuals in the database were randomly selected as training samples while the remaining samples without overlapping were used as test data. Consequently, we have 200 training images and 120 testing images for RBF neural network for each trial. Since the number of the ORL-FVC database is limited, we had performed the trial over 3 times to get the average authentication rate. Our achieved authentication rate is 96.55% (Table 2).

Table 2. Recognition rate of our proposed method

Test	Rate
1	97.25%
2	95.98%
3	96.43%
Mean	96.55%

In our paper, the effectiveness of the proposed method was compared with that of the mono-modal traits, typically human face recognition systems [21], and fingerprint recognition systems [22], of which the ZM has first 10 orders with 36 feature elements. From the comparative results of MMT shown in Table 3, it can be seen that the recognition rate of our multimodal system is much better than that of any other individual recognition, and that although the output of individual recognition may agree or conflict with each other, our system still searches for a maximum degree of agreement between the conflicting supports of the face pattern.

Table 3. The FAR,FRR and Accuracy values obtained from the monomodal traits

Trait	FRR(%)	FAR(%)	Accuracy
Face[21]	13.47	11.52	73.20
Fingerprint[22]	7.151	7.108	92.892

Also in our work, we conducted separated experiments on the technique of face, fingerprint, fusion at matching score and feature level. The comparison between the achieved accuracy of our proposed technique with that of each mentioned technique has indicated its striking usefulness and utility. (See in figure.7).

For the recognition performance evaluation, a False Acceptance Rate (FAR) and a False Rejection Rate (FRR) test were performed. These two measurements yield another performance measure, namely Total Success Rate (TSR):

$$\mathrm{TSR} = \left(1 - \frac{\mathrm{FAR+FRR}}{\text{total number of accesses}}\right) \times 100\% \tag{12}$$

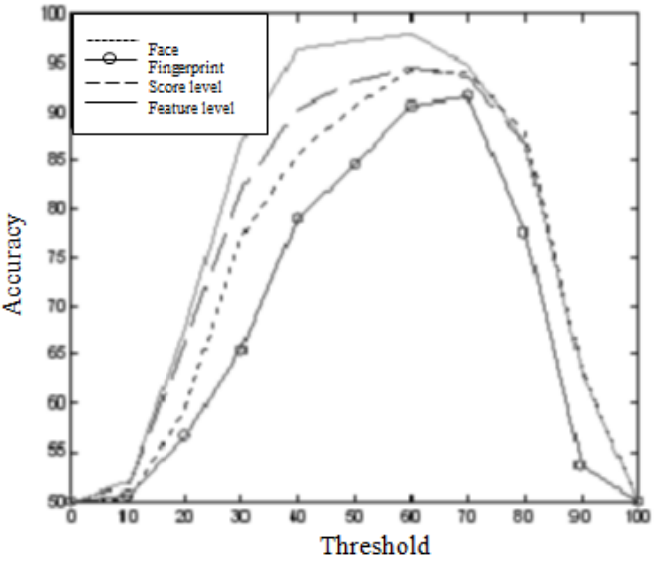

Fig. 7. The Accuracy curve of face, fingerprint, fusion at score and feature level

The system performance was evaluated by Equal Error Rate (EER) where FAR=FRR. A threshold value was obtained, based on Equal Error Rate criteria where FAR=FRR. Threshold value of 0.2954 was gained for ZM-RBF as a measure of dissimilarity.

Table 4 shows the testing results of verification rate with the first 10 order moments for ZM, based on their defined threshold value.

The results demonstrate that the application of ZM as feature extractors can best perform the recognition.

Table 4. Testing result of authentication rate of Multimodal

Method	**Thres**	**FAR(%)**	**FRR(%)**	**TSR(%)**
Proposed method	0.2954	4.95	1.12	96.55

4 Conclusion

This paper has outlined the possibility to augment the verification accuracy by integrating multiple biometric traits. In the paper, the authors have presented a novel approach in which both fingerprint and face images are processed with Zernike Moment- Radial Basis Functions to obtain comparable features. The reported experimental results have demonstrated a remarkable improvement in the accuracy achieved from the proper fusion of feature sets. It is also noted that fusing information from independent/ uncorrelated sources (face and fingerprint) at the feature level fusion enables better authentication than doing it at score level. This preliminary achievement does not constitute an end in itself, but suggests an attempt of a multimodal data

fusion as early as possible in parallel processing. However, the real feasibility of this approach, in a real application scenario, may heavily depend on the physical nature of the acquired signal; thus, it is assumed that further experiments on "standard" multimodal databases will allow better validation of the overall system performances.

References

1. Campbell Jr., J., Alyea, L., Dunn, J.: Biometric security: Government application and operations (1996), http://www.vitro.bloomington.in.us:8080/~BC/
2. Davies, S.G.: Touching big brother: How biometric technology will fuse flesh and machine. Information Technology @ People 7(4), 60–69 (1994)
3. Newham, E.: The Biometric Report. SJB Services, New York (1995)
4. Jain, K., Hong, L., Pankanti, S.: Biometrics: Promising Frontiers for Emerging Identification Market. Comm. ACM, 91–98 (February 2000)
5. Clarke, R.: Human identification in information systems: Management challenges and public policy issues. Information Technology @ People 7(4), 6–37 (1994)
6. Ross, A., Nandakumar, D., Jain, A.K.: Handbook of Multibiometrics. Springer, Heidelberg (2006)
7. Ross, A., Jain, A.K.: Information Fusion in Biometrics. Pattern Recognition Letters 24(13), 2115–2125 (2003)
8. Jain, A.K., Ross, A.: Multibiometric systems. Communications of the ACM 47(1), 34–40 (2004)
9. Zernike, F.: Physica (1934)
10. Du, S., Ward, R.: Wavelet based illumination normalization for face recognition. Department of Electrical and Computer Engineering. The University of British Coloumbia, Vancouver, BC, Canada; IEEE Transactions on Pattern Analysis and Machine Intelligence, 0-7803-9134-9 (2005)
11. Belhumeur, P.N., Hespanha, J.P., Kriegman, D.J.: Eigenfaces va. Fisherfaces: Recognition using class specific linear projection. IEEE Transactions on Pattern Analysis and Machine Intelligence 19(7), 711–720 (1997)
12. Cootes, T., Taylor, C., Cooper, D., Graham, J.: Active shape models-their training and applications. Computer Vision and Image Understanding 61(1), 38–59 (1995)
13. Cootes, T., Edwards, G., Taylor, C.: Active appearance models. IEEE Transactions on Pattern Analysis and Machine Intelligence 23(6), 681–685 (2001)
14. Mukundan, R., Ong, S.H., Lee, P.A.: Image analysis by Tchebichef moments. IEEE Transactions on Image Processing 10(9), 1357–1364 (2001)
15. Yap, P.T., Paramesran, R., Ong, S.H.: Image analysis by Krawtchouk moments. IEEE Transactions on Image Processing 12(11), 1367–1377 (2003)
16. Haddadnia, J., Faez, K.: Human face recognition using radial basis function neural network. In: Proc. 3rd International Conference on Human and Computer, HC 2000, Aizu, Japan, pp. 137–142 (September 2000)
17. Haddadnia, J., Ahmadi, M., Faez, K.: A hybrid learning RBF neural network for human face recognition with pseudo Zernike moment invariant. In: IEEE International Joint Conference on Neural Network, IJCNN 2002, Honolulu, Hawaii, USA, pp. 11–16 (May 2002)
18. Jang, J.-S.R.: ANFIS: Adaptive-Network-Based Fuzzy Inference System. IEEE Trans. Syst. Man. Cybern. 23(3), 665–684 (1993)

19. FVC. Finger print verification contest (2004), http://bias.csr.unibo.it/fvc2004.html
20. ORL. The ORL face database at the AT&T (Olivetti) Research Laboratory (1992), http://www.uk.research.att.com/facedatabase.html
21. Lajevardi, S.M., Hussain, Z.M.: Zernike Moments for Facial Expression Recognition. In: Internation Conference on Communication, Computer and Power, ICCCP 2009, Muscat, February 15-18, pp. 378–381 (2009)
22. Qader, H.A., Ramli, A.R., Al-Haddad, S.: Fingerprint Recognition Using Zernike Moments. The Internation Arab Journal of Information Technology 4(4) (October 2007)