N-Feature Neural Network Human Face Recognition

Javad Haddadnia [1], Karim Faez [2], Majid Ahmadi [3]

[1,3] Electrical and Computer Engineering Department, University of Windsor,
Windsor, Ontario, Canada, N9B 3P4
{javad, ahmadi}@uwindsor.ca
[2] Electrical Engineering Department, Amirkabir University of Technology, Tehran, Iran, 15914
kfaez@cic.aku.ac.ir

Abstract

This paper introduces a novel method for human face recognition that employs a set of different kind of features from the face images with Radial Basis Function (RBF) neural network called the Hybrid N-Feature Neural Network (HNFNN) human face recognition system. The face image is projected in each appropriately selected transform methods in parallel. The output of the RBF classifiers are fused together to make a decision. Experimental results for human face recognition confirm that the proposed method lends itself to higher classification accuracy relative to existing techniques.

1. Introduction

Face recognition may seem an easy task for humans, and yet computerized face recognition system still can not achieve a completely reliable performance. The difficulties arise due to large variation in facial appearance, head size, orientation and change in environment conditions. Such difficulties make face recognition one of the fundamental problems in pattern analysis. In recent years there has been a growing interest in machine recognition of faces due to potential commercial application such as film processing, law enforcement, person identification, access control systems, etc. A recent survey of the face recognition systems can be found in references [1-2].

A complete conventional human face recognition system should include three stages. The first stage involves detecting the location of face in arbitrary images [3-4]. The second stage requires extraction of pertinent features from the localized image obtained in the first stage. Finally the third stage involves classification of facial images based on the derived feature vector obtained in the previous stage.

In order to design a high accuracy recognition system, the choice of feature extractor is very crucial. Two main approaches to feature extraction have been extensively used in conventional techniques [2]. The first one is based on extracting structural facial features that are local structure of face images, for example, the shapes of the eyes, nose and mouth. The structure-based approaches deal with local information instead of global information. Therefor they are not affected by irrelevant information in an image. It has been shown that the structure-based approaches by explicit modeling of facial features have been troubled by the unpredictability of face appearance and environmental condition [2]. The second one is based on statistical approaches when features are extracted from the whole image and therefore use global information instead of local information. Since the global data of an image are used to determine the feature elements, data that are irrelevant to facial portion such as hair, shoulders and background may contribute to creation of erroneous feature vectors that can affect the recognition results [5].

In the field of pattern recognition, the combination of an ensemble of classifiers has been proposed to achieve image classification systems with higher performance in comparison with the best performance achievable employing a single classifier. This has been verified experimentally in the literature [6-7]. A number of image classification systems based on the combination of outputs of different classifier systems have been proposed. Different structures for combining classifier systems can be grouped in three configurations [8-9]. In the first group, the classifier systems are connected in cascade to create pipeline structure. In the second one, the classifier systems are used in parallel and their outputs are combined named it parallel structure. Finally the hybrid structure is a combination of the pipeline and parallel structures. In this paper, we propose a human face recognition system that can be designed based on hybrid structure classifier system to have evolutionary recognition results by developing the N-features and selecting them for the recognition problem. This human face recognition system uses available information and extracts more characteristics for face classification purpose by extracting different feature domains from input images. In this paper three different feature domains have been used for extracting features from input images. These include Pseudo Zernike Moment Invariant (PZMI) and Zernike Moment Invariant (ZMI) which produce the best result for human face recognition in comparison with other moments [10] and also Principal Component Analysis (PCA) [11].

Finally in this paper Radial Basis Function (RBF) neural network is used as the classifier. Recently RBF neural networks have found to be very attractive for many engineering problems. An important property of RBF neural networks is that they form a unifying link among many different research fields such as function approximation, regularization, noisy interpolation and pattern recognition. The increasing popularity of RBF neural networks is partly due to their simple topological structure, their locally tuned neurons and their ability to have a fast learning algorithm in comparison with the multi-layer feed forward neural networks [10][12]. The rest of this paper is organized as follows. The proposed human face recognition system is developed in section 2. Section 3 presents the feature domains. The classification technique is described in section 4. Finally section 5 and 6 presents the experimental results and conclusion.

2. The Proposed HNFNN

Fig. (1) shows a conventional human face recognition. This system uses one feature domain and one classifier. Usually neural network are used as classifier therefore this conventional method named Single Feature Neural Network (SFNN) human face recognition system. The Proposed human face recognition has been shown in Fig. (2). Unlike conventional human face recognition system the proposed HNFNN system is developed in five stages. In the first step, face localization process is done. To ensure a robust, accurate feature extraction that distinguishes between face and nonface region in an image, we require the exact location of the face region. In this paper we have used a modified version of shape information technique that presented in reference [3] for face localization. After face localization, in the second stage we have created a subimage, which contains information needed for recognition algorithm. By using a subimage, data that are irrelevant to facial portion are disregarded. In the third stage, different features are extracted in parallel from the derived subimage. These features are obtained from the different domains. The fourth stage required classification, which classify a new face image, based on the chosen features, into one of the possibilities. This is done for each feature domain in parallel as Fig. (2) shows. Finally the last stage combines the outputs of each neural network classifiers to construct the identification. In this paper majority method has been selected for decision strategy.

2.1. Face Localization

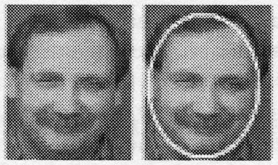
Many algorithms have been proposed for face localization and detection. A critical survey on face localization and detection can be found in reference [2]. The ultimate goal of the face localization is finding an object in an image as a face candidate that its shape resembles the shape of a face. Faces are characterized by elliptical shape and an ellipse can approximate the shape of a face. A technique is presented in [3], which finds the best-fit ellipse to enclose the facial region of the human face in a frontal view of facial image. In this algorithm an ellipse model with five parameters has been used. Initially connected component objects are determined by applying a region-growing algorithm. Consequently for each connected component object with a given minimum size, the best-fit ellipse is computed on the basis of its moments. To assess how well the connected component object is approximated by its best-fit ellipse, we define the new distances measure between the connected component object and the best-fit ellipse as follows:

$$\phi_i = P_{inside} / \mu_{0,0}$$

$$\phi_o = P_{outside} / \mu_{0,0}$$

where the P_{inside} is the number of background points inside the ellipse, $P_{outside}$ is the number of points of the connected component object that are outside of the ellipse and $\mu_{0,0}$ is the size of the connected component object [3]. The connected component object is closely approximated by its ellipse when ϕ_i and ϕ_o is as small as possible. We have named the threshold values for both ϕ_i and ϕ_o as Facial Candidate Threshold (FCT). Our experimental study indicates that when FCT is less than 0.1 the connected component is very similar to ellipse therefore it is a good candidate as a face region. If ϕ_i and ϕ_o are grater than FCT, there is no face region in the input image therefor we reject it as nonface image [13]. An example of applying this method for locating a face candidate and rejecting nonface image has been shown in Fig. (3). Subsequently the rest of the system processes the selected face candidates for recognizing.

$\phi_i = 0.065$, $\phi_o = 0.008$



$\phi_i = 0.062$, $\phi_o = 0.011$

$\phi_i = 0.15$, $\phi_o = 0.191$

Figure 3: Distinguishing between face and nonface using best-fit ellipse and FCT threshold

2.2. Subimage Formation

The subimage encloses the pertinent information around the face in an ellipse that was explained in section 2.1 while pixel value outside the ellipse is set to zero. Fig. (4) shows sample of selecting of face location and creating of subimage in feature extracting respectively. By using subimage, data that are irrelevant to facial portion such as hair, shoulders and background are disregarded and the speed of computing various features is increased due to smaller pixels content of the subimages.

(a) Face location

(b) subimage

Figure 4: Creating subimages from face images

3. N-Features Domains

In order to design a good face recognition system, the choice of feature extractor is very crucial. To design a system with low to moderate complexity the feature vectors should contain the most pertinent information about the face to be recognized. Face recognition system should be capable of recognizing unpredictability of face appearance and changing environment. In the proposed system, N different feature domains are extracted from the derived subimages in parallel. Therefore this approach can extract more characteristics of face images for classification purpose. In this paper we set N=3 and therefore three different kind of feature domains have been selected. These include PZMI, ZMI and PCA.

3.1. Pseudo Zernike Moment Invariant

The advantages of considering orthogonal moments are that they are shift, rotation and scale invariant and very robust in the presence of noise. We have used the PZMI for generating feature vector elements. Pseudo Zernike polynomials are well known and widely used in the analysis of optical systems. Pseudo Zernike polynomials are orthogonal set of complex-valued polynomials defined as [10][14]:

$$V_{nm}(x,y) = R_{nm}(x,y)\exp\left(jm\tan^{-1}\left(\frac{y}{x}\right)\right)$$

where $x^2 + y^2 \le 1$, $n \ge 0$, $|m| \le n$ is even and Radial polynomials R_{nm} are defined as:

$$R_{nm}(x,y) = \sum_{s=0}^{n-|m|} D_{n,|m|,s}\,(x^2+y^2)^{\frac{n-s}{2}}$$

$$D_{n,|m|,s} = (-1)^s \frac{(2n+1-s)!}{s!(n-|m|-s)!(n-|m|-s+1)!}$$

The PZMI of order n and repetition m, can be computed as follows:

$$PZMI_{nm} = \frac{n+1}{\pi} \sum_{(n-m-s)\,even,\,s=0}^{n-|m|} D_{n,|m|,s} \sum_{a=0}^{k}\sum_{b=0}^{m} \binom{k}{a}\binom{m}{b} (-j)^b CM_{2k+m-2a-b,\,2a+b}$$
$$+ \frac{n+1}{\pi} \sum_{(n-m-s)\,odd,\,s=0}^{n-|m|} D_{n,|m|,s} \sum_{a=0}^{d}\sum_{b=0}^{m} \binom{d}{a}\binom{m}{b} (-j)^b RM_{2d+m-2a-b,\,2a+b}$$

where $k=(n-s-m)/2$, $d=(n-s-m+1)/2$, $CM_{p,q}$ is the scale invariant central moments and $RM_{p,q}$ is the scale invariant radial geometric moments are defined as:

$$CM_{p,q} = \frac{\mu_{pq}}{M_{00}^{(p+q+2)/2}}$$

$$RM_{p,q} = \frac{\sum_x \sum_y f(x,y)(\hat{x}^2+\hat{y}^2)^{1/2}\,\hat{x}^p \hat{y}^q}{M_{00}^{(p+q+2)/2}}$$

where $\hat{x} = x - x_0$, $\hat{y} = y - y_0$ and M_{pq}, μ_{pq} and x_0, y_0 are defined as follow:

$$M_{pq} = \sum_x \sum_y f(x,y) x^p y^q$$

$$\mu_{pq} = \sum_x \sum_y f(x,y)(x - x_0)^p (y - y_0)^q$$

$$x_0 = M_{10} / M_{00}$$

$$y_0 = M_{01} / M_{00}$$

3.2. Zernike Moment Invariant

Zernike polynomials are orthogonal set of complex-valued polynomials defined as:

$$V_{nm}(x,y) = R_{nm}(x,y)\exp(jm\tan^{-1}(\frac{y}{x}))$$

where $x^2 + y^2 \le 1$, $n \ge 0$, $|m| \le n$ and $n - |m|$ is even and Radial polynomials { R_{nm} } are defined as:

$$R_{nm}(x,y) = \sum_{s=0}^{(n-|m|)/2} S_{n,|m|,s} (x^2 + y^2)^{\frac{n-2s}{2}}$$

where:

$$S_{n,|m|,s} = (-1)^s \frac{(n-s)!}{s!(\frac{n+|m|}{2} - s)!(\frac{n-|m|}{2} - s)!}$$

The complex Zernike moments [19] of order *n* and repetition *m* are given by:

$$ZMI_{nm} = \frac{n+1}{\pi} \sum_x \sum_y f(x,y) V^*_{nm}(x,y)$$

To utilize the shift invariant property of moments, we have used ZMI from the scale invariant central moments ($CM_{p,q}$) as follows [14]:

$$ZMI_{nm} = \frac{n+1}{\pi} \sum_{s=0}^{(n-|m|)/2} \sum_{a=0}^{b} \sum_{d=0}^{|m|} \binom{|m|}{d} \binom{b}{a} (-1)^d S_{n,|m|,s} CM_{n-2s-2a-d,2a+d}$$

where $b = (n - |m|)/2 - s$.

3.3. Principal Component Analysis

PCA is a well-known statistical technique for feature extraction. Each $M \times N$ image in the training set was row concatenated to form $MN \times 1$ vectors $\tilde{A}_i$. Given a set of N_T training images $\{\tilde{A}_i\}_{i=0,1,\ldots N_T}$ the mean vector of the training set was obtained as:

$$\overline{A} = \frac{1}{N_T} \sum_{i=1}^{N_T} \tilde{A}_i$$

The average vector was subtracted out from the training vectors to obtain:

$$A_i = \tilde{A}_i - \overline{A} \quad , i=1,2,3,\ldots N_T$$

An $N_T \times MN$ matrix A was constructed with the A_i^T as its row vectors. The singular value decomposition of A can then be written as:

$$V^T AU = |\Sigma | 0|$$

where Σ is an $N_T \times N_T$ diagonal matrix with singular values $s_i > 0$ arranged in descending order, and V and U are $N_T \times N_T$ and $MN \times MN$ orthogonal matrices, respectively. V is composed of the eigenvectors of AA^T, while U is composed of the eigenvectors of AA^T. These are related by:

$$\hat{U} = A^T V$$

where $\hat{U}$ consists of the eigenvectors of AA^T, which correspond to the non-zero singular values. This relation allows a smaller $N_T \times N_T$ eigenvalue problem for AA^T to be solved, and to subsequently obtain $\hat{U}$ by matrix multiplication.

The projection of a face vector onto the space of N_T eigenfaces results in an N_T-dimensional feature vector of projection weights. As PCA has the property of packing the greatest energy into the least number of principal components, the smaller principal components which are less than a threshold can be discarded with minimal loss in representational capability. This dimensionality reduction results in face weight vectors of dimensions $\tilde{N}_T < N_T$. An appropriate value of $\tilde{N}_T$ can be chosen by considering the Basis Restriction Error (BRE) as a function of $\tilde{N}_T$ [11]. This gradual decrease in error is significant for recognition techniques based on eigenfaces where storage and computational performance are directly related to N_T.

4. Classifier Design

Neural networks have been employed and compared to conventional classifiers for a number of classification problems. The results have shown that the accuracy of the neural network approaches equivalent to, or slightly better than, other methods. Also, due to the simplicity, generality and good learning ability of the neural networks, these types of classifiers are found to be more efficient [10][12]. Radial Basis Function (RBF) neural networks have found to be very attractive for many engineering problem because: (1) they are universal approximators, (2) they have a very compact topology and (3) their learning speed is very fast because of their locally tuned neurons. Therefore the RBF neural networks serve as an excellent candidate for pattern applications and attempts have been carried out to make the learning process in this type of classification faster than normally required for the multi-layer feed forward neural networks [12].

4.1. RBF Neural Network Structure

An RBF neural network structure is shown in Fig. (5), which has architecture similar to that of a traditional three-layer feed forward neural network. The construction of the RBF neural network involves three different layers with feed forward architecture. The input layer of this network is a set of *n* units, which accept the elements of an *n*-dimensional input feature vector. The input units are fully connected to the hidden layer with *r* hidden units. Connections between the input and hidden layers have unit weights and, as a result, do not have to be trained. The goal of the hidden layer is to cluster the data and reduce its dimensionality. In this structure hidden layer is named RBF units. The RBF units are also fully connected to the output layer. The output layer supplies the response of neural network to the activation pattern applied to the input layer. The transformation from the input space to the RBF-unit space is nonlinear, whereas the transformation from the RBF-unit space to the output space is linear.

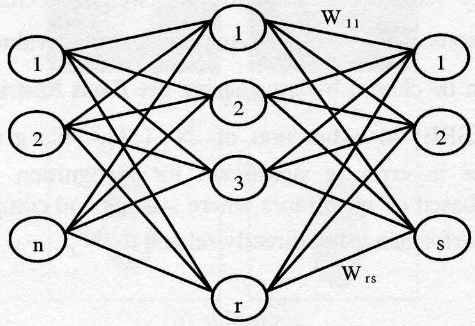


Figure 5: RBF neural network structure

The RBF neural network is a class of neural networks, where the activation function of the hidden units is determined by the distance between the input vector and a prototype vector. The activation function of the RBF units is expressed as follow [10][12]:

$$R_i(x) = R_i\left(\frac{\| x - c_i \|}{\sigma_i}\right) \quad , \quad i=1,2,\ldots,r$$

It should be noted that x is an *n*-dimensional input feature vector, c_i is an *n*-dimensional vector called the center of the RBF unit, σ_i is the width of RBF unit and r is the number of the RBF units. Typically the activation function of the RBF units is chosen as a Gaussian function with mean vector c_i and variance vector σ_i as follows:

$$R_i(x) = \exp\left(- \frac{\| x - c_i \|^2}{\sigma_i^2}\right)$$

Note that σ_i^2 represents the diagonal entries of covariance matrix of Gaussian function. The output units are linear and therefore the response of the j-th output unit for input x is given as:

$$y_j(x) = b(j) + \sum_{i=1}^{r} R_i(x) w_2(i,j)$$

where $w_2(i,j)$ is the connection weight of the i-th RBF unit to the j-th output node and $b(j)$ is the bias of the j-th output. The bias is omitted in this network in order to reduce network complexity. Therefore:

$$y_j(x) = \sum_{i=1}^{r} R_i(x) \times w_2(i,j)$$

4.2. RBF Based Classifier Design

RBF neural network classifier can be viewed as a function mapping interplant that tries to construct hypersurfaces, one for each class, by taking a linear combination of the RBF units. These hypersurfaces can be viewed as discriminant functions, where the surface has a high value for the class it represents and a low value for all others. An unknown input feature vector is classified as belonging to class associated with the hypersurface with the largest output at that point. In this case the RBF units' serve as components in a finite expansion of the desired hypersurface where the component coefficients (the weights) have to be trained [12][15].

For designing a classifier based on RBF neural network, we have set the number of input nodes in the input layer of neural network equal to the number of feature vector elements. The number of nodes in the output layer is set to the number of image classes. The RBF units are selected using the following clustering procedure [15]:

Step1: Initially the RBF units are set equal to the number of outputs.

Step2: For each class k, the center of RBF units (c_k) is selected as the mean value of the sample patterns belonging to the same class, i.e.

$$c_k = \frac{\sum_{i=1}^{N^k} x_k^i}{N^k}, \quad k=1,2,\ldots,s$$

where x_k^i is the i-th sample pattern belonging to class k and N^k is the number of samples pattern in the same class.

Step3: For each class k, compute the distance d_k^f from the mean c_k to the farthest sample pattern x_k^f in that class:

$$d_k^f = \| x_k^f - c_k \|, \quad k=1,2,\ldots,s$$

Step4: For each class k, compute the distance $dc(k,j)$ between the mean of the class and the mean of other classes and then find the minimum among the distances computed:

$$dc(k,j) = \| c_k - c_j \|$$

$$d_{min}(k,l) = \min(dc(k,j))$$

where j=1,2,...,s and $j \neq k$. Then we check the relationship between $d_{min}(k,l)$, d_k^f, and d_l^f. If $d_k^f + d_l^f \leq d_{min}(k,l)$ then class k is not overlapping with other classes. Otherwise, class k is overlapping with other classes and misclassifications may occur in this case.

Step5: If two classes are overlapped strongly, we first split one of the classes into two to remove the overlap. If the overlap is not removed the second class is also split. This requires addition of a new RBF unit to the hidden layer. This can be seen in Fig. (5).

Step6: Repeat steps 2 to 5 until all the training sample patterns are clustered correctly in the hidden layer.

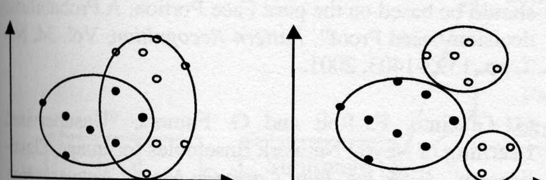


Figure 6: One class split into two classes

The above procedure determines the number of RBF units, *r*, in the RBF neural network structure. In addition, we select the mean values of all the clusters, which have been determined by the above procedure as the initial center of the RBF units.

Training of the RBF neural network involves estimating output connection weights, centers and widths of the RBF units. We use the HLA method, which combines the gradient method and the linear least squared method for training RBF neural network [15].

5. Experimental Results

To check the utility of our proposed algorithm experimental studies are carried out on the ORL database images of Cambridge University. 400 face images from 40 individuals in different states from the ORL database have been used to evaluate the performance of the proposed method. None of the 10 samples are identical to each other. They vary in position, rotation, scale and expression. In this database each person has changed his face expression in each of 10 samples (open/close eye, smiling/not smiling). For some individuals, the images were taken at different times, varying facial details (glasses/no glasses). Samples of database used are shown in Fig. (7).

Figure 7: Sample of face database

A sample of the proposed system with three different feature domains and the RBF neural network has been developed. In this example, for the PZMI and ZMI all moments from order 9 to 10 have been considered as feature vector elements. The feature vectors for these domains have 21 elements for the PZMI and 11 elements for the ZMI. Also for the PCA feature vector has been created based on the 30 largest PCA number for each image. A total of 200 images have been used to train and another 200 for test. Each training set consists of 5 randomly chosen images from the same class in the training stage. In the face localization step, shape information algorithm with FCT=0.1 has been applied to all images. Subsequently, subimage has been created from the derived localized face image. Recognition rate of 99.7% was obtained using this proposed technique.

To compare the effectiveness of the proposed method in comparison with the SFNN human face recognition systems, we have developed the SFNN systems using the PZMI+RBF [10], ZMI+RBF [10] and PCA+RBF. For these systems we have selected the PZMI as feature domain with order 9 and 10 which has 21 elements, ZMI with order 9 and 10 with 11 feature elements and finally PCA with 30 largest value. The comparison of the HNFNN with each individual classifiers as a function of class number has been shown in Fig. (8). From the results, these can be no doubt that the recognition rate of the HNFNN is much better than that of any individual classifiers. From this figure it is clear that the output of each individual classifier may agree or conflict with each other but the HNFNN search for a maximum degree of agreement between the conflicting supports of a face pattern.

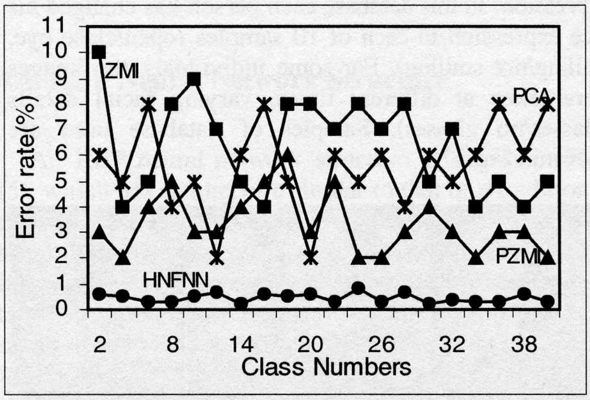


Figure 8: Error rate in the HNFNN and each individual classifier based on class number

Also to demonstrate the effectiveness of the human face recognition system by the proposed method, we have compared the proposed method with other algorithms. To make this comparison meaningful, an average overall error rate is defined as:

$$E_{ave} = \frac{\sum_{i=1}^{m} E_{NM}(i)}{mN_t}$$

where m is the number of experimental runs, each being performed on random partition of the database into sets, $E_{NM}(i)$ is the number of misclassification for the i-th run, and N_t is the number of total testing images for each runs. In our study the ORL database was used in the experiments and methods reported in [10], [16] and [17] were used for comparison purpose. Table (1) shows the results of this comparative study. In this table the SFNN denotes the Shape Information with Neural Network that was reported in [10], CNN is the Convolution Neural Network method used in [16] and also NFL for Nearest Feature Line method in [17].

Table 1: Error rate in different methods

Methods	E_{ave} %
CNN [16]	3.83
NFL [17]	3.125
SFNN [10]	1.323
Proposed method	0.283

6. Conclusion

This paper presented a novel method for the recognition of human faces in 2-Dimensional digital images. The proposed technique is based on the Hybrid N-Feature Neural Network (HNFNN) structure. An implementation example is given to demonstrate the feasibility of the HNFNN human face recognition system. It employs the RBF neural networks and three feature domains. These include PZMI, ZMI and PCA. The highest recognition rate of 99.7% with the ORL database was obtained using this proposed algorithm. Comparison with some of the existing traditional technique in the literatures on the same database indicates the usefulness of the proposed technique.

References

[1] M. A. Grudin, "On Internal Representation in face Recognition Systems", *Pattern. Recognition.* Vol. 33, No. 7, pp.1161-1177, 2000.

[2] J. Daugman, "Face Detection: A Survey", *Computer Vision and Image Understanding*, Vol. 83, No. 3, pp. 236-274, Sept. 2001.

[3] J. Haddadnia, K. faez, "Human Face Recognition Based on Shape Information and Pseudo Zernike Moment", *5th Int. Fall Workshop Vision, Modeling and Visualization*, Saarbrucken, pp. 113-118, Germany, Nov. 22-24, 2000.

[4] J. Wang and T. Tan, "A New Face Detection Method Based on Shape Information", *Pattern Recognition Letter*, Vol. 21, pp. 463-471, 2000.

[5] L. F. Chen, H. M. Liao, J. Lin and C. Han, "Why Recognition in a statistic-based Face Recognition System should be based on the pure Face Portion: A Probabilistic decision-based Proof", *Pattern Recognition*, Vol. 34, No. 7, pp. 1393-1403, 2001.

[6] G. Giacinto, F. Roli and G. Fumera, "Unsupervised Learning of Neyral Network Ensembles for Image Classification", *IEEE Int. Joint Conf. On Neural Network*, Vol. 3, pp. 155-159, 2000.

[7] J. Kittler, M. Hatef, R. P. W. Duin and J. Matas, "On Combining Classifier", *IEEE Trans. On Patt. Anal. and Mach. Intel.*, Vol. 20, pp. 226-239, March 1998.

[8] T. K. Ho, J. J. Hull and S. N. Srihari, "Decision Combination in Multiple Classifier Systems", *IEEE Trans. On Patt. Anal. and Mach. Intel.*, Vol. 16, No. 1, pp. 66-75, Jan. 1994.

[9] Y. Lu, "Knowledge Integrations in a Multiple Classifier System", *Applied Intelligence,* Vol. 6, No. 2, pp. 75-86, April 1996.

[10] J. Haddadnia, K. Faez, P. Moallem, "Neural Network Based Face Recognition with Moments Invariant", *IEEE Int. Conf. On Image Processing,* Vol. I, pp. 1018-1021, Thessaloniki, Greece, 7-10 October 2001.

[11] M. Truk and A. Pentland, "Eigenfaces for Recognition", *Journal Cognitive Neuroscience,* Vol. 3, No. 1, pp. 71-86, 1991.

[12] W. Zhou, "Verification of the nonparametric characteristics of backporpagation neural networks for image classi

[13] fication", *IEEE Trans. On Geo. and Remote Sensing,* Vol. 37, No. 2, pp. 771-779, March 1999.

[14] J. Haddadnia, M. Ahmadi, K. Faez, "An Efficient Method for Recognition of Human Face Recognition Using Higher Order Pseudo Zernike Moment Invariant", *The 5th IEEE Int. Conf. on Automatic Face and Gesture Recognition,* Washington, DC, USA, May 20-21, 2002, Accepted for presentation.

[15] C. H. The and R. T. Chin, "On Image Analysis by the Methods of Moments", *IEEE Trans. On Patt. Anal. and Mach. Intel.*, Vol. 10, No. 4, pp. 496-513, 1988.

[16] J. Haddadnia, M. Ahmadi, K. Faez, "A Hybrid Learning RBF Neural Network for Human Face Recognition with Pseudo Zernike Moment Invariant", *IEEE Int. Joint conf. on Neural Network,* Honolulu, HI, May 12-17, 2002, Accepted for presentation.

[17] S. Lawrence, C. L. Giles, A. C. Tsoi and A. D. Back, "Face Recognition: A Convolutional Neural Networks Approach", *IEEE Trans. on Neural Networks, Special Issue on Neural Networks and Pattern Recognition*, Vol. 8, No. 1, pp. 98-113, 1997.

[18] S. Z. Li and J. Lu, "Face Recognition Using the Nearest Feature Line Method", *IEEE Trans. on Neural Networks*, Vol. 10, pp. 439-443, 1999.

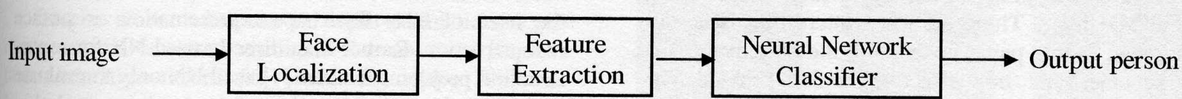


Figure 1: SFNN human face recognition system

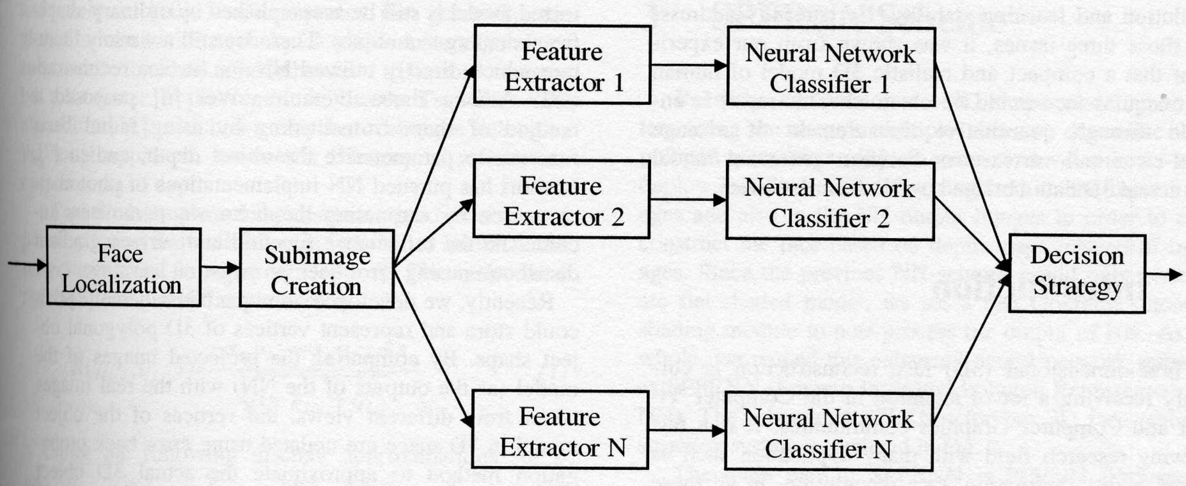


Figure 2: HNFNN based human face recognition system