Face Verification Based on Singular Value Decomposition and Radial Basis Function Neural Network+

Yunhong Wang, Tieniu Tan and Yong Zhu

National Laboratory of Pattern Recognition
Institute of Automation, Chinese Academy of Sciences, Beijing, P. R. China, 100080
{wangyh, tnt}@nlpr.ia.ac.cn

Abstract

Face is an important biometric feature for personal identification. This paper describes a new face verification methods based on singular value decomposition and RBF neural networks. The proposed method utilizes the positive samples and negative samples learning ability of RBF neural networks to improve singular values based face verification. Experiment results show that the novel face verification method is effective and possesses several desirable properties when it compared with many existing methods.

[key words] face identification, SVD, neural network

I. Introduction

Face verification plays an important role in biometrics based personal identification. A biometrics verification system is designed to verify or recognize the identity of a living person on the basis of his/her physiological characters, such as face, fingerprint, and iris, or some aspects of behavior such as handwriting or keystroke pattern. The need for reliable identification of interacting users is obvious. The biometrics verification technique acts as an efficient method and has wide applications in the areas of information retrieval, automatic

+This work is funded by research grants from the NSFC, the 863 Program and the Chinese Academy of Sciences.

banking, control of access to security areas, buildings, and so on. Compared with other biometric verification techniques, face recognition has the advantages of being passive and non-intrusive [3].

Great progress has been made in face recognition in the past 20 years. For almost all previously proposed techniques, the success of face recognition depends on the solution of two problems: representation and matching [1]. The representation of a pattern can be considered as feature extraction in pattern recognition. In [2], image features are divided into four groups: visual features, statistical pixel features, transform coefficient features, and algebraic features. The algebraic features represent intrinsic properties of an image and have good stability. Hong [2] suggested that the algebraic features are valid features in object recognition such as face recognition. He proposed a singular value decomposition (SVD) based recognition method which uses the singular values as the feature vectors. The effectiveness of SVD has been tested in [2] and [10] respectively. In [2], an error rate of 42.47% was recorded which was thought to be caused by the statistical limitations of the small samples. Cheng [10] proposed a human face recognition method based on the statistical model of small sample size that also used the singular values as the face features. In his paper, an optimal discriminate transformation is constructed to transform an original space of singular value (SV) vectors into a new space whose dimension is significantly lower than that of the original space to minimize the small sample size effect. That approach was tested on 64 facial images of eight people. Good discrimination ability was obtained with an accuracy rate of 100% [10]. It should be noted that in order to make the method independent of translation, rotation and scaling, the images were represented by Goshtasby's shape matrices. The

Goshtasby's shape matrices are invariant to translation, rotation, and scaling of the facial images and are obtained by polar quantization of the shape [11]. The above two methods have never been tested with large face databases and their effectiveness with large databases remains unknown (especially when there are variations in lighting and viewpoint). Furthermore, both methods use only singular values as face features.

In general, an unsupervised learning approach can not get a high recognition rate. Under conditions where we can not acquire a large number of face images for every person, utilizing all available samples is very important. This means that not only positive samples but also negative samples need to be learned. A radial basis function (RBF) neural network classifier makes it possible to learn both positive and negative samples. Since the structure of RBF neural networks determines the performance of classification, we should design the network structure to satisfy our requirements.

In order to utilize all available samples, a suitable RBF classifier which has learned both “positive” and “negative” samples in advance is needed. In this paper we proposed a RBF neural network based classification method whose features are the SVs of face images. Good verification performance is achieved by supervised learning.

In Section II the Singular Value Decomposition (SVD) is described. A SVD based feature extraction method is presented in Section III. In Section IV we introduce the RBF neural network and prove that this kind of neural network has the ability of learning both positive and negative samples. Section IV gives the design of the RBF neural network classifier of the face verification system. Experimental results will be given in Section V and some conclusions are drawn in Section VI.

II. Singular Value Decomposition

Singular value decomposition (SVD) provides a new way for extracting algebraic features from an image. SVD has been used in many fields such as data compression, signal processing and pattern analysis [4].

The main theoretical properties of SVD relevant to face image recognition are

- The SVD of a face image has good stability. When a small perturbation is added to an face image, large variance of its SVs does not occur.
- Singular values represent algebraic properties of an image [2]. To some extent, SV features possess algebraic and geometric invariance.

The properties of the singular values are described in detail in the following:

Theorem 1 (SVD) If $A \in R^{m\times n}$, then there exist orthogonal matrices

$$U = [u_1, \cdots, u_m] \in R^{m\times m}$$

and $$V = [v_1, \cdots, v_n] \in R^{n\times n}$$

such that $$U^T AV = diag(\sigma_1, \cdots, \sigma_p)$$

where $p = \min(m,n)$, $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_p \geq 0$.

σ_i, i=1,2,...p, are the singular values of A. The singular values are the square roots of the eigenvalues λ_i of AA^H or $A^H A$, that is $\sigma_i = \sqrt{\lambda_i}$.

Theorem 2 (The stability of SV) Assume $A^{m\times n}, B^{m\times n} \in R^{m\times n}$, and their singular values are $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_n$, $\tau_1 \geq \tau_2 \geq \cdots \geq \tau_n$, respectively, then $|\sigma_i - \tau_i| \leq \|A - B\|_2$

This means that when there is a disturbance at A, the variation of its singular values is not more than the $\|\cdot\|_2$ -norm of the disturbance matrix.

Theorem 3 (The scaling property) If the singular values of $A^{m\times n}$ are $\sigma_1, \sigma_2, \cdots, \sigma_k$, the singular values of $\alpha * A^{m\times n}$ are $\sigma_1^*, \sigma_2^*, \cdots, \sigma_k^*$, then

$$|\alpha|(\sigma_1, \sigma_2, \cdots, \sigma_k) = (\sigma_1^*, \sigma_2^*, \cdots, \sigma_k^*)$$

Theorem 4 (The rotation invariant property) If P is a unitary matrix, then the singular values of PA are the same as those of A.

The above properties of SVD are very desirable in face recognition, especially when images are taken under different noise and viewpoint conditions.

III Face Image Feature Extraction

Singular values represent important attributes of a matrix. As images may be regarded as matrices, singular values can serve as image features in image similarity evaluation.

We can prove that the difference of two matrices can be presented in terms of singular values.

If

$$A = USV^T = \sum_{i=1}^{k} \sigma_i u_i v_i^T \tag{1}$$

$$B = P\Sigma Q^T = \sum_{i=1}^{k} \gamma_i p_i q_i^T \tag{2}$$

then

$$\|A\|_2 = \sqrt{\sigma_1^2 + \sigma_2^2 + \cdots + \sigma_k^2} \tag{3}$$

$$\|B\|_2 = \sqrt{\gamma_1^2 + \gamma_2^2 + \cdots + \gamma_k^2} \tag{4}$$

and

$$\|A\|_2 + \|B\|_2 \geq \|A - B\|_2 \geq \|A\|_2 - \|B\|_2 \tag{5}$$

That is

$$\sqrt{\sigma_1^2+\sigma_2^2+\cdots+\sigma_k^2}+\sqrt{\gamma_1^2+\gamma_2^2+\cdots+\gamma_k^2}\geq\|A-B\|_2\geq\left|\sqrt{\sigma_1^2+\sigma_2^2+\cdots\sigma_k^2}-\sqrt{\gamma_1^2+\gamma_2^2+\cdots\gamma_k^2}\right| \tag{6}$$

After face images have been normalized, (so that they have similar scales, lighting, etc.) singular values can serve as image features for personal verification.

In practical systems, we can ensure the scale of face images by detection and transformation.

IV. Design of RBF Classifier

Two visually distinct images may have similar singular values but the U and V of SVDs are different. In order to avoid misclassification, we can learn the samples that are easy to be confused in advance. A RBF neural network has the ability of learning both “positive” samples and “negative” samples, so it is ideal to be used as the classifier in face based identity verification. An important issue is to design an appropriate RBF neural network structure and to extract the “negative” samples.

The general structure of a RBF neural network is shown in Fig. 1.

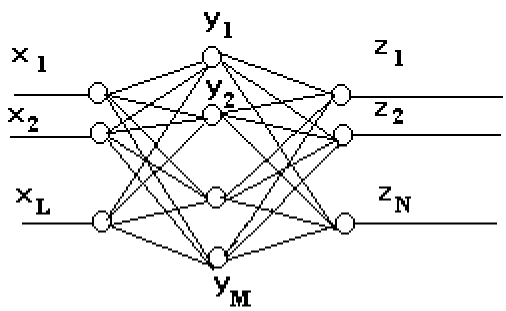

Fig. 1. The structure of RBF neural network

The output of the jth hidden node is

$$\mathbf{y}_h=\Phi(\|\mathbf{X}_k-\mathbf{C}_j\|)\quad \text{j=1,2,...}\ N_0 \tag{7}$$

where $\mathbf{X}_k$ is a N-dimensional input vector, $\mathbf{C}_j$ is the jth RBF neural network's center, N_0 is the number of hidden units, and $\Phi(\cdot)$ is a nonlinear, radial symmetric function whose center is C_j. In this paper, Gauss function is used as the basis function. So the output of the hidden layer is:

$$\mathbf{y}_{\mathbf{h}k} = \Phi(\|\mathbf{X_k} - \mathbf{C}_j\|) = \exp\left[-\sum_{i=1}^{N} \frac{(x_i - C_{ij})^2}{2\rho_j^2}\right] \tag{8}$$

The output of the ith output unit of the RBFN is:

$$z_{ki} = \sum_h w_{ih} \Phi(\|\mathbf{X}_k - \mathbf{C_j}\| + w_{k0} \tag{9}$$

The design of the RBF neural network is to choose the hidden units of the network and determine the centers and shape parameters, where $\mathbf{C}_j$ and ρ_j are obtained by fuzzy clustering.

The singular values of each matrix may not the most appropriate features for face identification as different persons' face images may have similar singular values. This problem can be solved by supervised learning. We should select the features easy to be confused in advance, then process them in supervised learning. These features can be obtained by fuzzy clustering.

According to the classical clustering method, the data set X is divided into some ordinary subsets, namely S_i (i=1,2,...,c), μ_{S_i} indicates membership of x_k belonging to a cluster.

$$\mu_{s_i}(x_k): X \rightarrow \{0,1\} \tag{10}$$

According to the fuzzy clustering method, the data set X is divided into some fuzzy subsets, namely $\underline{S_i}$ (i=1,2,...c).The extent of sample x_k being subordinate to a sort is expressed by the membership function $\mu_{\underline{S_i}}$.

$$\mu_{\underline{S_i}}(x_k): X \rightarrow [0,1] \tag{11}$$

So the advantage of the fuzzy clustering is obvious. It can indicate both the centers of each cluster and the linkage of different clusters.

Because the fuzzy clustering method has many advantages, we use it to select center and shape parameters of the network. We should notice that the shape parameters reveal the range in which the input samples are considered to be similar to the typical samples. The appropriate shape parameters should be chosen in accordance with the distribution of the sample space. For example, the shape parameter ρ should be small in cases where the samples are thickly scattered; it should be large in cases where the samples are thinly scattered. We use the results of the membership function of the fuzzy clustering. Each ρ_j is obtained by choosing the samples whose membership function are greater than 0.7 in each kinds, calculating the distance between the center and the selected samples, then averaging these distances. Suppose that there are L samples in a kind which correspond to this condition, then

$$\rho_j = \frac{1}{L}\sum_{i=1}^{L} d_{ij} \tag{12}$$

where d_{ij} is the distance between the selected samples and the center parameters.

Based on the above idea, we choose the center and shape parameters by fuzzy clustering.

The number of clusters should be given in advance. In this paper the fuzzy cluster number depends on the negative samples. We form a RBF neural network classifier for each person. For example, to establish a RBF neural network classifier for person A, the steps are as follows.

(1) Choose 3-5 clustering centers from the training database of face images of A.

(2) Set the distance threshold to be l, Suppose that the distance from a sample to a clustering

center is d. If d<l, label the sample.

(3) Cluster all the labeled samples with the fuzzy clustering method and compute the clustering centers and shape parameters.

(4) Train the classifier with all the samples in the database.

The RBF neural network designed in this way can learn both “positive” and “negative” samples to avoid mis-classification caused by inaccurate singular values.

The property of the learning ability of the RBF neural network is proved as follows.

For Parzen Kernel window function to estimate the probability density, the estimation is:

$$\hat{p}_{N_i}(x/\omega_i) = \frac{1}{N_i}\sum_{j=1}^{\omega_j} K(x - x_j^{(i)}) \tag{13}$$

where N_i is the number of persons, $K(\bullet)$ is the Parzen window function. Gaussian function is employed here, that is

$$K(x - x_j^{(i)}) = \frac{1}{(2\pi)^{n/2}\alpha(N_i)^{n/2}} \exp[-\frac{1}{2\alpha(N_i)}(x - x_j^{(i)})^T (x - x_j^{(i)})] \tag{14}$$

where $\alpha(N_i)$ is window length.

For RBF neural networks

$$y_i = \sum_{k=1}^{H} w_{ik} K(\frac{x - x_k}{\alpha}) = \sum_{x_k \in w_i} K(\frac{x - x_k}{\alpha}) + \sum_{x_k \notin w_i} w_{ik} K(\frac{x - x_k}{\alpha}) \tag{15}$$

where H is the number of hidden units. It can be found from the above equation that both the “positive” and “negative” samples can be learned.

In summary, in our face image verification system, face features are extracted by means of singular value decomposition and face verification by a RBF neural network classifier.

V. Experimental Results

We use the ORL database which contains a set of faces taken between April 1992 and April 1994 at the Olivetti Research Laboratory in Cambridge. There are 10 different images for each of the 40 distinct subjects. There are variations in facial expression (open/closed eyes, smiling/non-smiling), and facial details (glasses/no glasses). All images were taken against a dark homogeneous background with the subjects in an up-right, frontal position, with tolerance for some tilting and rotation of up to about 20 degrees. The images are greyscale with a resolution of 92×112. Some of the images are shown in Fig. 2.

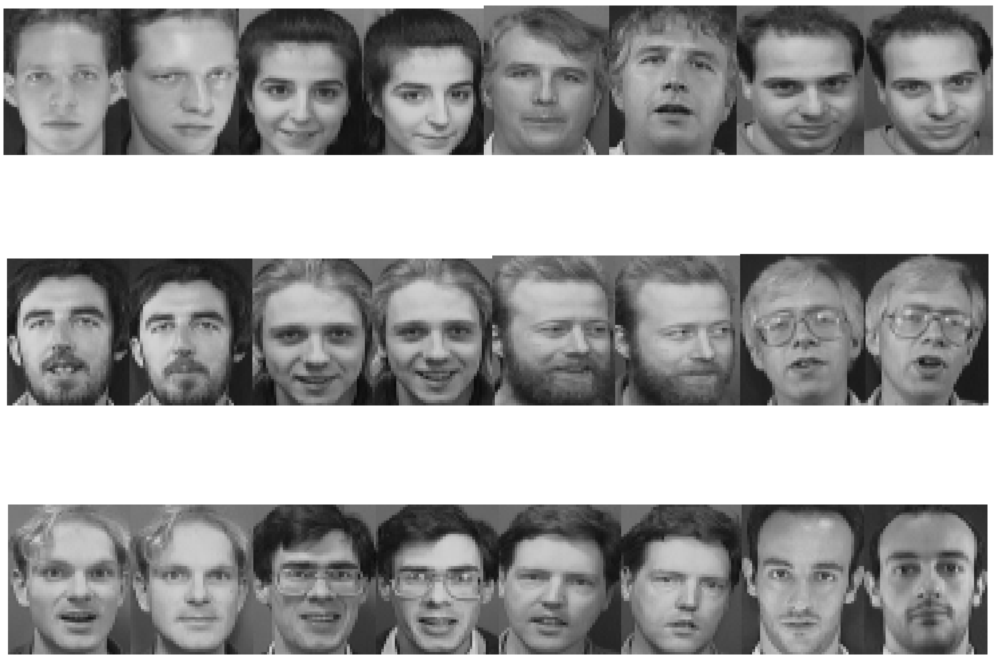

Fig. 2 Some face images of the ORL face database

Experiments were performed with 6 training images and 4 test images for each person. The average correct verification rate is about 80.9% for just using the singular values in distance matching (i.e., the Euchlidean distance classifier). The reject rate is 5.3%. Better results were achieved when applying singular values with the RBF neural network classifier. The recognition rate is 92% and the rejection rate is 4.2%. The improvement is mostly due to the RBF neural network’s learning ability. For example, the SVs of person s9 are very similar to those of person s38. In the process of training we can utilize the images of s38 as the "negative" samples for s9 while constructing the s9's classification plane. Some samples, which are difficult to classify, can be learned in advance. Supervised learning improves recognition accuracy. In addition to the high recognition accuracy, the proposed method is efficient in computation. We implemented the proposed face identification algorithm on a PII 400 PC computer. The identification of one face image takes less than one second (the algorithm is expected to run even faster after code optimization). The eigenface approach [15] to the same identifying problem takes much longer than the proposed method. The performances of the two methods are similar when we test with the ORL database.

VI. Conclusions and Future Work

In this paper we have proposed a SVD and RBF neural network based method for face verification. Real-time face identification is necessary in most practical applications. The proposed method can process face images (including training and identifying) in high speed and obtain good results. Its effectiveness and good performance has been proven by experiments.

The SVD based verification method achieves good results in the ORL database because the training and testing images have been normalized. The proposed identification technique improves the correct verification rates by exploiting the RBF neural network’s learning ability. Both the positive samples and negative samples can be learned in advance by the RBF neural network. As we know, in practice the face image database for each person is often small.

Small sample size is not sufficient for statistic or neural network recognition methods to form the judge plane to determine if the sample belongs to a known person or not [9]. The RBF neural network forms judge planes by using not only positive samples but also negative samples. The classic SVD method for object recognition suffers from the disadvantage of unsupervised learning and could not obtain high recognition accuracy. In the proposed method, both "positive" samples or "negative" samples can be learned in advance. So the SVD and RBF neural network based face identification method described in this paper overcomes the disadvantage of unsupervised learning in the exiting SVD based face recognition algorithms.

References:

[1] Jun Zhang, Yong Yan, and Martian Lades, "Face Recognition: Eigenface, Elastic Matching, and Neural Nets", Proceedings of the IEEE, Vol.85, No.9, 1997.

[2] Z. Hong, "Algebraic feature extraction of image for recognition", Pattern Recognition, Vol.24, pp211-219, 1991.

[3] Steve Lawrence, Lee Giles, Ah Chung Tsoi, Andrew D. Back, "Face recognition: A Convolutional Neural Network Approach", IEEE Trans. on Neural Networks, Special Issue on Neural Networks and Pattern Recognition, 1998.

[4] V.C. Klema, "The Singular Value Decomposition: Its Computation and Some Applications", IEEE Trans. Automatic Control, Vol. 25, pp164-176, 1980.

[5]Sami Romdhani, "Face recognition Using Principle Components Analysis", Master thesis, University of Glasgow.

[6]C. Nadal, R. Legault and C. Y. Suen, "Complementary Algorithms for the Recognition of

Totally Unconstrained Handwritten Numerals", Proc. 10th International conference on Pattern Recognition., Vol.A, pp434-449, 1990.

[7] Tin Kam Ho, Jonathan J. Hull, Sargur N. Srihari, "Decision Combination in Multiple Classifier Systems", IEEE Trans. PAMI, Vol.16, No.1, pp66-75, 1994.

[8] Roberto Brunelli and Daniele Falavigna, "Person Identification Using Multiple Cues", IEEE, Trans., PAMI, Vol.17, No.10, 1995.

[9] Sarunas J. Raudys and Anil K. Jain, "Small Sample Size Effects in Statistical Pattern Recognition: Recommendations for Practitioners", IEEE Trans., PAMI, Vol.13, No.3, 1991.

[10] Yong-Qing Cheng, ect., "Human face recognition method based on the statistical model of small sample size", SPIE Proc: Intell. Robots and Compu. Vision, Vol.1607, pp85-95, 1991.

[11] Y. Cheng, K. Liu, J. Yang, and H. Wang, 'A Robust algebraic method for human face recognition," in Proc. 11th Int., Conf. on Patt. Recog., pp221-224, 1992.

[12] F. Kimura and M. Shridhar, "Handwritten number recognition based on multiple algorithms", Pattern Recognition, Vol.24, No.10, pp969-983, 1991.

[13] B. Efron, "The efficiency of logistic regression compared to normal discriminate analysis," J. Amer. Statist. Ass., Vol.70, No.352, pp892-898, 1975.

[14] G. Strang, Linear Algebra and Its Applications, 3rd ed. Orlando, FL: Harcourt Brace Jovanovich, 1986.

[15] Matthew A., Turk and Alex P. Pentland, "Face Recognition Using Eigenfaces", ICPR, pp586-591, 1991