A Multiple Eigenspaces Constructing Method and Its Application to Face Recognition

Wu-Jun Li, Bin Luo, Chong-Jun Wang,
Xiang-Ping Zhong, and Zhao-Qian Chen

National Laboratory for Novel Software Technology,
Nanjing University, Nanjing 210093, P.R. China
{liwujun, chjwang}@ai.nju.edu.cn
{luobin, chenzq, chjwang }@nju.edu.cn

Abstract. The well-known eigenface method uses a single eigenspace to recognize faces. However, it is not enough to represent face images with large variations, such as illumination and pose variations. To overcome this disadvantage, many researchers have introduced multiple eigenspaces into face recognition field. But most of these methods require that both the number of eignspaces and dimensionality of the PCA subspaces are a priori given. In this paper, a novel self-organizing method to build multiple, low-dinensinal eigenspaces from a set of training images is proposed. By *eigenspace-growing* in terms of low-dimensional eigenspaces, it completes clustering images systematically and robustly. Then each cluster is used to construct an eigenspace. After all these eigenspaces have been grown, a selection procedure *eigenspace-selection* is used to select the ultimate resulting set of eigenspaces as an effective representation of the training images. Then based on these eigenspaces, a framework combined with neural network is used to complete face recognition under variable poses and the experimental result shows that our framework can complete face recognition with high performance.

1 Introduction

Face recognition is to identify or verify one or more persons in the given still or video images of a scene using a stored database of faces [1]. Due to various applications in the areas of pattern recognition, image processing, computer vision, and cognitive science and so on, face recognition has gained much attention in recent years. Existing approaches for face recognition can be classified into two categories [2]: geometric feature-based methods and appearance-based methods. The geometric feature-based methods, such as elastic bunch graph matching [3] and active appearance model [4], make use of the geometrical parameters that measure the facial parts; whereas the appearance -based methods use the intensity or intensity-derived parameters. As a representative of this approach, the eigenface approach [5] has become the benchmark of face recognition techniques.

The FERET evaluation [6], however, shows that the performance of a face recognition system may decrease seriously with the change of illumination condition or pose. To recognize faces under variable pose, Pentland et al [7] have proposed a multiple eigenspaces method, which builds view-specific eigenspaces. For

L. Wang, K. Chen, and Y.S. Ong (Eds.): ICNC 2005, LNCS 3611, pp. 55 – 64, 2005.

recognition of human faces with any view in a certain viewing angle range, Huang et al [8] have introduced neural network ensemble to multiple eigenspaces. To address the variable illumination issue, Li et al [9] [10] have proposed illumination invariant methods by building multiple eigenspaces in terms of illumination directions. Kim et al [11] proposed a mixture-of-eigenfaces method to recognize face images with pose and illumination variations. All these approaches with multiple eigenspaces outperformed those using a single eigenspace when the face images are of variations in illumination, pose, and expression and so on.

Nevertheless, all these methods require that both the number of eignspaces and dimensionality of the PCA subspaces are a priori given. But in real world, when imaging conditions are affected by a diversity of factors, it is unreasonable to manually divide the images into several groups and then construct an eigenspace for each. Leonardis et al [12] have proposed a self-organizing framework to construct multiple eigenspaces with two procedures called *eigenspace-growing* and *eigenspace-selection*, and tested it on a number of standard image sets with significant performance. In this paper, a novel self-organizing method to build multiple, low-dinensinal eigenspaces from a set of training images is proposed. Using the same terminology as [12], we divide the process into two procedures named *eigenspace-growing* and *eigenspace-selection*. Procedure *eigenspace-growing* of our method is similar to the corresponding one in [12]. Procedure *eigenspace-selection*, however, is novel which is simpler and more understandable. By *eigenspace-growing* in terms of low-dimensional eigenspaces, it completes clustering images systematically and robustly. Then each cluster is used to construct an eigenspace. After all these eigenspaces have been grown, a selection procedure *eigenspace-selection* is used to select the ultimate resulting set of eigenspaces as an effective representation of the training images. Subsequently based on these eigenspaces, a framework combined with neural network is used to complete face recognition. Experimental result indicates that the multiple eigenspaces constructing method of this paper is very effective and our face recognition framework can recognize face images of huge variations with a high recognition ratio.

The rest of this paper is organized as follows. In section 2, we present our self-organizing method to construct multiple eigenspaces. In section 3, we propose a face recognition framework based on neural network and multiple eigenspaces. Section 4 is the empirical study. Finally in section 5, we summarize the contributions of this paper and discuss future work.

2 Construct Multiple Eigenspaces

Given a set of training images, in order to divide them into several groups each of which can be effectively represented by a low-dimensional eigenspace, there are two parts of the problem that have to be solved: to decide the number of eigenspaces and the effective dimension number of each eigenspace. This paper introduces a novel self-organizing method to construct multiple, low-dinensinal eigenspaces from a set of training images systematically and robustly. The method is divided into two procedures: *eigenspace-growing* and *eigenspace-selection*. Procedure *eigenspace-growing* completes grouping images and constructs an eigenspace for each attained cluster while procedure *eigenspace-selection* is used to select a subset of these attained eigenspaces which can effectively represent the training images.

To describe our method conveniently, we first explain some notations. X is a given set of training images and the number of images in X is n, namely, $|X| = n$. $G_j^{(t)}$, $E_j^{(t)}$ and $p_j^{(t)}$ denote, respectively, the jth subset of X, the jth eigenspace, and the effective dimension number of the jth eigenspace, where the superscript ${}^{(t)}$ denotes that the *eigenspace-growing* procedure is at the tth iteration. δ_{ij} is the distance between the original image x_i and its reconstruction $\hat{x}_{ij}$ from the eigenspace j, $\delta_{ij} = \left\| x_i - \hat{x}_{ij} \right\|$. ρ_j is the reconstructive error of the jth eigenspace,

$$\rho_j^2 = \frac{1}{\left|G_j\right|} \sum_{x_i \in G_j} (x_i - \hat{x}_{ij})^T (x_i - \hat{x}_{ij}) = \frac{1}{\left|G_j\right|} \sum_{x_i \in G_j} \delta_{ij}^2 .$$

2.1 Seed Formation

To initiate the growing, a large number of initial subsets of images (seeds), denoted by $G_j^{(0)}$ (assume $|G_j^{(0)}| = k_j\,(k_j \ll n)$), are randomly selected from X and the number of seeds is also random but within a reasonable range according to n. The smallest effective dimension $p_j^{(0)} = 0$ which means that the seed's eigenspace $E_j^{(0)}$ is the mean image $\bar{x}_j = \frac{1}{\left|G_j^{(0)}\right|} \sum_{x_i \in G_j^{(0)}} x_i$.

Not all of the generated seeds are useful for the *eigenspace-growing* procedure. So before growing, some of them must be ruled out based on the comparison of the reconstructive error ρ_j with a predefined threshold value. The closer the images included in one seed are, the smaller ρ_j is, hence only seeds with similar images get small ρ_j and are accepted for further growth.

2.2 Eigenspace Growing

For every eigenspace $E_j^{(t)}$, procedure *eigenspace-growing* finds images in X which is compatible to it and adds the image into $G_j^{(t)}$. For every image not included in $G_j^{(t)}$, it is projected onto $E_j^{(t)}$ and the feature coefficients are attained. Then, we reconstruct it from the feature coefficients, and calculate the reconstructive error δ. After all the reconstructive errors are gotten, we sort all the images with respect to the error δ. If the lowest error is above a threshold value, it indicates that there is no image compatible to this eigenspace and the growth of this eigenspace is terminated. Otherwise, we temporarily add the image with the lowest error into $G_j^{(t)}$, and get the temporary $G_j^{(t+1)}$. Then, we construct the temporary eigenspace and calculate its corresponding reconstructive error $\rho_j^{(t+1)}$ based on current effective dimension number.

Then we can decide:

- If $\rho_j^{(t+1)}$ is below a threshold value, we accept current eigenspace as the permanent $E_j^{(t+1)}$ and use it as the base for the growth of next iteration.
- Otherwise, we increase current eigenspace's dimension by one and recalculate the error $\rho_j^{(t+1)}$:
  - ➢ If the error decreases significantly, we accept current eigenspace and current increased dimension with current set of images.
  - ➢ Otherwise, the temporarily added image is rejected, $G_j^{(t+1)} = G_j^{(t)}$, $E_j^{(t+1)} = E_j^{(t)}$, and the growth of this eigenspace is terminated.

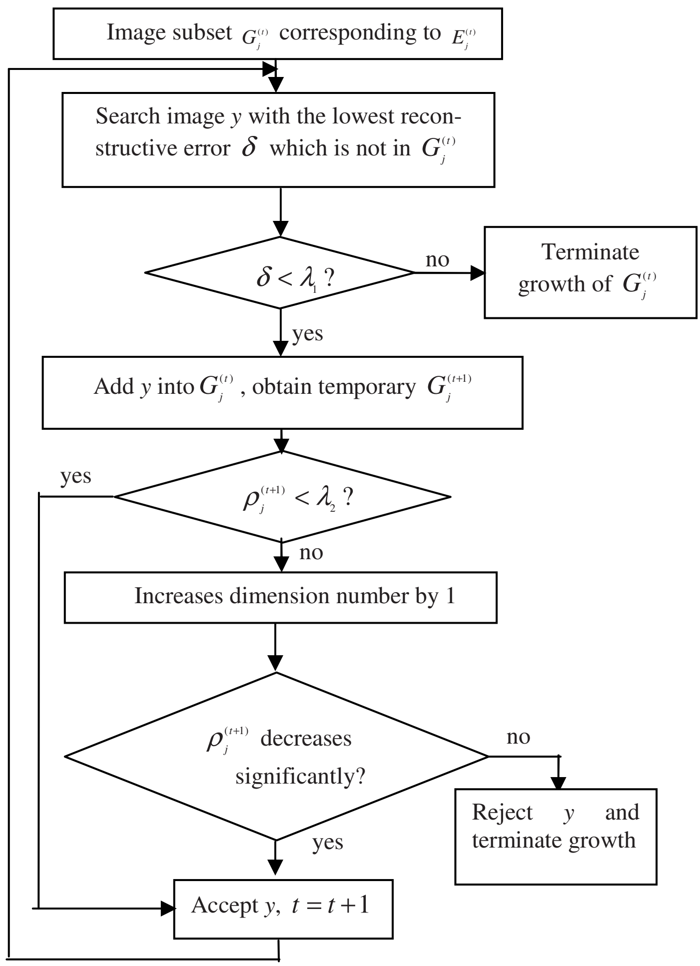

Fig. 1. Eigenspace growing for an individual eigenspace

This completes one iteration of the procedure *eigenspace-growing* which will terminate in a finite number of iterations since n is finite.The procedure *eigenspace-growing* is depicted in Fig.1.

2.3 Eigenspace Selection

Because the eigenspaces are initiated with a large number of random seeds and every seed grows independently, several of the eigenspaces are partially or even completely overlapped in terms of the constituting subsets of images. So certain redundant eigenspaces must be ruled out.

In order to select a set of eigenspaces with large variation in terms of the constituting subsets of images among them, we propose a method based on the idea: If most images in a subset corresponding to an eigenspace are included by the union of subsets corresponding to the other eigenspaces, that eigenspace is rejected. The *eigenspace-selection* procedure is described as follows:

Assume that ultimately k eigenspaces are constructed by procedure *eigenspace-growing*. We sort all of the eigenspaces in an ascending order with respect to the number of the images included in the subset corresponding to each eigenspace. We assume the sorted result is $E_1, E_2, E_3, \ldots, E_k$ and the corresponding subsets are $G_1, G_2, G_3, \ldots, G_k$. For i = 1 to k, let $G_{it} = [G_{i+1} \ldots G_k]$, where G_{it} is the union of $G_{i+1} \ldots G_k$. If G_i is a subset of G_{it} or there are only a small number of images in G_i not included by G_{it}, G_i is ruled out.

After *eigenspace-selection* procedure is completed, we can attain the necessary eigenspaces that effectively represent the training set.

3 Face Recognition Framework

If the number of eigenspaces is I, the framework for face recognition is shown in Fig.2. Extracting the feature vectors of each image in all I eigenspaces, we can get I eigenvectors. Then by combining all I eigenvectors one by one, we can get the combined eigenvector of the image which is used as the input of a BP-neural network with one hidden layer whose output points out the identity of the image.

At the test phase, given a probe image p, if it belongs to one of the objects in the gallery, the output should tell its identity, otherwise the output is *unrecognizable*. So the output is a binary vector in which one bit stands for a specific object. If the framework is designed to recognize k people, the output vector has $k+1$ bits in which the first k bits stand for k objects to recognize and the last bit stands for *unrecognizable* objects rejected by the framework (also called negative samples). If the identity of the face image is j, the jth bit of the output vector is *1* and the other k bits are *0*. And if the identity of the face image is *unrecognizable*, the $(k+1)$th bit is *1* and the other k bits are *0*.

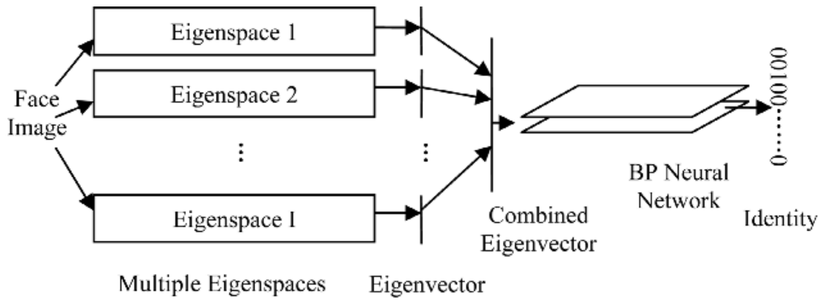

Fig. 2. Framework for face recognition

4 Experiment

To evaluate our framework, we apply it to recognize faces under variable poses. However, it is suitable for recognizing faces with more complex variation, such as arbitrary variations in pose, illumination, and expression and so on.

4.1 Data Acquisition and Preprocessing

In our experiment, we have collected indoor face images using a Logitech QuickCam® Pro 3000 camcorder. To collect face images with different poses, we let subjects sit in front of the camcorder and rotate their heads horizontally from the left side to the right side between ± 40 degrees so that in the face images both eyes are always visible. We collected face images of 15 subjects. For all the 15 subjects, we collected 60 images for each subject. Several face images with both eyes being visible in FERET face base [6] are added into the training set for constructing multiple eigenspaces.

To eliminate the effect of the non-face region variations on the recognition performance, we crop the face area from the whole image and perform recognition on the cropped face area. After locating the face and the eyes, we use the method in [8] to estimate the pose of the face in the image. Fig.3 shows a human head seen from above, the distance a between the projection of the mid-point of two eyes and the center of the face, and the radius of the head r (suppose the head has the same shape as a circle) can be attained through face and eye detection. Then we estimate the pose θ by $\theta = \arcsin(a/r)$. First we calculate the mid-point of the two eyes, then extend from the mid-point to the left side by $w(1-\sin\theta/2)/\cos\theta$, extend to top by $w/\cos\theta$, and crop a $3w/\cos\theta$ by $3w/\cos\theta$ area as the face image, then resize it to an image of 80 by 80 pixels, and eliminate its background with a mask of 80 by 80 pixels shown in Fig.4. Fig.5 is some cropped images with various poses from one of the persons in our face base. All training set and test set of this paper are the normalized images just like those shown in Fig.5.

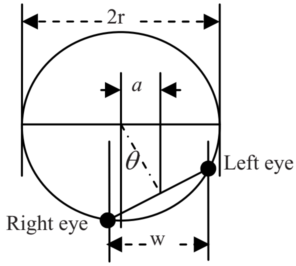

Fig. 3. Pose estimation

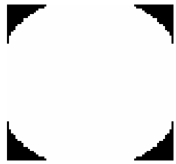

Fig. 4. Mask image

Fig. 5. The cropped face areas

4.2 Experimental Result

Our experiment is designed to recognize 5 subjects in our database, and use another 5 subjects as a "*rejection*" subject, i.e., negative examples. In the test phase, if the given image is from the first 5 subjects, the output will tell the identity of the subjects or else if the image is from the second 5 subjects, the system will simply reject the image, marking it as "*unrecognizable*".

As a comparison experiment, we also test the conventional method, which uses neural networks trained on face images with specific poses. In order to train these neural networks, we separate face images into different views using the method in [8]. We group the images into 7 sets: (-35 to -25), (-25 to -15), (-15 to -5), (-5 to +5), (+5 to +15), (+15 to +25), (+25 to +35), and label images falling into each set as -30, -20, -10, 0, +10, +20, and +30 degrees, respectively. That is to say, in this paper, when we say that one image is of 10 degrees, we actually mean that the pose of the face in the image is between 5 and 15 degrees.

Then we create an individual eigenspace with 20 dimensions for each different view by the standard eigenface method in [5], namely, we construct 7 eigenface sets for the -30, -20, -10, 0, +10, +20, and +30 degrees images, respectively. The average faces of seven eigenspaces are shown in Fig.6.

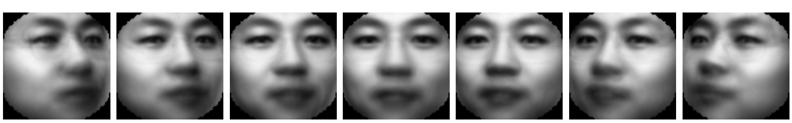

Fig. 6. Average faces of eigenspaces

Table 1. Recognition performance of conventional neural network

Test set / Neural network	-30	-20	-10	0	10	20	30	All
Num of images	18	26	18	34	19	25	21	161
-30 network	100%	88%	78%	62%	63%	72%	62%	74%
-20 network	94%	100%	94%	65%	63%	64%	67%	77%
-10 network	83%	85%	100%	82%	74%	64%	62%	78%
0 network	83%	69%	89%	97%	95%	72%	76%	83%
10 network	72%	77%	89%	85%	100%	92%	76%	84%
20 network	72%	69%	67%	74%	95%	100%	95%	81%
30 network	61%	50%	61%	53%	53%	84%	100%	65%

Table 2. Recognition performance of our framework

Test Set	-30	-20	-10	0	10	20	30	All
Num of images	18	26	18	34	19	25	21	161
Our framework	83%	92%	89%	94%	95%	92%	90%	91%

After the seven eigenspaces are constructed, we train seven back-propagation (BP) neural networks, with training data of -30, -20, -10, 0, +10, +20, and +30 degrees images, respectively. All these networks have 20 input units, 15 hidden units and 6 output units. The training set includes 10 subjects' 542 face images in all and the test set includes 161 face images of the same 10 subjects. The training set and the test set have no intersection. Each column in Table 1 stands for the test images of a specific view, and each row shows the recognition ratio of each neural network tested on images with different poses.

We applied our method presented in section 2 to construct multiple eigenspaces on the training set with various poses including 300 images from FERET face base [6] and 100 images collected by us. Through *eigenspace-growing* and *eigenspace-seletion*, six eigenspaces are finally constructed. The subsets of images corresponding to different eigenspaces have some images with the same pose. So the eigenspaces constructed by our method represent the training set with a globally consistent description. After the six eigenspaces are constructed, we use the same training set (542 images) to train the framework described in section 3 and test it on the same test set (161 images). The recognition result is shown in Table 2.

From Table 1, we can see that if there is an accurate pose estimator and the test image is fed to the right neural network, the recognition rate is above 99% on average, as shown by the diagonal line of Table 1. However, the recognition performance will decrease seriously if the pose estimation is not accurate enough. For example, if the images of -20 degree are fed to the 30 degree network, the recognition ratio is as low as 50%. Furthermore, in real world, the imaging condition is very complex, with arbitrary illumination, pose and expression variations and we do not know which one puts the most significant impact on one specific image. So we do not know how many eigenspaces should be constructed and what the dimension number should be to effectively represent the distribution of face images.

Nevertheless, in our framework, the pose (or other variations) estimation can be skipped. The input image with arbitrary pose can be fed into the framework directly to complete recognition. Moreover, in spite of whatever conditions under which the images are collected, the eigenspaces that represent the images effectively can be constructed in a self-organizing way. In the experiment of this paper, the recognition ratio (91%) of our framework is much higher than the average recognition ratio (77%) of the conventional method without pose estimation. This indicates that our face recognition framework combined with the multiple eigenspaces constructing method is very effective.

5 Conclusion

We have proposed a novel method to construct multiple eigenspaces from a set of training images and applied it to a face recognition framework combined with neural network. The experimental result indicates that our face recognition framework can complete recognition with a high accuracy.

There are many works left to do in the near future. Firstly, the face base is too small and the images in it only have pose variation. We will collect a large face base with more complex imaging conditions to evaluate the multiple eigenspaces constructing method and the face recognition framework. Secondly, the method proposed by Leonardis et al in [12] is not used to recognize faces and our method is also not tested on their image base. We plan to do comparisons with other methods on more data sets. Moreover, we will compare our face recognition framework with some other face recognition methods. Thirdly, the thresholds, such as λ_1, will affect the final resulting set of eigenspaces. We hope to find out the relationship between these thresholds and the ultimate result.

Acknowledgements

The research in this paper used the FERET database of face images collected under the FERET program. The National Natural Science Foundation of P.R.China under grant No. 60273033 and the Natural Science Foundation of Jiangsu Province of China under grant BK2003067 supported this research.

References

1. Zhao, W., Chellappa, R., Rosenfeld, A., Phillips, P. J.: Face Recognition: a Literature Survey. ACM Computing Surveys 35 (2003) 399-458
2. Brunelli, R., Poggio, T.: Face Recognition: Features Versus Templates. IEEE Transactions on Pattern Analysis and Machine Intelligence 15 (1993) 1042-1052
3. Laurenz, W., Jean-Marc, F., Norbert, K., Christoph, v. d. M.: Face Recognition by Elastic Bunch Graph Matching. IEEE Transactions on Pattern Analysis and Machine Intelligence19 (1997) 775-779

4. Edwards, G. J., Cootes, T. F., Taylor, C. J.: Face Recognition Using Active Appearance Models. In: Proceedings of the 5th European Conference on Computer Vision, vol. 2. Freeburg, Germany (1998) 581-595
5. Turk, M., Pentland, A.: Eigenfaces for Recognition. Journal of Cognitive Neuroscience 3 (1991) 71-86
6. Phillips, P. J., Wechsler, H., Huang, J., Rauss, P. J.: The FERET Database and Evaluation Procedure for Face-recognition Algorithms. Image and Vision Computing 16 (1998) 295-306
7. Pentland, A., Moghaddam, B., Starner, T.: View-based and Modular Eigenspaces for Face Recognition. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Seattle, WA (1994) 21-23
8. Huang, F. J., Zhou, Z.-H., Zhang, H.-J., Chen, T.: Pose Invariant Face Recognition. In: Proceedings of the 4th IEEE International Conference on Automatic Face and Gesture Recognition, Grenoble, France (2000) 245-250
9. Li, W.-J., Wang, C.-J., Xu, D.-X., Chen, S.-F.: Illumination Invariant Face Recognition Based on Neural Network Ensemble. In: Proceedings of the 16th IEEE International Conference on Tools with Artificial Intelligence (ICTAI 2004), Boca Raton, Florida (2004) 486-490
10. Li, W.-J., Wang, C.-J., Xu, D.-X., Luo, B., Chen, Z.-Q.: A Study on Illumination Invariant Face Recognition Methods Based on Multiple Eigenspaces. Lecture Notes in Computer Science, Vol. 3497. Chongqing, China (2005) 131-136
11. Kim, H.-C., Kim, D., Bang, S. Y.: Face Recognition Using the Mixture-of-eigenfaces Method. Pattern Recognition Letters 23 (2002) 1549-1558
12. Leonardis, A., Bischof, H., Maver, J.: Multiple Eigenspaces. Pattern Recognition 35 (2002) 2613-2627