A Study on Illumination Invariant Face Recognition Methods Based on Multiple Eigenspaces

Wu-Jun Li[1], Chong-Jun Wang[1], Dian-Xiang Xu[2], Bin Luo[1], and Zhao-Qian Chen[1]

[1]National Laboratory for Novel Software Technology
Nanjing University, Nanjing 210093, P.R.China
Email: {liwujun, chjwang}@ai.nju.edu.cn
[2]Department of Computer Science
North Dakota State University, Fargo, ND58105, USA
Email: Dianxiang.xu@ndsu.nodak.edu

Abstract. It is a challenge to recognize faces under variable poses or illumination directions. In the area of multiview face recognition, many experimental results have shown that the performance of approaches based on multiple eigenspaces is higher than the performance of those based on a single eigenspace. This paper presents two multiple illumination eigenspaces-based methods, RDEB and BPNNB, for solving the variable illumination problem of face recognition. The experiment shows that the methods have a high recognition ratio. In particular, BPNNB has outperformed the assumptive method which knows the illumination directions of faces and completes recognition in the specific eigenspace using eigenface method[2] for each face subset with a specific illumination direction.

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]. With the change of illumination condition or pose, the performance of a face recognition system may decrease significantly [3]. To recognize faces under variable poses, Pentland et al [4] have proposed a multiple view-specific eigenspaces method. For a given probe face image, it first determines the orientation of the target object by selecting the eigenspace which best describes the input image according to the residual description error (the “distance-from-face-space” metric [2]). Then recognition is done on the properly selected viewspace.

It is known that lighting changes impose a greater impact on image variation than different personal identity. To address the variable illumination issue, we have developed a multiple illumination eigenspaces method (called NNEB) [5]. Based on NNEB, this paper proposes a residual description error based method, called RDEB, and a Back-Propagation neural network based method, called BPNNB. These methods, evaluated with the IAE face database [5], have demonstrated a high recognition ratio. In particular, the BPNNB method outperformed an assumptive method which knows the illumination directions of faces and completes recognition in the specific eigenspace using eigenface method [2] for each face subset with a specific illumination direction.

The rest of this paper is organized as follows. In section 2, we briefly introduce multiple illumination eigenspaces and the NNEB-method. Section 3 presents the RDEB method. Section 4 describes the BPNNB method. Section 5 is the empirical study. Section 6 concludes the paper.

2 Multiple Illumination Eigenspaces and Neural Network Ensemble Based (NNEB) Method

Assume that the face images in a training set, say *X*, can be divided into a number of subsets, say $\{X1,X2,...,XI\}$, according to the illumination directions when images are imaged, i.e. $X=\{X1,X2,...,XI\}$, where $Xi=\{x_{i1},x_{i2},...,x_{in_i}\}$, *I* is the number of different illumination directions, x_{ij} denotes the *jth* image corresponding to the *ith* illumination direction, and n_i is the number of face images corresponding to the *ith* illumination direction. We can build *I* eigenspaces [2], each describing a region of the face space that is corresponding to a specific imaging illumination direction. Thus, we can get the corresponding eigenspace set, $S=\{S_1,S_2,...,S_I\}$, where S_i stands for an eigenspace corresponding to the *ith* illumination direction. If $I > 1$, we have multiple illumination eigenspaces.

The NNEB method [5] extracts the feature vector of each image in the corresponding eigenspace. Illumination direction-specific backpropagation (BP) neural networks with one hidden layer are trained on the extracted feature vectors. Decision making is done by an ensemble module (Figure 1) that

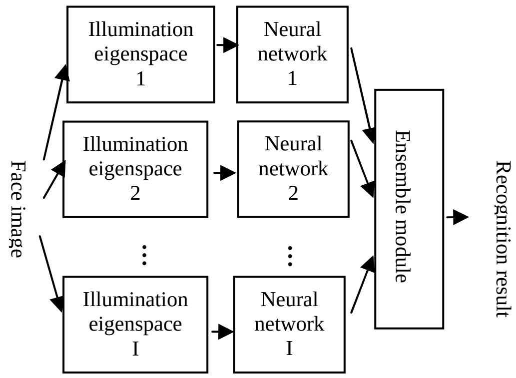

Fig.1. Neural Network Ensemble Architecture

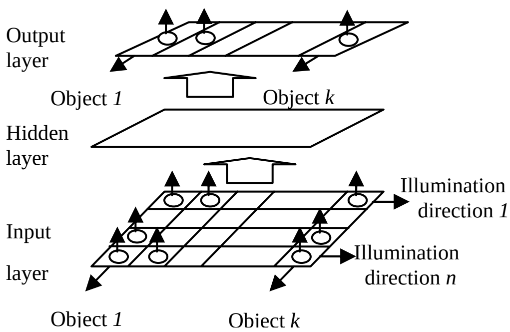

Fig.2. The Second-layer Neural Network

combines the illumination direction-specific neural networks (called first layer neural networks, which are trained on the feature vectors of the corresponding images extracted in the corresponding eigenspaces). At the test phase, given one input image, we feed it into these different channels and obtain a final decision from the ensemble module. The ensemble module is another BP- neural network called the second layer neural network, as shown in Figure 2. The input vector is the cascade of all the output vectors of the first layer. Suppose *n* eigenspaces are used, and the number of output units of each neural network in the first layer is *k* (each output unit stand for an object to be recognized). Thus, there are *q* ($q=n*k$) input units for the second layer. If these *q* input units are arranged into an $n*k$ array, each column of the array will stand for a specific object to recognize, whereas each row of the array will stand for a specific illumination direction. The second layer is used as a classifier that has *k* output units, each of which stands for an object. At the training phase, we force the output unit corresponding to the correct person to be “1” and all the other units’ values to be “0”. The first layer neural networks are trained as classifiers and used as regression estimators.

3 The Residual Description Error Based (RDEB) Method

The RDEB-method is an extension to the multiple eigenspaces method in [4]. First, multiple illumination eigenspaces are trained as described in section 2. Second, the illumination direction of the target object is determined by selecting the eigenspace which best describes the input image. This is accomplished by evaluating the residual description error [2] using each illumination space's eigenvectors. Once the proper illumination direction is determined, the face image is projected into that eigenspace, and then recognized. The RDEB-method is outlined as follows:

(1) Build a set of eigenspaces: If the number of illumination directions is *I*, we have a set of *I* illumination eigenspaces denoted as $S=\{S_1,S_2,...,S_I\}$.

(2) Extract the set of feature vectors of the face images in the gallery: Each image in the gallery $G=\{G1,G2,...,GI\}$ (Gi stands for the subset of gallery with an illumination direction *i*) is projected into the corresponding eigenspace of a specific illumination direction. So we can extract the set of feature vectors denoted as $C=\{C_{g1},C_{g2},...,C_{gI}\}$, in which $C_{gi}=\{c_{gi_1},c_{gi_2},...,c_{gi_{ni}}\}$ is a subset of C with an illumination direction *i* and c_{gi_j} is the feature vector of the *jth* face image with an illumination direction *i* by projecting into the specific eigenspace and *ni* is the number of images in Gi.

(3) Compute the feature vectors of the probe face images: Given a probe face image (say *p*), we project it into every illumination eigenspace and can get *I* feature vectors ($c_{pi}=S_i^T \times p,(i=1,2,...,I)$). Then the eigenspace corresponding to *p* is determined by computing the ratio of Signal-to-Noise of the rebuilt image. For any eigenspace $S_i(i=1,2,...,I)$, the rebuilt image is denoted as $p_i=S_i \times c_{pi}$, then the ratio of Signal-to-Noise corresponding to that eigenspace is $R_i=10\lg(\frac{||p||^2}{||p-p_i||^2})$. If $R_j=\max\{R_i|i=1,2,...,I\}$, the eigenspace corresponding to *p* is S_j and the vector $c_{pj}=S_j^T \times p$ is the feature vector of *p*.

(4) Complete recognition: For any element in set C_{gj}, if $dis(c_{pj},c_{gj_k})=\min\{dis(c_{pj},c_{gj_i})\,|\,i=1,2,...,ni\}$ ($dis(x,y)$ stands for the distance between vector *x* and *y*), the probe face image *p* and gj_k in gallery are considered as imaging from the same object. The measure of distance in this paper is Euclidean distance.

4 The Back-Propagation Neural Network Based (BPNNB) Method

The architecture of BPNNB-method is shown in Figure 3. If the number of illumination directions of the face images is *I*, we can build *I* illumination eigenspaces. 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.

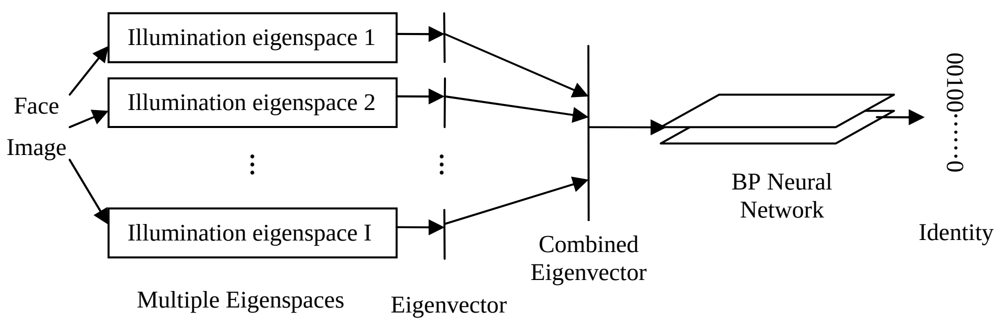

Fig.3. Architecture Base on Back-Propagation Neural Network

At the test phase, given a probe image *p*, if it belongs to one of the object 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 architecture 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 architecture (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*.

5 Experiment

While a number of benchmark face bases for face recognition, such as FERET database [3] and PIE database [7], are available, none of them are suitable for face recognition methods that are based on multiple illumination eigenspaces and neural networks, due to the need of multiple images of each object with each specific illumination direction. To evaluate our approaches, we have built the IAE (Illumination And Expression.) face base [5], a collection of 4320 face images of 27 subjects - 20 men and 7 women - among whom 15 persons wore glasses. There are four illumination directions to image the images. We call the front of a face 90’ direction, the left of a face 0’ direction. Accordingly, there are 30’ and 60’ directions in the left front of a face. For all the 27 subjects, there are 40 images for every subject in each illumination direction, so there are 4320 images in total. All images are color with 352*288 pixels in size. For the experiment of this paper, all color images were transformed into gray level images. Sample images in IAE face base are shown in Figure 4 (if the colors of the images were visible, you would tell that the illumination directions were obviously different). The rows correspond to the directions of 0’, 30’, 60’, and 90’, respectively.

Fig.4. Samples in IAE Face Base

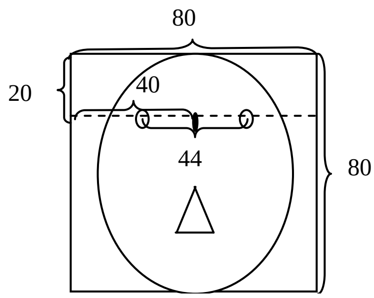

Fig.5. Standard Face Image

The face images were normalized before they were presented to the system for training or testing. The normalized version of a face image satisfies such constraints that the face could be appropriately cropped. The constraints include that the size of the image was fixed, the line between the two eyes was parallel to the horizontal axis, and the inter-ocular distance (distance between the two eyes) was set to a fixed value. The normalized face images in this paper satisfy the constraints in Figure 5. We conducted face detection using the method in [6] and then rotated and resized the row images using the bilinear method. After that, face images were cropped to satisfy the constraints in Figure 5. In the end, the appropriately cropped face images were histogram equalized. Hence all the training and testing sets in this paper were normalized and histogram equalized gray level face images.

The experiment was intended to recognize 10 subjects in the IAE face base and use another 10 subjects as a “*rejection*” subject, i.e., negative samples. Four eigenspaces, each corresponding to an illumination direction, were built. The training set for each eigenspace consisted of 270 face images among which each of the 27 persons contributed 10 images. Then we trained 4 BP-neural networks with one hidden layer, each corresponding to an illumination direction. The number of input units of each neural network was equal to the number of the dimensions of the corresponding eigenspace. Each neural network had 11 output units among which the first 10 units were corresponding to 10 subjects to recognize and the last one was corresponding to the 10 subjects to reject, i.e., the negative samples. At the test phase, if a given face image was from the first 10 subjects, the system would tell the identity of the subject; if the image was from another 10 subjects to reject, the system would reject the image and mark it as "*unrecognizable*". The training set for each neural network was composed of 300 images with a corresponding illumination direction. Among the 300 images, each of the 10 persons to be recognized contributed 20 images, and each of the 10 persons to be rejected contributed 10 images. None of the 1200 (300*4) images have been used to build eigenspaces.

The training set (gallery images) for the architecture of neural network ensemble, RDEB-method and BPNNB-method comprised of all the 1200 images used to train the first layer neural networks. The testing set (probe images) consisted of 800 images, 200 images from each illumination direction. Among each 200 images, each of the 20 subjects contributed 10 images. None of the 800 images was in the training set.

Taking different numbers (i.e. 10, 20, 30, 40 and 50) of dimensions for each eigenspace, we tested all the three method mentioned above (NNEB-method, RDEB-method and BPNNB-method) using the testing set with unknown illumination directions. For comparison, we also tested an assumptive method which knows the illumination directions of faces and completes recognition in the specific eigenspace using eigenface method [2] for each face subset with a specific illumination direction.

Note that the parameters of neural networks have a great effect on the performance of neural networks and the number of units in the hidden layer is the most important one among all the parameters. So our experiment only took the number of units in the hidden layer into account. This number was adjusted several times and the best result was selected as the final evaluation of recognition performance.

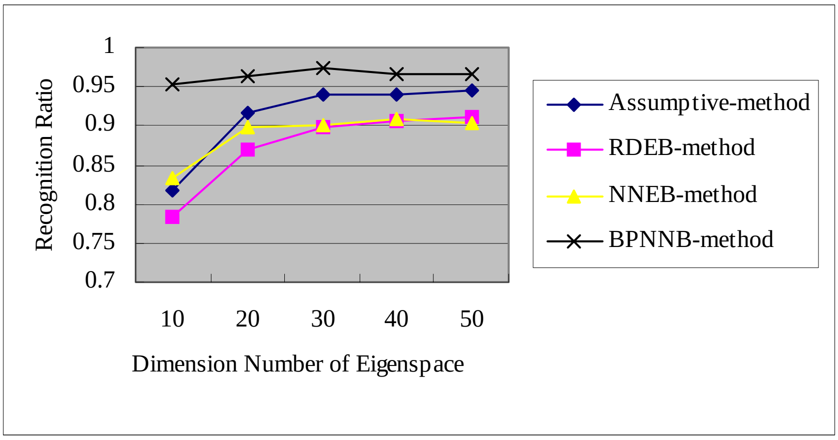

Fig.6. Recognition Performance

The experimental result is shown in Figure 6, which shows that all the three methods based on multiple eigenspaces achieved high recognition performance. As the number of dimensions of eigenspace increases, the performance mostly increases. The performance of NNEB-method and BPNNB-method increases with a bigger acceleration when the number of dimension of eigenspace is under 20 compared to over 20. While the threshold of RDEB-method and the Assumptive-method is 30. So the methods using neural network can obtain features of more discriminative ability with a smaller dimension number. Moreover, BPNNB-method proposed in this paper achieved much higher performance than other methods, even compared to the Assumptive-method, even if the number of dimension of eigenspace is small (near 10). We have also applied BPNNB-method to address the problem caused by variable poses (Face base and experimental method are analogous to [8]), and achieved analogous experimental result.

The absolute recognition ratio is not very high and the highest recognition ratio of BPNNB is 97.4%. It is mainly because, during the process of image collection, the heads of subjects were not fixed and they might sway slightly, so some of the images were blurry when collected by a camera. Meanwhile, the sway of the head has caused the variety of pose and illumination intensity.

6 Conclusions

We have presented two new face recognition methods based on multiple eigenspaces. They can complete recognition with high recognition accuracy. In the real world, however, illumination is complex. The illumination directions and the illumination intensity are essentially arbitrary. This is a major issue of future research. Moreover, a practical face recognition system should address both illumination and pose problems at the same time. A possible approach is to construct multiple eigenspaces from a large set of training images in a self-organizing way [9].

Acknowledgements

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., and Phillips, P. J. Face recognition: a literature survey. CVL Technical Report, University of Maryland, 2000.
2. Turk, M., and Pentland, A. Eigenfaces for recognition. Journal of cognitive neuroscience, 3(1): 71-86, 1991.
3. Phillips, P. J., Wechsler, H., Huang, J., and Rauss, P. J. The FERET database and evaluation procedure for face-recognition algorithms. Image and Vision Computing, 16(5): 295-306, 1998.
4. Pentland, A., Moghaddam, B., and Starner, T. View-based and modular eigenspaces for face recognition. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Seattle, WA, 21-23, 1994.
5. Li, W-J., Wang, C-J., Xu, D-X., and Chen, S-F. Illumination Invariant Face Recognition Based on Neural Network Ensemble. In Proceedings of the 16^{th} IEEE International Conference on Tools with Artificial Intelligence (ICTAI 2004), pp.486-490, Boca Raton, Florida, 2004. IEEE computer society.
6. Wu, J., Zhou, Z-H. Efficient Face Candidates Selector for Face Detection, Pattern Recognition, 36(5): 1175-1186, 2003.
7. Terence, S., Simon, B., and Maan, B. The CMU pose, illumination, and expression (PIE) database. In Proceedings of the Fifth International Conference on Face and Gesture Recognition, 2002.
8. Huang, F. J., Zhou, Z-H., Zhang, H-J., and Chen, T. Pose invariant face recognition. In Proceedings of the 4th IEEE International Conference on Automatic Face and Gesture Recognition, 245-250, France, 2000.
9. Leonardis, A., Bischof, H., and Maver, J. Multiple eigenspaces. Pattern Recognition, (35): 2613-2627, 2002.