PRACTICAL FACE RECOGNITION SYSTEM USING EIGENFACE ALGORITHM*

CHEN Gang QI Fei-Hu
(Department of Computer Science & Engineering, Shanghai Jiaotong University, Shanghai 200030, China)

Abstract A method taking multi-samples as sub-modes and grouping face modes into partial intersection ones was proposed to reduce computation and improve system extension property. In combination, the sum rule based on Bayesian theory was used. The face recognition experiments with the ORL and AR face databases showed that the eigenface algorithm using multi-samples reached a high recognition rate and a reasonable time cost. Grouping face modes makes training become a distributed computation job which reduces time cost for training and brings the convenience of system extension when new face modes are to be added.
Key words eigenface, face recognition, pattern recognition.

实用人脸识别系统的本征脸法实现*

陈 刚 戚飞虎
(上海交通大学计算机科学与工程系, 上海, 200030)

摘要 将人脸模式的多个样本作为子模式, 并将较多的人脸模式部分相交地分组, 以减小计算量和便于识别系统的扩展. 在结合方法上采用基于贝叶斯理论的和结合规则. 使用 ORL 和 AR 人脸图像库的人脸识别实验表明, 本征脸法在采用多样本训练后, 获得了较高的识别率和较短的识别时间; 分组训练使识别系统可采用分布并行计算, 从而减少训练时间, 同时使系统在增加新人脸模式时便于扩展.
关键词 本征脸法, 人脸识别, 模式识别.

Introduction

Human face is the most common pattern in vision. Face recognition[1], which can be directly applied in many fields such as credit card, driver license, passport and personal identification, is one of the important applications of pattern recognition theory.

Traditional face recognition methods recognize faces by matching geometrical features[2], such as the length of mouth. The recognition rate is low because human face isn't a rigid object but has complex expressions. In recent years, some new algorithms based on whole image information have been introduced. Elastic matching[3] that is the improvement of templates matching has high recognition rate but is very slow. Artificial neural nets[4, 5] are paid attention to because of their abilities of learning and robustness. Using algebraic features, eigenface[3, 6] is simple and easy to be realized.

Making use of whole information of images, eigenface takes image as matrix and uses eigenvalues of matrix and corresponding eigenvectros to recognize faces. However, the method trained by one sample per person has a recognition rate of about 80% and the computation increases sharply as the number of face modes increases. In this article, we first introduce eigenface algorithm briefly, then take multi-samples of a person as sub-modes in training to get a high recognition rate; group face modes into partial intersection ones to reduce computation and improve system extension proper-

* 国防预研基金(96. J2. 4. 2)资助项目
稿件收到日期 1998-09-07, 修改稿收到日期 1998-12-17

* The project supported by the Preliminary Research Foundation of National Defense(No. 96. J2. 4. 2).
Received 1998-09-07, revised 1998-12-17

ty; in combination, the sum rule based on Bayesian theory[7] is used. Finally, experiments using ORL and AR face databases[8] are carried out and conclusion is drawn in Section 4.

1 Eigenface

Suppose there are N images in face database and they are represented by vectors $X_1, X_2, \ldots, X_N$ (L-dimensional vectors). So face average image is $X_{ave} = \frac{1}{N}\sum_{i=1}^{N} X_i$ and mean deviation images can be written as:

$$X'_i = X_i - X_{ave} \quad i = 1, 2, \ldots, N \tag{1}$$

Then covariance matric C can be computed;

$$C = \frac{1}{N}\sum_{i=1}^{N} X'_i (X'_i)^T. \tag{2}$$

The eigenvalue λ_k and corresponding eigenvectors u_k of matrix C thus can be computed. These eigenvectors form a vector space which can represent the feature information of face images. All the face images in the database are projected onto the vector space to from respective vectors $Y_1, Y_2, \ldots, Y_N$,

$$(Y_i)^T = [y_{i1} y_{i2} \ldots y_{iL}], i = 1, 2, \ldots N:$$

$$y_{ij} = (u_j)^T X'_i, \ j = 1, 2, \ldots L \tag{3}$$

For an unknown face image X, the projection vector Y of the difference between X and X_{ave} is:

$$y_j = (u_j)^T (X - X_{ave}), j = 1, 2, \ldots L$$

Then vector Y is compared with all projection vectors $Y_1, Y_2, \ldots, Y_N$ and recognition is achieved by some distance rules. For example, Euclidean distance $e_i = \| Y - Y_i \|$ $(i = 1, 2, \ldots, N)$ is often used. The image is recognized as the n-th mode:

$$n = \arg(\min_{i=1,2,\ldots,N}(e_i)). \tag{4}$$

The size of matrix C is $L \times L$ which is very large even for a small size image. For example, a 24×28 size image will make matrix C to be $(24 \times 28) \approx 4.5 \times 10^5$ size. So reduction method is often used.

First, mean deviation images are piled into a matrix:

$$X' = [X'_1, X'_2, \ldots, X'_N] \tag{5}$$

Then covariance matrix C can be written as:

$$C = \frac{1}{N} X' (X')^T. \tag{6}$$

Thus by linear algebra theory, computation of eigenface λ_k and corresponding eigenvectors u_k of matrix $X'(X')^T$ can be translated into computation of eigenface λ_k and corresponding eigenvectors v_k of matrix $(X')^T X'$. The size of $(X')^T X'$ is $N \times N$ which is far less than $L \times L$, in this way, computation is reduced. After v_k is known, u_k can be computed.

$$u_k = \frac{1}{\sqrt{\lambda_k}} X' v_k, \tag{7}$$

2 Recognition System

A general face recognition system is shown in Fig. 1.

Face recognition system consists of five functional parts: training, location, preprocessing, feature extracting and recognition. The training module trains face images in database to get parameters for recognition. It's the most important part of all modules and relies greatly on algorithms. The location module locates face in an image for further processing. Under certain photographic conditions, this step can be left out. The preprocessing module preprocesses face images, such as size normalization and noise filter. Different methods are chosen in applications as needed. The feature extracting module extracts features of

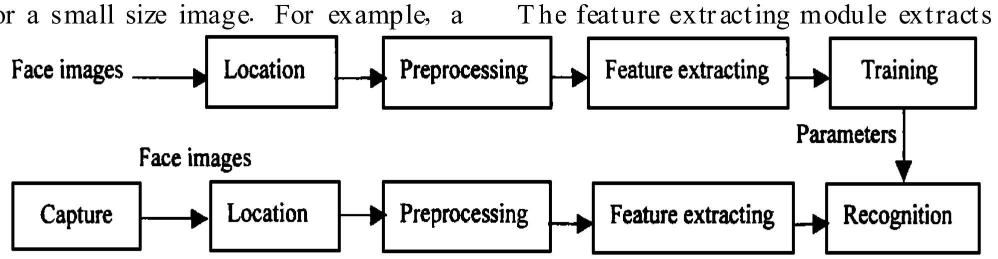

Fig. 1 The diagram of face recognition system
图 1 人脸识别系统框图

face images. How to get stable and effective features is one of the key points in face recognition. The recognition module recognizes unknown faces using the parameters trained. In this paper we focus our discussion on training, feature extracting and recognition modules and assume that others have already been done.

The performance of a face recognition system can be described by the number of persons, recognition rate, refusal rate and time cost. Recognition rate is the probability of correct recognition of some person in the face database. Refusal rate is the probability of correct refusal of some person not in the face database. Generally, recognition rate is used. In a system, recognition rate and refusal rate are inconsistent and compromise should be made between them. A practical system always reaches some performance under certain conditions; the more conditions (that also means less variance), the higher the recognition rate becomes. We study recognition of face images under simple background with variance (such as illumination, expression and head pose) and compare recognition systems with or without refusal. The following is the consideration of realization of a practical recognition system:

(1) A practical system should be simple to be realized. That's why eigenface algorithm is chosen.

(2) A practical system should have a high recognition rate. Eigenface trained by one sample per person has a recognition rate of about 80% that is still low. So multi-samples are adopted: multi-samples of a same person are taken as sub-modes in training. An unknown face image is recognized as the mode if the face image is considered to be one of the mode's sub-samples.

(3) A practical system should have the convenience of extension when new face modes are to be added. We get this property by grouping the face modes. When new modes are added, the most of old groups can be used and only a few new groups should be trained. The groups are independent of each other and can be trained in distributed computers, which reduces training time cost greatly. Grouping also reduces the size of covariance matrix and makes it easy to compute eigenvalues and corresponding eigenvectors.

Fig. 2 A grouping method without intersection
图 2　无相交的分组方案

One of the combination methods puts face modes into m-groups without intersection, see Fig. 2. After training, for an unknown face, parameters of m-groups are used to compute distance e_i and the face is recognized as the n-th mode:

$$n = \arg_i(\min_i(e_i)).$$

We proposed a combination method which puts face modes into m-groups with partial intersection, see Fig. 3. After training, for an unknown face, Bayesian sum-rule combination of intersection parts and parameters of m-groups are used to compute distance e_i and the face is recognized as the n-th mode:

$$P(X_i/EC_j) = \begin{cases} 0 & X_i \notin EC_j \\ \dfrac{1/e_i^2}{\sum_j (1/e_j^2)} & X_i \in EC_j \end{cases}, \quad (8)$$

$$n = \arg_i(\max_i \sum_j P(X_i | EC_j)) \quad (9)$$

From Figs. 2 and 3, we know that both methods have the convenience of extension. The groups are independent of each other, so distributed computation can be applied to reduce training time cost.

3 Experiments and Discussion

The ORL (the Olivetti and Oracle Research Lab's face image database) and AR (the AR face database) databases are used in the experiments.

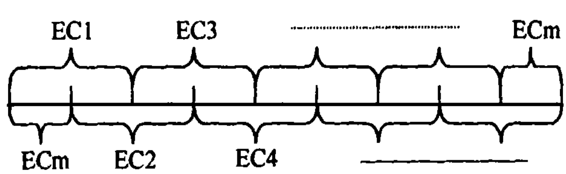

Fig. 3 A grouping method with partial intersection
图 3　部分相交的分组方案

Fig. 4 Three persons' face images in ORL database
图 4 ORL 人脸图像库中 3 个人的脸像

Fig. 5 Three persons' face images in AR database
图 5 AR 人脸图像库中 3 个人的脸像

40 persons in ORL and 79 persons in AR face database are selected to form a new database of 119 persons with 9 samples per person, in which 100 persons are used for recognition and 19 persons for refusal. The samples of a same person are taken at differenet times, varying the lighting, head position, facial expressions (open/closed eyes, smiling/serious) and facial details (glasses/no glasses). The ORL images are gray ones with 256 levles and their size is 92×112 (width×height) in pixel. The face images are resized to 24×28 (width ×height) by bilinear interpolation method to reduce computation. Three persons' face images in ORL are listed in Fig. 4. The AR images are 24-bits, 768×576 size color images. The images are first grayed, segmented and then resized to 24×28 by bilinear interpolation method. Figure 5 shows three persons' images in AR database.

To demonstrate the extension property, experiments are carried out for 60, 80 and 100 face

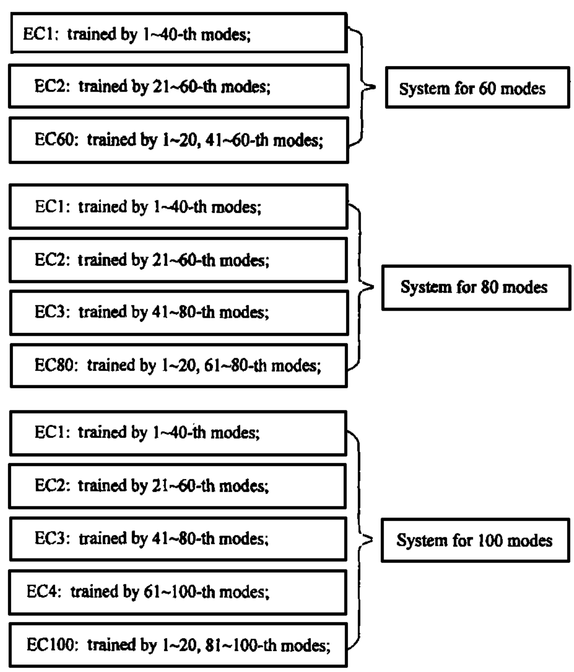

Fig. 6 Sample of system extension
图 6 系统扩展示例

modes(1 sample per person and 3 samples per person), see Fig. 6. Systems with refusal and without refusal are also compared. For comparison of recognition rate, refusal rate is required to be greater than 99%. Table 1 shows recognition rates of each eigenface classifiers EC_i trained by grouped images.

The test images include the face images for training and so are the followings. Table 1 shows that recognition rate of eigenface trained by one sample per person is much lower than that trained by three samples. So using multi-samples as submodes in training is effective. Then applying the two methods the groups are combined and Table 2 gives the result.

From Table 2, we know that our partial intersection method is superior to that without intersection with or without refusal. Its recognition rate with refusal, which is often required in practical systems, is much higher than the latter. This benefits from partial intersection and Bayesian sum-rule combination, which improves the system robustness.

The experiments used a computer PⅡ400 and Matlab 5.2 interpreter. The average time cost of eigenface is 0.26s. The average time cost for training one group, such as EC1, is 48.6s. If a compiler like C language is used, the time cost could be further reduced. Figure 7 shows some sample eigenfaces got in the program.

Table 1 Recognition result of each group
表 1　分组识别结果

Group	EC0 (1～20)	EC1 (1～40)	EC2 (21～60)	EC3 (41～80)	EC4 (61～100)	EC60 (1～20, 41～60)	EC80 (1～20, 61～80)	EC100 (1～20, 81～100)
Recognition rate (one sample, without refusal)	77.8%	78.05%	76.68%	79.18%	74.18%	75.82%	79.73%	77.77%
Recognition rate (three samples, without refusal)	93.9%	93.6%	90.82%	93.9%	93.6%	91.95%	95.27%	95.27%
Recognition rate (three samples, with refusal rate > 99%)	86.1%	85%	82.77%	91.95%	93.6%	80%	90.28%	89.18%

Table 2 Recognition result after groups combined
表 2　分组结合后的识别结果

System	60 modes		80 modes		100 modes	
	Without intersection	Partial intersection	Without intersection	Partial intersection	Without intersection	Partial intersection
Groups used	EC0, EC2	EC1, EC2, EC60	EC1, EC3	EC1, EC2, EC3, EC80	EC0, EC2, EC4	EC1, EC2, EC3, EC4, EC100
Recognition rate (three samples, without refusal)	87.22%	90.73%	91.25%	92.09%	89.67%	91.67%
Recognition rate (three samples, with refusal rate > 99%)	77.78%	90.55%	85.41%	91.95%	84.11%	91.33%

Fig. 7 Eigenface samples
图7 本征脸示例

4 Conclusion

Taking image as matrix, eigenface algorithm uses eigenvalues and corresponding eigenvectors in recognition. The method has advantage of not needing geometric features of eyes, noses and mouths, but doesn't reach high recognition rate when single sample image per person is used for training. Another problem is that the larger the number of face modes is, the more complex the computation becomes. So we proposed a method taking multi-samples as sub-modes and grouped face modes into partial intersection ones to reduce computation and improve system extension property. In combination, we used the sum rule based on Bayesian theory. The face recognition experiments with the ORL and AR face databases showed that:

(1) Eigenface using multi-samples reaches a high recognition rate and a reasonable time cost;

(2) Partial intersection grouping and Bayesian sum-rule combination improve system robustness, especially for a system with refusal;

(3) Grouping makes training become a distributed computation job which reduces time cost for training;

(4) When new face modes are to be added, the system has the convenience of extension and still reaches a high recognition rate.

REFERENCES

[1] Wechsler H, Phillips P J, Bruce V, et al. Face Recognition from Theory to Applications. Berlin, Heidelberg: Springer-Verlag, 1998, 402- 411, 537- 546

[2] Mirhosseini A R, YAN Hong, Lam K M, et al. Human face image recognition: an evidence aggregation approach, Computer Vision and Image Understanding, 1998, 71(2): 213—230

[3] ZHANG Jun, YAN Yong, Lades M. Face recognition: eigenface, elastic matching, and neural nets. Proceedings of the IEEE, 1997, 85(9): 1423—1435

[4] Gutta S, Huang J, Imam I F, et al. Face and hand gesture recognition using hybrid classifiers. In: Pentland A, Essa I, Proceedings of the Second International Conference on Automatic Face and Gesture Recognition, Killington, Vermont, 1996; Los Alamitos, California: IEEE Computer Society Press, 1996, 164—169

[5] Lawrence S, Giles C L, Tsoi A C, et al. Face recognition: a convolutional neural network approach, IEEE Trans. on Neural Networks, 1997, 8(1): 98—113

[6] Belhumeur P N, Hespanha J P, Kriegman D J. Eigenfaces vs. fisherfaces: recognition using class specific linear projection, IEEE Trans. on PAMI, 1997, 19(7): 711—720

[7] Foggia P, Sansone C, Tortorella F, et al. Multiclassification: reject criteria for the Bayesian combiner. Pattern Recognition, 1999, 32: 1435—1447

[8] Martinez A, Benavente R. The AR face database, CVC Technical Report # 24, June 1998