Abstract A face recognition scheme is proposed, wherein a face image is preprocessed by pixel averaging and energy normalizing to reduce data dimension and brightness variation effect, followed by the Fourier transform to estimate the spectrum of the preprocessed image. The principal component analysis is conducted on the spectra of a face image to obtain eigen features. Combining eigen features with a Parzen classifier, experiments are taken on the ORL face database.

Key words face recognition; Fourier transform; principal component analysis; Parzen classifier; pixel averaging; energy normalizing

In recent years, automatic personal authentication and recognition become increasingly important because the security control is required everywhere. By virtue of stability and dependability, biometrics-based technologies such as speech recognition, fingerprint verification, iris authentication, and face recognition attract much attention. Among them, face recognition becomes one of the focuses due to its amicability and low-cost. However, the performance of a practical face recognition system is sensitive to variations in expressions, viewpoints, and dissimilar backgrounds. Therefore, how to reduce the variation effects becomes a challenging issue in the related study. Till now, numerous relevant algorithms have been proposed for feature extraction and classifier design, two of the most important modules in a face recognition system, to circumvent the problem. However, no single algorithm is proved successful in every application, while combinations of different methods show different performances.

This paper presents a scheme to solve the above problem. It includes preprocessing of a face image by pixel averaging and energy normalizing to reduce data dimension and brightness variation, the Fourier transform of a preprocessed image to estimate the shift invariant spectrum, and robust eigen feature extraction with the principle component analysis (PCA). For each subject, the trained eigen features are utilized to design a Parzen classifier. The eigen feature of a test face image is employed as input to the designed classifiers to obtain recognition results. Experimental results on the Olivetti Research Laboratory (ORL) face database show the feasibility of the presented scheme.

1 Preprocessing

The presented face recognition scheme is depicted in Fig.1. It includes three key modules: preprocessing, feature extraction based on the fast Fourier transform (FFT) and the PCA, and classification. This section addresses the relevant preprocessing technique.

Generally, face recognition directly using an original face image is too computationally intensive to be practically implemented, and hence data reduction is one of the key tasks. In preprocessing, we adopt a pixel averaging method to down-sample an original face image. Pixel averaging is a local spatial processing algorithm. The basic idea is to substitute several neighbored pixels by a single pixel, gray level of which is obtained by averaging those of the neighbored pixels. Specifically, suppose that there is an \(N_0 \times M_0\) face image \(g_0(x_0, y_0)\), and the pixel averaged one is an \(N \times M\) image \(g(n, m)\), then

\[
g(n, m) = \frac{1}{D_{x,y}} \sum_{x, y} g_0(x, y) \quad 1 \leq n \leq N, 1 \leq y \leq M
\]  

(1)
where \( S \) is the set of pixels in a selected block and \( D \) is the total number of the pixels in \( S \).

In practical applications, illumination condition affects the performance of a face recognition system\[1\]. To reduce the effect, we normalize a down-sampled image into an image that has unit energy. For \( g(n, m) \) in Eq.(1), the energy-normalized version is represented by \( I(n, m) = g(n, m)/\|g(n, m)\| \), where \( \|g(n, m)\| \) is the Frobenius norm. Properly, \( I(n, m) \) is the preprocessed version of the original face image \( g(n_0, m_0) \).

2 Feature Extraction

In feature extraction, the objective is to reduce data dimension so that the extracted feature is as representative as possible. Many successful feature extraction paradigms have been developed and they are grossly divided into two basic categories, geometric model-based approaches and statistical approaches. An approach of the former type extracts geometric features such as areas, distances and angles related to eyes, nose, mouth, and chin\[2\]. However, the geometric features are very sensitive to variations in facial expressions and details. On the other hand, a statistical approach attempts to capture and define the face as a whole to solve the above problem; it is widely studied due to the advantages in speed, simplicity, and relative insensitivity to small or gradual changes in the face images. Among the popular methods, the PCA and the linear discriminant analysis (LDA) are two classic tools\[3-7\]. However, many LDA-based algorithms suffer from the so-called “small sample size problem”, wherein the number of available samples is generally far smaller than the dimensionality of samples\[3\]. In addition, a statistical approach ignores the importance of local information in a face image.

In our discussion, feature extraction is conducted by the Fourier transform to stick out local information represented by variation in frequency, and PCA is successively performed to reduce data dimension.

Since a preprocessed two-dimensional face image \( I(n, m) \) may be considered as an element of the \( NM \)-dimensional vector space, it is row concatenated to form an \( NM \)-dimensional column vector \( z \). The spectra of \( z \) is estimated by discrete Fourier transform (DFT) as follows

\[
w(k) = \sum_{n=0}^{N-1} z(n) \exp \left(-\frac{j2\pi nk}{MN} \right) \]

where \( z(n) \) is the \( (n+1) \)-th element of the vector \( z \) and \( w(k) \) the \( (k+1) \)-th one of \( w \), with \( k = 0, 1, \cdots, NM-1 \). In realization, the FFT algorithm is adopted and the dimension of \( w \) may be reduced to one half according to symmetry properties of the DFT, and thus the following spectrum vector is obtained

\[
x = \left[ w(0) \quad w(1) \quad \cdots \quad w(L) \right]^T \]

(3)

where \( T \) denotes the transpose of a vector or a matrix, \( L = NM/2 \) for even \( NM \) and \( L = NM/2 + 0.5 \) for odd \( NM \). According to the shift invariant property of the DFT, the spectra vector in Eq.(3) is spatially shift invariant and thus insensitive to face translation within the image plane.

For ordinary face image size and down-sampling factor, generally the dimension of the spectra vector is still very large. In our study, the PCA is adopted for further data reduction.

Assume that the spectra vector set of the training face images is \( \{ x_i, i = 1, 2, \cdots, P \} \). The main idea is to find the vectors which best account for the distribution of the data set\[4\]. For the training spectra vector set, the covariance matrix is defined by

\[
C = \frac{1}{P} \sum_{i=1}^{P} \tilde{x}_i \tilde{x}_i^T
\]

(4)

where

\[
\tilde{x}_i = (x_i - \bar{x}), \quad \bar{x} = \frac{1}{P} \sum_{i=1}^{P} x_i
\]

(5)

From the covariance matrix \( C \), the PCA seeks a set of \( Q \) \( (Q<<L) \) orthogonal vectors associated with the \( Q \) eigenvalues of \( C \) to optimally describe the distribution of the data. However, the matrix \( C \) is \( L \) by \( L \) and thus it is computationally intensive to determine its \( L \) eigenvectors and eigenvalues. Practically, the number of training images, \( P \), is smaller than the dimension of a spectra vector, \( L \). To reduce the computational load, we apply singular value decomposition (SVD) to the smaller \( L \) by \( P \) matrix, \( \tilde{X} = [\tilde{x}_1 \quad \tilde{x}_2 \quad \cdots \quad \tilde{x}_P] \). The associated singular values of \( \tilde{X} \) allow us to rank the eigenvectors according to their usefulness in characterizing the variation among the training spectra vectors. In our study, \( Q \) orthogonal eigenvectors corresponding to the \( Q \) largest singular values are adopted to characterize the variation and span a space called eigensubspace. The value of \( Q \) is determined by a pre-selected suppression ratio \( \eta \) as follows

\[
\eta = \frac{\lambda_1 + \lambda_2 + \lambda_3 + \cdots + \lambda_Q}{\lambda_1 + \lambda_2 + \lambda_3 + \cdots + \lambda_P}
\]

(6)
where $\lambda_1 \geq \lambda_2 \geq \lambda_3 \geq \cdots \geq \lambda_r$ are absolutes of the $P$ singular values of $X$.

Once the set of $Q$ orthogonal eigenvectors $\{u_j, j = 1, 2, \ldots, Q\}$ is obtained, the eigen feature vector, $\xi$, of a face image is formed by projecting the spectra vector $x_i$ onto the spanned eigensubspace. The $j$-th element of $\xi$ is determined by $\xi(j) = u_j^T(x_i - \bar{x})$ for $j = 1, 2, \ldots, Q$.

### 3 Parzen Classifier Design

In our scheme, a Parzen classifier is adopted for eigen feature classification. The Parzen classifier is a kernel density estimator, with which a nonlinear function is approximated by the superposition of a set of kernels. For a pattern class, its Gaussian kernel density estimator, with which a nonlinear eigen feature classification. The Parzen classifier is a kernel-based Parzen classifier is determined by the width $\lambda$. For a pattern class, its Gaussian function is approximated by the superposition of a set $\{Q_j, j = 1, 2, \ldots, R\}$. Once a Parzen classifier for a training face image and the relevant kernel width $\lambda$ is substituted by

$$f_\lambda(y, s) = \frac{1}{R} \sum_{j=1}^{R} \frac{1}{(s \sqrt{2\pi})^D} \exp\left(-\frac{||y - \xi_j||^2}{2s^2}\right) \quad (7)$$

where $\xi_j$ is the $Q$-dimensional eigen feature vector of a training face image and $y$ is that of a test one. In our discussion, Parzen classifier design means to estimate the relevant kernel width $s$ using the training feature set $\{\xi_j, i = 1, 2, \ldots, R\}$. Once a Parzen classifier for each subject is designed, the face of an unknown subject could be verified or recognition by the outputs from the classifiers of all library subjects.

Among the proposed algorithms for Parzen classifier design, the one proposed by Duin adopts a maximum likelihood principle to estimate the kernel width $s$, wherein the commonly used likelihood function

$$L(s) = \prod_{j=1}^{R} f_\lambda(\xi_j, s) = \prod_{j=1}^{R} \frac{1}{(s \sqrt{2\pi})^D} \exp\left(-\frac{||\xi_j - \xi||^2}{2s^2}\right)$$

is substituted by

$$L'(s) = \left[\prod_{j=1}^{R} \frac{1}{(s \sqrt{2\pi})^D} \exp\left(-\frac{||\xi_j - \xi||^2}{2s^2}\right)\right]^{\frac{1}{s^D}}$$

Setting the derivative of the logarithm of $L'(s)$ to zero yields

$$\frac{1}{R} \sum_{j=1}^{R} \frac{||\xi_j - \xi||^2}{s^D} \exp\left(-\frac{||\xi_j - \xi||^2}{2s^2}\right) - \frac{Q}{s} = 0$$

Based on the above equation, a crude but rather fast way to obtain an approximate solution is to substitute the superposition in the denominator by average, so that the above equation collapses to

$$\frac{1}{R} \sum_{j=1}^{R} \frac{||\xi_j - \xi||^2}{s^D} = \frac{Q}{s} \quad (8)$$

The root of Eq.(8) is found by choosing $s$ as

$$s = \sqrt{\frac{1}{Q \sum_{j=1}^{R} ||\xi_j - \xi||^2} - 1} \quad (9)$$

### 4 Experimental Results

We take experiments on the ORL face database to demonstrate the feasibility of the proposed scheme. The database contains face images of 40 subjects with 10 images per subject. Some of the images are taken in different periods and thus subjected to variations in brightness and facial pose, expression, and details (with/without glasses).

We use the training eigen features of all subjects to estimate the Parzen classifier parameter $s$ for common use, since the training samples of each subject are too small to obtain a good estimation of $s$; while for each subject, a Parzen classifier is constructed by its own training eigen features.

Our experiment examines the effect of training size on recognition rate. Three to seven prototype images per subject are randomly selected for training, the rest ones for test. The recognition rates of 20 different runs are averaged and listed in Tab.1, wherein the suppression ratio $\eta$ is set to 91%. For convenience of comparison, Tab.1 lists the results of six different schemes, they are pixel averaging (PA) and PCA plus respectively the Euclidean distance classifier (ED) and the Parzen classifier (PZ), wavelet transform (WT) plus respectively ED and PZ, and our preprocessing method (PN) and FFT and PCA plus respectively ED and PZ. Obviously, recognition rate increases with the number of training images and our scheme outperforms other listed ones. When five images per subject are used for training, the error rates of our scheme (2.40%) are comparable to those of some current methods on the same database, such as the individual PCA (5.0%)\(^6\), the Bayesian classifier for SVD coefficients (7.5%)\(^6\), the direct LDA (about 6.0%)\(^7\), the kernel density estimation (KDE) classifier.
for eigenfaces (6.55%)\cite{8}, and a kernel-based nonlinear representor for eigenfaces (7.0%)\cite{9}.

<table>
<thead>
<tr>
<th>Scheme</th>
<th>Training images per subject</th>
</tr>
</thead>
<tbody>
<tr>
<td>PA+PCA+ED</td>
<td>83.69 87.69 90.38 91.16 91.00</td>
</tr>
<tr>
<td>PA+PCA+PZ</td>
<td>87.16 91.88 94.78 95.97 96.54</td>
</tr>
<tr>
<td>WL+ED</td>
<td>82.80 87.15 88.90 90.25 90.13</td>
</tr>
<tr>
<td>WL+PZ</td>
<td>87.66 92.71 95.28 96.25 96.46</td>
</tr>
<tr>
<td>PN+PCA+ED</td>
<td>88.63 91.35 93.08 94.63 94.63</td>
</tr>
<tr>
<td>PN+PCA+PZ</td>
<td>92.07 95.42 97.60 98.41 98.29</td>
</tr>
</tbody>
</table>

5 Conclusions

This paper presents a face recognition scheme consisted of a novel preprocessing technique, an FFT and PCA based feature extracting algorithm, and a Parzen classifier. Experimental results on the ORL face database show the feasibility of the proposed novel scheme.

Acknowledgement

The related work is partly supported by the Young Scholar Foundation of UESTC (No.JX200402).

References


Brief Introduction to Author(s)

SUN Xin (孙鑫) was born in 1980. She is now pursuing her M.S degree at UESTC. Her research interest is in signal processing. E-mail: uestcsunxin@163.com.

LIU Bing (刘兵) was born in 1978. He is now pursuing his M.S degree at UESTC. His research interest is in signal processing. E-mail: storm_liu@163.com.

LIU Ben-yong (刘本永) was born in 1966. His is now an associate professor and Ph.D at UESTC. His research interests include pattern recognition, signal processing, and machine learning. E-mail: byliu@uestc.edu.cn.