Class-Specific Subspace-Based Two-Dimensional Principal Component Analysis for Face Recognition

P. Sanguansat, W. Asdornwised, S. Jitapunkul
Department of Electrical Engineering
Chulalongkorn University
Bangkok, Thailand
parinya-san@student.chula.ac.th
{widhyakorn.a, somchai.j}@chula.ac.th

S. Marukatat
National Electronics and Computer Technology Center (NECTEC)
Pathumthani, Thailand
sanparith.marukatat@nectec.or.th

Abstract

In this paper, we proposed a class-specific subspace-based Two-Dimensional Principal Component Analysis (2DPCA) for face recognition. In 2DPCA, 2D face image matrices do not need to be previously transformed into a vector. In this way, the spatial information can be preserved. Moreover, 2DPCA can achieve higher performance than PCA both in face recognition and face representation task. However, both PCA and 2DPCA are unsupervised techniques, no information of class labels are considered. Therefore, the directions that maximize the scatter of the data might not be as adequate to discriminate between classes. In recognition task, a projection is required to emphasize the discrimination between classes. The Face-Specific Subspace (FSS) was proposed in concept of class-specific subspace. Each subspaces learned from the training images which correspond to only one class, thus the number of these subspaces is equal to the number of classes. Since the information of class labels are considered in FSS, so the discriminant power can be improved. We apply 2DPCA to class-specific concept in our framework which consists of two methods: the first one, we apply FSS to 2DPCA method and the second one, we use the Bilateral-projection-based 2DPCA (B2DPCA) instead of 2DPCA. The B2DPCA does not only allows further reducing of the dimension of feature matrix of 2DPCA-based but also improving the classification accuracy. Experimental results on Yale face database showed an improvement of our proposed techniques over the conventional 2DPCA.

1. Introduction

In face recognition, a great number of successful face recognition systems have been developed and reported in the literature [8, 9, 1, 13, 14]. Among these works, the linear subspace techniques, such as Principal Component Analysis (PCA) and Linear Discriminant Analysis (LDA) are the most popular ones. The PCA's criterion chooses the subspace in the function of data distribution while LDA chooses the subspace which yields maximal inter-class distance, and at the same time, keeping the intra-class distance small. In general, LDA extracts features which are better suitable for classification task. Both techniques intend to project the vector representing face image onto lower dimensional subspace, in which each 2D face image matrix must be previously transformed into vector and then a collection of the transformed face vectors are concatenated into a matrix. This is the cause of three serious problems in particular approaches. First of all, the covariance matrix, which collects the feature vectors with high dimension, will leads to the curse of dimensionality, which computing is very memory and time-consuming. Secondly, the spatial structure information could be lost when the column-stacking vectorization and image resizeation are applied. Finally, normally in face recognition task, the available number of training samples is relatively very small compared to the feature dimension, so the covariance matrix which estimated by these features will be singular, called singularity problem or Small Sample Size (SSS) problem [3]. Especially, as a supervised technique, LDA has a tendency to overfitting in SSS problems.

Various solutions have been proposed for solving the SSS problem [2, 6, 4, 1, 13, 14]. Among these LDA extensions, Fisherface [1] and the discriminant analysis of principal components framework [13, 14] demonstrates a significant improvement when applying LDA over principal components from the PCA-based subspace. Since both PCA and LDA can overcome the drawbacks of each other. PCA is constructed around the criteria of preserving the data distribution. Hence, it is suited for face representation and re-
construction from the projected face feature. However, in the classification tasks, PCA only normalize the input data according to their variance. This is not efficient since the between classes relationship is neglected. In general, the discriminatory power depends on both within and between classes relationship. LDA considers these relationships via scatter matrices analysis of within and between-class scatter matrices. Taking this information into account, LDA allows further improvement. Especially, when there are prominent variation in lighting condition and expression. However, LDA has certain two drawbacks when directly applied to the original input space [14]. First of all, some non-face information such as image background are regarded by LDA as the discriminant information. This causes misclassification when the face of the same subject is presented on different background. Secondly, when SSS problem has occurred, the within-class scatter matrix be singular, so-called the singularity problem. Projecting the high dimensional input space into low dimensional subspace via PCA can solve these LDA problems. Nevertheless, all of above techniques, the spatial structure information still be lost.

Recently, Yang et al. [11] proposed an original technique called Two-Dimensional Principal Component Analysis (2DPCA), in which the image covariance matrix is computed directly on image matrices so the spatial structure information can be preserved. This yields a covariance matrix whose dimension just equals to the width of the face image. This is far smaller than the size of covariance matrix in PCA. Therefore, the image covariance matrix can be better estimated and full rank. Evidently, the experimental results in [11] have shown the improvement of 2DPCA over PCA on several face databases.

Subsequently, Bilateral-projection-based 2DPCA (B2DPCA) was proposed in [12, 5]. The bilateral projection scheme was applied there via left and right multiplying projection matrices. In this way, the eigenvalue problem was solved two times. One corresponds to the column direction and another one corresponds to the row direction. As a result, the feature matrix and computation time is reducing. Especially, the B2DPCA achieves the improvement in recognition accuracy than 2DPCA.

However, 2DPCA and B2DPCA is a unsupervised technique that is no information of class labels are considered. Therefore, the directions that maximize the scatter of the data from all training samples might not be as adequate to discriminate between classes. In recognition task, a projection is required to emphasize the discrimination between classes. In [7], the extension of Eigenface, PCA-based, was proposed by using alternative way to represent by projecting to Face-Specific Subspace (FSS). In conventional Eigenface method, face images are analyzed on the features extracted in a low-dimensional space learned from all training samples from all classes. While each subspaces of FSS learned from training samples from one class. In this way, the FSS representation can provide a minimum reconstruction error. The reconstruction error is used to classify the input data via the Distance From FSS (DFFSS), less DFFSS means more probability that the input data belongs to the corresponding class.

The remainder of this paper is organized as follows: In Section 2, the 2DPCA concept are described. The B2DPCA is introduced in Section 3. In Section 4, the FSS technique is described. In Section 5, the experimental results are presented for the Yale face image database to demonstrate the effectiveness of out proposed techniques compared to 2DPCA. Finally, conclusions are presented in Section 6.

2. Two-Dimensional Principal Component Analysis

The Two-Dimensional Principal Component Analysis (2DPCA) in [11], the image covariance matrix \( G \) was defined as

\[
G = E[(A - EA)^T (A - EA)],
\]

where \( A \) represents the face image. This is much smaller than the size of real covariance matrix needed in PCA, therefore can be computed more accurately on small training set. Given a database of \( M \) training image matrices \( A_e, e = 1, \ldots, M \) with same dimension \( m \) by \( n \). The matrix \( G \) is computed in a straightforward manner by

\[
G = \frac{1}{M} \sum_{k=1}^{M} (A_k - \bar{A})^T (A_k - \bar{A}),
\]

where \( \bar{A} \) denotes the average image, \( \bar{A} = \frac{1}{M} \sum_{e=1}^{M} A_e \). Let \( x_1, \ldots, x_d \) be \( d \) selected largest eigenvectors of \( G \). Each image \( A \) is projected onto these \( d \) dimensional subspace. The projected image \( Y \) is then an \( m \) by \( d \) matrix given by \( Y = AX \) where \( X = [x_1, \ldots, x_d] \) is a \( n \) by \( d \) projection matrix.

A nearest neighbor classifier is used for classification. The distance between two arbitrary feature matrices, \( Y_i = [y_{i1}, y_{i2}, \ldots, y_{ip}] \) and \( Y_j = [y_{j1}, y_{j2}, \ldots, y_{jp}] \) is defined by

\[
d(Y_i, Y_j) = \sum_{n=1}^{p} ||y_{in} - y_{jn}||,
\]

where \( ||y_{in} - y_{jn}|| \) denotes the Euclidean distance between the two principal component vectors \( y_{in} \) and \( y_{jn} \). Suppose that the feature matrices of \( M \) training samples are \( Y_{il}, l = 1, 2, 3, \ldots, M \), and that each of these samples is assigned to class \( \omega_k \). Given a feature matrix of test sample \( Y_{test} \), if

\[
d(Y_{test}, Y_l) = \min_{j} d(Y_{test}, Y_j)
\]

then the test sample \( Y_{test} \) is belong to class \( \omega_k \).
3. Bilateral-projection-based 2DPCA

The Bilateral-projection-based 2DPCA (B2DPCA) in [5, 12], bilateral projection scheme was applied to 2DPCA via left and right multiplying projection matrices as follows

\[
B = Z^TAX,
\]

where \(B\) is a feature matrix which extracted from image \(A\) and \(Z\) is a left multiplying projection matrix. Similar to the right multiplying projection matrix \(X\) in Section 2, matrix \(Z\) is a \(m\) by \(q\) projection matrix that obtained by choosing the eigenvectors of mean covariance matrix \(H\) corresponding to the \(q\) largest eigenvalues. The matrix \(H\), which corresponds to the column direction of images, can be evaluated by

\[
H = \frac{1}{M} \sum_{c=1}^{M} (A_c - \bar{A})(A_c - \bar{A})^T.
\]

Therefore, the dimension of feature matrix is decreasing from \(m \times d\) to \(q \times d\). In this way, the computation time also be reducing. Moreover, the recognition accuracy of B2DPCA is better than 2DPCA. A nearest neighbor classifier is used for classification like in 2DPCA but the distance between \(B_i\) and \(B_j\) is defined by

\[
d(B_i, B_j) = \sqrt{\sum_{m=1}^{q} \sum_{n=1}^{d} (b_{(m,n)} - b_{(m,n)}')^2}.
\]

Since the feature matrices are extracted from both row and column directions of images, thus the distance will be sum of distances of each elements \(b_{(m,n)}\) in feature matrix \(B\).

4. Face-Specific Subspace

In [7], the original Face-Specific Subspace (FSS) was proposed for applying to Eigenface, PCA-based method. The difference from traditional method is the covariance matrix of the \(k^{th}\) class is evaluated from training samples of the \(k^{th}\) class, individually. The \(k^{th}\) FSS was represented as a 4-tuple: the projection matrix, the mean of the \(k^{th}\) class, the eigenvalues of covariance matrix and the dimension of the \(k^{th}\) FSS. For identification, the input sample is projected to all FSSs and then reconstruct by that FSS. If reconstruction error which obtained from the \(k^{th}\) FSS is minimum then the input sample is belong to the \(k^{th}\) class, so called Distance From FSS (DFFSS). In this paper, the FSS is adapted to 2DPCA and B2DPCA frameworks.

4.1. FSS-Based 2DPCA

Let \(G_k\) be the image covariance matrix of the \(k^{th}\) FSS. Then \(G_k\) can be evaluated by

\[
G_k = \frac{1}{M} \sum_{A_c \in \omega_k} (A_c - \bar{A}_k)(A_c - \bar{A}_k)^T,
\]

where \(\bar{A}_k\) is the average image of class \(\omega_k\). The \(k^{th}\) projection matrix \(X_k\) is a \(n\) by \(d_k\) projection matrix which composed by the eigenvectors of \(G_k\) corresponding to the \(d_k\) largest eigenvalues. The \(k^{th}\) FSS of 2DPCA was represented as a 3-tuple:

\[
\gamma_k^{2DPCA} = \{X_k, \bar{A}_k, d_k\}
\]

4.2. FSS-Based B2DPCA

For the image covariance matrix \(H\), which corresponds to the column direction of images, can be also evaluated by

\[
H_k = \frac{1}{M} \sum_{A_c \in \omega_k} (A_c - \bar{A}_k)(A_c - \bar{A}_k)^T.
\]

Then the \(k^{th}\) FSS of B2DPCA was represented as a 5-tuple:

\[
\gamma_k^{B2DPCA} = \{Z_k, X_k, \bar{A}_k, d_k, q_k\},
\]

where \(Z_k\) is left multiplying projection matrix, which composed by the eigenvectors of \(H_k\) corresponding to the \(q_k\) largest eigenvalues, and \(X_k\) is right multiplying projection matrix, which composed by the eigenvectors of \(G_k\) corresponding to the \(d_k\) largest eigenvalues.

4.3. Classification Method

Let \(S\) be a input sample and \(U_k\) be a feature matrix which projected to the \(k^{th}\) FSS, by

\[
U_k = \begin{cases} W_kX_k, & \text{in case 2DPCA} \\ Z_k^T W_k Z_k, & \text{in case B2DPCA} \end{cases},
\]

where \(W_k = S - \bar{A}_k\). Then the reconstruct image \(W_k^r\) can be evaluates by

\[
W_k^r = \begin{cases} U_kX_k^T, & \text{in case 2DPCA} \\ Z_k U_k X_k^T, & \text{in case B2DPCA} \end{cases}
\]

Therefore, the DFFSS is defined by reconstruction error as follows

\[
\varepsilon_k(W_k^r, S) = \sum_{m=1}^{n_{row}} \sum_{n=1}^{n_{col}} |w_{(m,n)}^r - s_{(m,n)}|.
\]

If \(\varepsilon_k = \min_{1 \leq k \leq K} (\varepsilon_k)\) then the input sample \(S\) is belong to class \(\omega_k\).
Figure 1. Five sample images of one subject in the Yale database.

Table 1. Comparisons of 2DPCA, B2DPCA, 2DPCA+FSS and B2DPCA+FSS on Yale database

<table>
<thead>
<tr>
<th>Method</th>
<th>Accuracy (%)</th>
<th>d</th>
<th>q</th>
<th>Dimension</th>
</tr>
</thead>
<tbody>
<tr>
<td>2DPCA</td>
<td>87.78</td>
<td>5</td>
<td>N/A</td>
<td>100 × 5</td>
</tr>
<tr>
<td>B2DPCA</td>
<td>92.22</td>
<td>23</td>
<td>22</td>
<td>22 × 23</td>
</tr>
<tr>
<td>2DPCA+FSS</td>
<td>92.22</td>
<td>5</td>
<td>N/A</td>
<td>100 × 5</td>
</tr>
<tr>
<td>B2DPCA+FSS</td>
<td>94.44</td>
<td>1</td>
<td>1</td>
<td>1 × 1</td>
</tr>
</tbody>
</table>

5. Experimental results

In this section, we experimentally evaluate our proposed framework by using four methods: 2DPCA, B2DPCA, 2DPCA+FSS and B2DPCA+FSS. 2DPCA is used as baseline method for comparison on the well-known Yale [10] database. The Yale database contains 165 images of 15 subjects. There are 11 images per subject, one for each of the following facial expressions or configurations: center-light, with glasses, happy, left-light, without glasses, normal, right-light, sad, sleepy, surprised, and wink. Five sample images of one person from the Yale database are shown in Fig. 1. Each image was manually cropped and resized to 100 × 80 pixels.

In all experiments, the five image samples (centerlight, glasses, happy, leftlight, and noglasses) are used to train, and the six remaining images (normal, rightlight, sad, sleepy, surprised and wink) for test. We vary the number of principal component vectors $d$ and $q$ from 1 to 30 since the highest recognition accuracy lie on this interval. The experimental results are shown in Table 1. The recognition accuracy of B2DPCA+FSS is higher than the other methods. Especially, B2DPCA+FSS use only one feature vectors to obtain the highest recognition rate.

6. Conclusions

In this paper, the class-specific method of both 2DPCA and B2DPCA for face recognition are proposed, based on FSS. The advantage of our proposed method is the excellent performance over conventional 2DPCA and B2DPCA under variations in expression and illumination. Evidently, the experimental results have shown the improvement recognition accuracy on well-known face database. However, the disadvantage of the proposed method is it requires more memory for storing each classes and recognition time.

References