Face Recognition and Representation by Tensor-based MPCA Approach

Yun-Hee Han
Dept. of Control, Instrumentation, and Robot Engineering
Chosun University
Gwangju, Korea
Yhhan1059@gmail.com

Keun-Chang Kwak
Dept. of Control, Instrumentation, and Robot Engineering
Chosun University
Gwangju, Korea
Kwak@chosun.ac.kr

Abstract — This paper is concerned with face recognition and representation by Tensor-based Multilinear Principal Component Analysis (T- MPCA) approach. This method extracts feature directly from the tensor representation rather than the vector representation. Many image and video data are naturally tensor objects. For example, colour images are 3D (third-order tensor) objects with column, row, and colour modes. Currently, most face recognition algorithms are performed by 2D gray images. Therefore, video sequences with gray image can be viewed as third-order tensors with column, row, and time modes. In order to evaluate the performance, we perform face recognition based on face database constructed in test bed environments. The experimental results on face database reveal that the presented method shows a better performance in comparison with the well-known methods in distant-varying environments.

Keywords- tensor-based MPCA; face recognition; video sequence; distant-varying environments

I. INTRODUCTION

Today, face recognition being regarded as a fundamental technology of biometrics has been applied to a variety of areas, including image processing, computer vision, and pattern recognition. The most representative recognition techniques frequently used in conjunction with face recognition are Principal Component Analysis (PCA) and Fisher’s Linear Discriminant Analysis (LDA). The PCA approach known as eigenface method is a popular unsupervised statistical technique that supports finding useful image representations. It also exhibits optimality when it comes to dimensionality reduction [1-3]. However, the PCA is not ideal for classification purposes mainly because of the fact it retains unwanted variations occurring due to lighting and facial expression. There are numerous extensions to the standard PCA method. Meanwhile, the LDA method known as fisherface method is a supervised learning approach whose functioning depends on class-specific information [4-5]. This statistically motivated method maximizes the ratio of between-class scatter and within-class scatter and is also an example of a class-specific learning method. Again, there are various enhancements made to the LDA.

The PCA-based and LDA-based approaches extract features which expressed high dimensionality vector from two-dimensionality image. Naturally, color video sequences are fourth-order tensors with the addition of a color mode. In the most active area of biometrics research, namely, that of face recognition, 3-D face detection and recognition using 3-D information with column, row, and depth modes, in other words, a third-order tensor, has emerged as an important research direction [6–8]. However, face images is naturally two-order tensor. Therefore, it is appropriate to extract tensor representation directly than vector representation. On the other hand, video sequences with 2D gray images can be viewed as third-order tensors with column, row, and time modes.

Therefore, in order to solve the third-order tensors problem, we shall use Tensor-based Multilinear Principal Component Analysis (T-MPCA) algorithm proposed by Lu [9]. This method is a multilinear algorithm performing dimensionality reduction in all tensor modes seeking those bases in each mode that allow projected tensors to capture most of the variation present in the original tensors [9]. Furthermore, this has been successfully applied to human gait recognition. In this paper, we perform face recognition and representation using T-MPCA approach under distant-varying environments. In order to evaluate performance of presented method, face database was constructed in test bed environments. Finally, we demonstrate the superiority and effectiveness of proposed method.

II. PCA AND LDA

Principal Component Analysis (PCA) is a well-known technique commonly exploited in multivariate linear data analysis. The main underlying concept is to reduce the dimensionality of a data set while retaining as much variation as possible in the data set. Let a face image be a two-dimensional array of pixels. The corresponding image is viewed as a vector with coordinates that results from a concatenation of successive rows of the image. Denote the training set of faces by. Define the corresponding covariance matrix in the standard manner

$$R = \frac{1}{N} \sum_{i=1}^{N} (z_i - \bar{z})(z_i - \bar{z})^T = \Phi \Phi^T$$

(1)

$$\bar{z} = \frac{1}{N} \sum_{i=1}^{N} z_i$$

(2)
Let \( E = (e_1, e_2, \cdots, e_r) \) be a matrix formed by “r” eigenvectors corresponding to the “r” largest eigenvalues. Thus, for a set of original face images \( Z \), their reduced feature vectors \( X = (x_1, x_2, \cdots, x_N) \) are obtained by projecting them into the PCA-transformed space following the linear transformation

\[
x_i = E^T(z_i - \overline{z})
\]

The classification task is cast in the reduced PCA space as all the computations of the distances are carried out there. Given two images \( z' \) and \( z \) in the original \( n^2 \)-dimensional space, a distance between them is defined in the form

\[
e(z, z') = \|x - x'\|
\]

where \( x \) and \( x' \) are PCA-transformed feature vectors of face image \( z \) and \( z' \), respectively.

Denote the training set of \( N \) face images by \( Z = (z_1, z_2, \cdots, z_N) \). The feature vectors \( X = (x_1, x_2, \cdots, x_N) \) can be obtained by projecting \( Z \) into the PCA-transformed space as follows

\[
x_i = E^T(z_i - \overline{z})
\]

where \( E = (e_1, e_2, \cdots, e_r) \) contain the \( r \) eigenvectors corresponding to the \( r \) largest eigenvalues.

The next stage is based upon the use of the FLD and can be described as follows. Consider \( c \) classes existing in the problem with \( N \) samples. The between-class and within-class scatter matrices are defined in the form

\[
S_b = \sum_{i=1}^{c} N_i (m_i - \overline{m})(m_i - \overline{m})^T
\]

\[
S_w = \sum_{i=1}^{c} \sum_{x \in C_i} (x - m_i)(x - m_i)^T
\]

where \( N_i \) is the number of samples occurring in \( i \)’th class \( C_i \) and \( \overline{m} \) is the mean of all samples, while \( m_i \) denotes the mean of class \( C_i \). The optimal projection matrix \( W_{FLD} \) is chosen as follows

\[
W_{FLD} = \arg \max_{W} \frac{W^T S_b W}{W^T S_w W} = [w_1 \ w_2 \ \cdots \ w_m]
\]

\[
S_b w_i = \lambda_i S_w w_i \ i = 1, 2, \ldots, m
\]

where \( [w_i | i = 1, 2, \cdots, m] \) is the set of discriminant vectors of \( S_b \) and \( S_w \) corresponding to the \( c-1 \) largest generalized eigenvalues. Thus, the feature vectors \( V = (v_1, v_2, \cdots, v_N) \) for any face images \( z_i \) can be then computed as

\[
v_i = W_{FLD}^T x_i = W_{FLD}^T E^T (z_i - \overline{z})
\]

III. TENSOR-BASED MPCA

Tensor-based Multilinear Principal Component Analysis (T-MPCA) algorithm is introduced in detail for dimensionality reduction in [9]. In this paper, the MPCA algorithm is used for face recognition. Before describing MPCA, the notations in [9] will be used in this paper. A Nth-order tensor is denoted as \( A \in R^{I_1 \times I_2 \times \cdots \times I_N} \). It is addressed by \( N \) indices \( i_m, n = 1, \ldots, N \), and each \( i_m \) addresses the n-mode of \( A \). The n-mode product of a tensor \( A \) by a matrix \( U \in R^{I_m \times I'_m} \), is denoted as \( A \times_n U \), is a tensor with entries

\[
(A \times_n U)(i_1, i_2, \ldots, i_N) = \sum_{i_m} A(i_1, i_2, \ldots, i_N) \cdot U(j_m, i_m)
\]

Following standard multiliner algebra, any tensor \( A \) can be expressed as the product

\[
A = S \times_1 U^{(1)} \times_2 U^{(2)} \times_3 U^{(3)} \cdots \times_N U^{(N)}
\]

where \( S = A \times_1 U^{(1)} \times_2 U^{(2)} \times_3 U^{(3)} \cdots \times_N U^{(N)} \) and \( U^{(n)} = (u^{(n)}_1 u^{(n)}_2 \cdots u^{(n)}_M) \) is an orthogonal \( I_n \times I_n \) matrix.

Let \( \{A_m, m=1, \ldots, M\} \) be a set of \( M \) tensor samples in \( R^{I_1} \otimes R^{I_2} \otimes \cdots \otimes R^{I_N} \). The total scatter of these tensors is defined as \( \Psi = \sum_{m=1}^{M} A_m - \overline{A}^T \), where \( \overline{A} \) is the mean tensor calculated as \( \overline{A} = \frac{1}{M} \sum_{m=1}^{M} A_m \). Give all the other projection matrices \( \widetilde{U}^{(n)} \) consist of the \( P_n \) eigenvectors corresponding to the largest \( P_n \) eigenvalues of the matrix \( \prod_{n=1}^{N} P_n \). Feature vectors attained by classification commuted following equation

\[
Y = X \times_1 \widetilde{U}^{(1)} \times_2 \widetilde{U}^{(2)} \times_3 \widetilde{U}^{(3)} \cdots \times_N \widetilde{U}^{(N)}
\]

The pseudo-code implementation of the MPCA algorithm is as follows [9]

[Step 1] Center the input samples as \( \overline{X} = X - \overline{X} \), where \( \overline{X} = \frac{1}{M} \sum_{m=1}^{M} X_m \) is the sample mean.

[Step 2] Calculate the eigen-decomposition of \( \Phi^{op} = \sum_{m=1}^{M} X_m \cdot \overline{X}^{T} \) and set \( \widetilde{U}^{(n)} \) to consist of the eigenvectors corresponding to the most significant \( P_n \) eigenvalues, for \( n = 1, \ldots, N \).

[Step 3] Calculate
\[
\{ \tilde{Y}_m = \tilde{X}_m \times U^{(1)} \times \cdots \times U^{(N)} , m = 1, \ldots, M \}.
\]

- Calculate \( \Psi_{k_l} = \sum_{m=1}^{M} Y_{m_l} \).

  (the mean \( \bar{Y} \) is all zero since \( \tilde{X}_m \) is centered).

- For \( k = 1 : K \)
  - For \( n = 1 : N \)
    * Set the matrix \( \tilde{U}^{(n)} \) to consist of the \( P_n \) eigenvector of the matrix \( \Phi^{(n)} \), corresponding to the largest \( P_n \) eigenvalues.
    - Calculate \( \{ \tilde{Y}_m, m = 1, \ldots, M \} \) and \( \Psi_{k_l} \).
    - If \( \Psi_{k_l} \neq \Psi_{k_l-1} < \eta \), break and go to Step 4.

[Step 4] The feature tensor after projection is obtained as \( \{ \tilde{Y}_m = \tilde{X}_m \times \tilde{U}^{(1)} \times \cdots \times \tilde{U}^{(N)} , m = 1, \ldots, M \} \).

IV. EXPERIMENTAL RESULTS

This section reports on the comprehensive of experiments and draws conclusions as to performance of the presented method. The goal of this experimentation is to compare the performance of proposed method with the well-known methods in template matching. For experiments, we captured sequence images through web camera in test bed environments. We used frontal face images with distant-varying in 1 meter and 2 meter. This face database includes 300 different images coming from 10 individuals. Here, we consider the size of enrolment data, processing time, user convenience, recognition performance, and distant-varying. Among total database, we used 50 face images obtained from 1-meter for enrolment and the remaining 250 face images for test stage. Each image was digitized and resized by a 45x40 whose gray levels ranged between 0 and 255. Fig. 1 shows some example of face database constructed in distant-varying environments.

The input is a third-order tensor. Fig. 2 visualizes three modes such as spatial column space, row space, and the time space.

The size of data set and processing time are in listed Table I-II. As listed in Table 2, the results obtained by the proposed method showed the faster processing times in comparison with the previous approaches.

| TABLE I. SIZE OF TEST DATA SET AND FEATURE VECTOR ACCORDING TO DISTANCE VARIATION |
|-----------------|-----------------|-----------------|-----------------|
|                  | Test data set   |                  | Feature vector  |
|                  |                 | 1m              |                 | 2m              |
| PCA              | 500×1800        | 2000×1800       | 500×100         | 2000×100        |
| T-MPCA           | 45×40×10×50     | 45×40×10×200    | 50×100          | 200×100         |

| TABLE II. COMPARISON OF TRAINING TIME AND PROCESSING TIME (CPU Intel i7 860, 2.80 GHz, 3 GB RAM SYSTEM) |
|------------------------------------------------|-----------------|-----------------|-----------------|
|                  | Training time   | Processing time |
|                  |                 | 1m              |                 | 2m              |
| PCA              | 0.371 sec       | 0.201 sec       | 0.849 sec       |
| LDA              | 0.380 sec       | 0.458 sec       | 1.968 sec       |
| T-MPCA           | 0.276 sec       | 0.019 sec       | 0.074 sec       |

Fig. 3 shows the performance comparisons of PCA, LDA, and T-MPCA. As shown in Fig. 3, the experimental result obtained by T-MPCA approach yielded a good performance in comparison to PCA, while that of LDA outperformed T-MPCA for 2 meter. Fig. 4 shows the classification results by the number of eigen tensor faces.

![Figure 1. Some of face database with distance variation.](image)
V. CONCLUSION

In this paper, we introduce non-cooperative user identification system in distant-varying environments by T-MPCA. The proposed methods have yielded similar results like previous approaches. On the other hand, the processing time is superior to previous approaches. We will plan to apply to real-time systems by complementing the recognition performance of the proposed method.

ACKNOWLEDGMENT

This research was financially supported by the Ministry of Education, Science Technology (MEST) and Korea Institute for the Advancement of Technology (KIAT) through the Human Resource Training Project for Regional Innovation.

REFERENCES