Pattern Recognition 43 (2010) 1106–1115

Contents lists available at ScienceDirect

Pattern Recognition

journal homepage: www.elsevier.de/locate/pr

A feature extraction method for use with bimodal biometrics

Yong Xu [a], David Zhang [a,b,*], Jing-Yu Yang [c]

[a] Bio-Computing Research Center, Shenzhen Graduate School, Harbin Institute of Technology, Shenzhen, China
[b] Biometrics Research Centre, Department of Computing, The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong
[c] School of Computer Science and Technology, Nanjing University of Science and Technology, Nanjing, China

ARTICLE INFO

Article history:
Received 27 May 2008
Received in revised form
1 September 2009
Accepted 8 September 2009

Keywords:
Principal component analysis
Feature extraction
Matrix-based complex PCA (MCPCA)
Palmprint identification
Biometrics

ABSTRACT

Bimodal biometrics has been found to outperform single biometrics and are usually implemented using the matching score level or decision level fusion, though this fusion will enable less information of bimodal biometric traits to be exploited for personal authentication than fusion at the feature level. This paper proposes matrix-based complex PCA (MCPCA), a feature level fusion method for bimodal biometrics that uses a complex matrix to denote two biometric traits from one subject. The method respectively takes the two images from two biometric traits of a subject as the real part and imaginary part of a complex matrix. MCPCA applies a novel and mathematically tractable algorithm for extracting features directly from complex matrices. We also show that MCPCA has a sound theoretical foundation and the previous matrix-based PCA technique, two-dimensional PCA (2DPCA), is only one special form of the proposed method. On the other hand, the features extracted by the developed method may have a large number of data items (each real number in the obtained features is called one data item). In order to obtain features with a small number of data items, we have devised a two-step feature extraction scheme. Our experiments show that the proposed two-step feature extraction scheme can achieve a higher classification accuracy than the 2DPCA and PCA techniques.

1. Introduction

Bimodal approaches to biometric recognition tasks, that is, approaches that combine two biometric traits to perform personal authentication, have been found to produce better results than single biometrics using a single trait alone [1–10]. For example, combination use of voice and face [1], of face and fingerprint [3,7,11], of ear and face [4], of 3D face and hand [5], of hand shape and texture [6], of both right and left hands with the natural layout [8], of different color components of ear images [12], depth and intensity information of faces [13], facial appearance and motion [14], 2D and 3D palmprint features [15] all illustrates that combination of two biometric traits is better than the sole use of either of them. Combination of multimodal biometrics including bimodal approaches can be categorized into three classes, i.e. feature level fusion [16], matching score level fusion and decision level fusion [17–19] in terms of the phase where the fusion operation occurs. To a great extent, biometrics can be referred to as one kind of image-based classification problems. A number of other applications such as image retrieval [20,21] and character recognition [22] can be also partially considered as image-based classification problems. Many algorithms such as SVM, the kernel method and the fuzzy algorithm [20,21,23] have been applied to these problems. In both bimodal and single-trait biometric approaches, traits are usually initially represented as high-dimensional images with dimension reduction techniques being applied to the images prior to authentication in order to represent them using lower-dimensional data. Principal component analysis (PCA) technique has been widely used to reduce image dimensionality but PCA was originally developed for one-dimensional data [24–28]. As a result, when we apply PCA to images, it is first necessary to transform the image matrices into their corresponding one-dimensional vectors. This usually leads to inaccurate estimation of the covariance matrix [27]. A further drawback of using PCA for image-based applications is that it is computationally expensive.

A noticeable extension of the PCA technique is two-dimensional PCA (2DPCA), which directly extracts features from image matrices [27,29–31]. Rather than projecting a corresponding vector, this technique projects an image matrix onto one transforming axis to obtain one image feature in the form of vector. The 2DPCA can estimate the covariance matrix more accurately than PCA [27]. Another extension of the PCA technique is complex-PCA [32] which uses a complex vector to represent information from the sample of one subject. Such denotations can be used to produce a complex covariance matrix, i.e. the generative matrix of the complex-PCA based feature extraction procedure. Complex-PCA then takes the eigenvector of the generative matrix as a transforming axis and

* Corresponding author. Tel.: +852 2766 7271; fax: +852 2774 0842.
E-mail addresses: laterfall@hitsz.edu.cn (Y. Xu), csdzhang@comp.polyu.edu.hk (D. Zhang), yangjy@public1.ptt.js.cn (J.-Y. Yang).

0031-3203/$ - see front matter © 2009 Elsevier Ltd. All rights reserved.
doi:10.1016/j.patcog.2009.09.013

projects onto it the complex vector denoting the sample of a subject, thereby obtaining a feature (a complex number) of the complex vector. The feature extraction results of complex-PCA represents subjects well in a low-dimensional space but, being derived from PCA, it shares its drawbacks such as being computationally expensive. There are other feature extraction methods related to PCA. Fisher discriminant analysis (FDA), for example, can be implemented as two successive PCA processes [33]. Kernel PCA (KPCA) is also based on the PCA methodology. In some extent, kernel fisher discriminant analysis (KFDA) [34–36] is also related to PCA. In each case, when applied to images, these methods except for 2DPCA should first transform the image matrices into their corresponding one-dimensional vectors. We also note that two new improvements to PCA, the incremental learning algorithm of bidirectional principal component analysis (BDPCA) [37] and the class-augmented PCA [38] have recently been proposed, respectively. The proposed incremental learning algorithm allows PCA to be computationally efficiently implemented in the case where new training samples arrive at any time. The class-augmented PCA allows the class information to be encoded, producing features more appropriate for classification.

This paper proposes matrix-based complex PCA (MCPCA), a novel mathematically tractable matrix-based feature extraction method for bimodal biometrics. MCPCA first denotes bimodal biometric traits such as the left and the right palmprint trait using a complex matrix and then extracts features from the complex matrix. The new method offers a unified framework for previous matrix-based PCA techniques and is supported by a theory which promises that the technique will be optimal among the class of methods capable of directly extracting features from complex matrices. We tested MCPCA on a number of biometric recognition applications and compared it with both 2DPCA and PCA and found that it classifies more accurately than 2DPCA and can also obtain low-dimensional features that represent subjects well. This paper has the following main contributions to the area of biometrics. The first is that it proposes for the first time the matrix-based complex PCA technique for biometric fusion, which provides readers with a novel feature extraction technique. The second contribution is that the paper thoroughly analyzes and presents the theoretical foundation of the proposed method. From the viewpoint of applications, the proposed method is significant in the following two ways: First, it enables us to directly deal with bimodal biometrics using a simple matrix-based method. Second, the proposed method extends the principal component analysis methodology. Moreover, experimental results show that the proposed method is a valid and attractive technique for implementing bimodal biometrics. On the other hand, the proposed method does require that the two matrices representing the two biometric traits should be of the same dimensions. In real-world applications this requirement may not be satisfied. However, as shown in the experiments, it is possible to overcome the problem by resizing one of the two matrices or cropping larger images. Another potential disadvantage of the proposed method is that naive MCPCA might produce more data items than other feature extraction methods. However, as shown in Section 4, the proposed two-step MCPCA can overcome this disadvantage.

The remainder of the paper is organized as follows. In Section 2 we present the matrix-based complex PCA technique. In Section 3 we describe its theoretical properties. In Section 4 we describe a two-step feature extraction scheme for MCPCA. In Section 5, we describe some experiments. Section 6 offers our conclusion.

2. MCPCA: a matrix-based complex PCA technique

In this section, we propose and formulate our novel feature extraction method, matrix-based complex PCA (MCPCA), so called since it appears to be formally analogous to PCA and is based on complex matrices. Suppose there are two matrices representing two different biometric traits of one subject. Further, assume that the two matrices are of the same dimensions. If the two matrices associated with the k-th subject are A_k and B_k, we can define a complex matrix C_k as $C_k = A_k + iB_k$. Hence, C_k can denote the k-th subject. The complex matrix allows us to integrate two representations of one subject into a simple denotation. This approach is especially suitable for real-world applications such as palmprint-based personal identification where the use of complex matrices allows subjects to be more reliably identified by using the information from two palmprints (one from each palm) rather than from just one palm as in most palmprint authentication approaches.

The basic idea of MCPCA is that feature extraction of a complex matrix is performed by projecting the complex matrix rather than the corresponding vector onto complex vectors (i.e. transforming axes). As a result, for a complex matrix, one transforming axis produces one feature in the form of a vector. If multiple transforming axes are used, we will obtain as many features, in the form of vectors, as there are transforming axes. As presented later, transforming axes are eigenvectors of the generative matrix of MCPCA.

In feature extraction, a feature is the result produced by projecting the sample onto the transforming axis. If C is a complex matrix denoting one bimodal biometric sample of a subject and X is a transforming axis in the form of a complex vector, then the feature extraction result of C with regard to X will be $Y = CX$. Because C and X are respectively a complex matrix and complex vector, Y is also a complex vector. Suppose that the complex matrix C has the dimension of $m \times n$ and the complex vector X has the dimension of $n \times 1$, then Y has the dimension of $m \times 1$. The matrix presented in formula (1) can be used as the generative matrix of MCPCA.

$$G = E(C^H C). \tag{1}$$

We regard the eigenvectors of this matrix as transforming axes. In this paper, the superscript H denotes the conjugate transpose of a complex matrix. This means $C^H = \overline{C}^T$. In practice, G can be evaluated by $G = \sum_{k=1}^{N} C_k^H C_k / N$, where N is the number of complex matrices. The total number of available eigenvectors is the same as the number of columns of the complex matrix. Suppose that the dimension of every complex matrix is $m \times n$, then there are a total of n available transforming axes. We can select $t (t \le n)$ transforming axes from the n transforming axes. Since each transforming axis can produce a feature in the form of a vector for a complex matrix by projecting the complex matrix onto the transforming axis, we can obtain t features in the form of vector, which can form a novel lower dimensional $m \times t$ matrix. We can also say that the t transforming axes transform the $m \times n$ complex matrix into an $m \times t$ complex matrix. An issue closely related to this is the matter of deciding which eigenvectors should be used as the transforming axes. Indeed, we should select the eigenvectors corresponding to the largest eigen-values as transforming axes to extract features of samples, as is done in the implementation of PCA.

MCPCA can be regarded as a unified framework of matrix-based PCA techniques and within this framework we can present 2DPCA as a special form of MCPCA. This is because if the imaginary parts of complex matrices representing sample data are null, MCPCA will be degraded and identical to 2DPCA as proposed in [27].

As we know, GLRAM [39], 2DLDA [40,41] and BDPCA [42] also have been used to directly extract features from matrices. 2DLDA is an extension of linear discriminant analysis (LDA) developed for one-dimensional vector data. GLRAM can be viewed as a more general technique on matrix decomposition. BDPCA is similar to MCPCA as follows: both of them are based on the principal component analysis methodology and use two procedures to

extract features from matrix data [42]. On the other hand, BDPCA and MCPCA are proposed for real matrix samples and complex matrix samples, respectively. MCPCA is different from 2DLDA in the following two aspects: first, MCPCA is indeed an unsupervised method in which the information of class label of samples is not exploited in the training phase whereas 2DLDA is a supervised method. Similar to LDA, when exploiting the statistical information of training samples to produce the projecting axis, 2DLDA also take into account the class label of training samples. Second, MCPCA is suitable for complex-matrix data whereas 2DLDA is developed for real matrices. MCPCA is different from GLRAM in the following two aspects: first, GLRAM has no closed-form solution and it usually needs a high computational cost whereas MCPCA has the closed-form solution and can be implemented easily. Second, besides MCPCA and GLRAM are proposed for complex-matrices and real matrices, respectively; they have different motivations. While MCPCA projects a complex matrix onto a complex vector to produce the resultant vector, it requires that the resultant vector can reconstruct the complex matrix with the minimum error. However, GLRAM attempts to use the product of three matrices L, M_i, R to approximate a real matrix A_i and requires that there be the minimum approximate error and both L and R be matrices with orthogonal columns.

3. Theoretical foundation of MCPCA

In this section, we consider the theoretical properties of MCPCA and show that MCPCA is optimal among the class of techniques that are able to directly extract features from complex matrices.

Property 1. *G is a Hermitian matrix.*

Proof. According to (1), it is clear that $(G)^H = G$. Hence, G is a Hermitian matrix. Furthermore, the following Property 2 also holds. □

Property 2. *All the eigen-values of G defined by* (1) *are real numbers.*

As we know, the covariance matrix of PCA has only non-negative real eigen-values and the eigenvectors of the covariance matrix of PCA are used as transforming axes for feature extraction. It is possible to use the eigen-values to evaluate the effectiveness on feature extraction of the corresponding eigenvectors. Indeed, it is not hard to prove that, for PCA, when selecting the eigenvectors of the covariance matrix as transforming axes, the corresponding eigen-values should be as large as possible. The following theorem presents the details of a similar principle for MCPCA.

Theorem 1. *For MCPCA-based feature extraction, the eigenvector of G corresponding to the largest eigen-value is the top-priority transforming axis. If we need t transforming axes to transform the samples into a new space, the best candidates are the t eigenvectors of G associated with the first t largest eigen-values, as they have the lowest reconstruction error.*

Proof. Suppose that the complex matrix C can be accurately expressed by $C = \sum_{i=1}^{s} Y_i X_i^H$ with the orthogonal constraint of $X_i^H X_j = 0 (i \neq j)$, $X_i^H X_j = 1 (i = j)$, where Y_i and X_i are both complex vectors. Right multiplying $C = \sum_{i=1}^{s} Y_i X_i^H$ by X_i allows us to obtain $CX_i = \sum_{i=1}^{s} Y_i X_i^H X_i$. Because of the orthogonal constraint condition on X_i, we further have $Y_i = CX_i$. Let $\hat{C} = \sum_{i=1}^{t} Y_i X_i^H$, $t \leq s$. Here s is defined as follows: if the size of the matrix is $m \times n$, s will be numerically equivalent to n. This also means that MCPCA can produce at most n transforming axes and n "features".

$\hat{C}$ can be constructed after X_i, Y_i are obtained, so $\hat{C}$ is called a reconstruction matrix of C. We can formulate the mean-square error between C and $\hat{C}$ as follows: $E(\|C - \hat{C}\|_F^2) = E[tr((C - \hat{C})(C - \hat{C})^H)]$, where $\|\cdot\|_F^2$ denotes the square of the Frobenius norm of a matrix. $E(\|C - \hat{C}\|_F^2)$ is also called the squared reconstruction error. Due to the orthogonal constraint of X_i and the formula $Y_i = CX_i$, we have $E(\|C - \hat{C}\|_F^2) = \sum_{i=t+1}^{s} X_i^H G X_i$, $G = E(C^H C)$. Obviously, the optimal X_i should minimize $E(\|C - \hat{C}\|_F^2)$ and satisfy the above orthogonal constraint of X_i. We solve this minimization problem using the Lagrange multiplier method. Defining the Lagrangian function $L = \sum_{i=t+1}^{s} X_i^H G X_i - \lambda_i (X_i^H X_i - 1)$ and requiring the derivative of L with regard to X_i to be zero yield $GX_i = \lambda_i X_i$. This implies that if we require $E(\|C - \hat{C}\|_F^2)$ to reach its extremes, the corresponding X_i must be the eigenvectors of G. Note that since $GX_i = \lambda_i X_i$ implies $X_i^H G X_i = \lambda_i X_i^H X_i = \lambda_i$, we have $E(\|C - \hat{C}\|_F^2) = \sum_{i=t+1}^{s} \lambda_i$. Assume that the eigen-values of G are $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n \geq 0$ and the corresponding eigenvectors are $X_1, X_2, \ldots, X_n$. If $X_1, X_2, \ldots, X_t$ are selected as t transforming axes, then the reconstruction error will be $\sum_{i=t+1}^{n} \lambda_i$, which is exactly the minimum construction error for feature extraction using t transforming axes. Thus, Theorem 1 is certain.

Theorem 1 can also be intuitively illustrated. Taking the eigenvectors corresponding to the largest eigen-values as transforming axes also implies that the obtained features capture the data components that most vary. In other words, we might say either that the corresponding features of different original samples will exhibit large statistical differences or that these transforming axes can capture the most representative information from different original samples. As a result, the features obtained using these transforming axes are able to reconstruct the original samples with only a small error. We might also note that the feature components that the principal component analysis methodology produces are statistically not correlated. Thus, PCA is also regarded as a decorrelation technique.

Since the generative matrix of MCPCA is generated from the complex matrix consisting of two real matrices, the generative matrix and feature extraction result appear to be statistically associated not only with the information from the two real matrices themselves but also with their cross-bred information. This can be formally demonstrated as follows: Let A, B be two real matrices of the same dimensions. If C is a complex matrix and defined as $C = A + iB$, then we have $C^H C = (A + iB)^H (A + iB) = (A^T - iB^T)(A + iB) = A^T A + B^T B + i(A^T B - B^T A)$. $A^T B$ and $B^T A$ represent the cross-bred information of A, B. Usually $A^T B - B^T A$ is not equal to zero, so we say that $C^H C$ contains the cross-bred information of A, B. Since we use A, B to denote two biometric traits and the generative matrix of MCPCA is $G = E(C^H C)$, transforming axes generated from MCPCA i.e. eigenvectors of G will be statistically directly related not only to the two biometric traits themselves but also to the cross-bred information generated from them. As a result, the feature extraction result will be statistically associated with the cross-bred information mentioned above since it is obtained by projecting the complex matrix denoting the sample of a subject onto the transforming axis of MCPCA. □

4. Two-step MCPCA to produce features with few data items

One of the goals of feature extraction is the reduction of the dimensionality of original samples. However, one difficulty with

MCPCA as presented to this point is that the features it obtains may contain more data items than those obtained using traditional PCA and as a result classification using MCPCA will be less efficient than PCA classification. The reason for this is as follows. Referring to each real number in the obtained features as one data item, we note that one real number consists of one data item whereas one complex number consists of two data items. Each feature obtained using MCPCA will be one complex vector consisting of a number of complex numbers, whereas every feature obtained using PCA is just a real number. So, if MCPCA and PCA extract an equal number of features from one sample, MCPCA features will contain many more data items than PCA features and classification of MCPCA-features will be correspondingly less efficient.

To improve the efficiency of classification of MCPCA-features we apply an additional feature extraction process. As noted in Section 3, when we extract a feature for a complex matrix using a transforming axis, the feature extraction result will be a complex vector. If $t(t \leq n)$ transforming axes are used for feature extraction, we will obtain t complex vectors. These t complex vectors can form an $m \times t$ complex matrix. As a result, we say that MCPCA transforms a bimodal biometric sample represented by an $m \times n$ complex matrix into an $m \times t$ complex matrix. We call these $m \times t$ complex matrices feature matrices and call the corresponding feature extraction the first feature extraction process. We can also carry out another feature extraction process (the second feature extraction process) to further transform the obtained $m \times t$ feature matrices into a lower-dimensional $q \times t(q \leq m)$ complex matrix. This is presented as follows. We denote B_i as the feature matrix of the i-th training sample. A complex matrix P is defined as $P = \sum_{i=1}^{N} B_i B_i^H / N$. Obviously, B_i is an m by t complex matrix and P is an m by m complex matrix. We regard P as the generative matrix and solve for its eigenvectors and eigen-values. Selecting a number of eigenvectors of P as transforming axes, we can carry out feature extraction for all the B_i. When selecting transforming axes, the corresponding eigen-values should be as large as possible. After the second feature extraction process is implemented, the $m \times t$ feature matrix B_i is further transformed into a $q \times t(q \leq m, t \leq n)$ matrix, where q is the number of the transforming axes used in the second feature extraction process. Hereafter we refer to feature extraction using both the first and second feature extraction processes as two-step MCPCA and refer to the first feature extraction process as naive MCPCA.

When we use the features obtained using two-step MCPCA to classify samples, we can classify more efficiently because these features have fewer data items than those obtained using naïve MCPCA. The computational complexities of the classification processes associated with naive MCPCA and two-step MCPCA can be expressed as $O(mt)$ and $O(qt)$ respectively, where $q \leq m$. Moreover, as presented in Section 5, when producing the best performance two-step MCPCA obtains the features that include fewer data items than the features obtained using 2DPCA. This also means that two-step MCPCA can classify more efficiently than 2DPCA.

PCA, 2DPCA and MCPCA all should solve the eigenvectors of the corresponding generative matrices. Indeed, when implementing the above methods, the main computational burden is caused by the solving of the eigenvectors. For different methods, we analyze the computational complexity of solving the eigenvectors as follows. We assume that each image is an $m \times n$ matrix, then for PCA using bimodal biometric traits, the generative matrix is $2mn \times 2mn$, and the computational complexity of solving the eigenvectors is $O(m^3 n^3)$. As for complex-PCA, since the generative matrix is an $mn \times mn$ complex matrix, the computational complexity of solving the eigenvectors is $O(m^3 n^3)$. Thus, it is clear that MCPCA is computationally more efficient than complex-PCA. For a general 2DPCA method using a single modal biometric image, since the generative matrix has the dimension of $n \times n$, the computational complexity of solving the eigenvectors is $O(n^3)$. For example, this is the computational complexity of left 2DPCA and right 2DPCA (Section 5.1) when solving the corresponding eigenvectors. Since score 2DPCA first implements the 2DPCA method for each of the two biometric traits, it also has the computational complexity $O(n^3)$ when solving the eigenvectors of the generative matrix. For combination 2DPCA, because two $m \times n$ biometric images are combined to form an $m \times 2n$ matrix and the generative matrix is a $2n \times 2n$ matrix, the computational complexity of solving the eigenvectors is also $O(n^3)$. As for MCPCA, when solving the eigenvectors of the corresponding generative matrices, the first and the second feature extraction processes have the respective computational complexities of $O(n^3)$ and $O(m^3)$. Thus, PCA using bimodal biometric traits and complex-PCA have the highest computational complexity. If there is only a small difference between m and n, the computational complexity of MCPCA will be similar to that of combination 2DPCA, score 2DPCA, and a general 2DPCA method using a single modal biometric image. We can conclude that when solving the eigenvectors of the generative matrix, MCPCA is much more efficient than PCA using bimodal biometric traits and its computational complexity is similar to that of combination 2DPCA, score 2DPCA and 2DPCA using a single biometric trait.

In summary, the second feature extraction process has the following rationales: First, it allows us to use less data to represent the original bimodal biometric traits and consequently the classification procedure will take a shorter time. Second, experimental results (presented later) show that generally the second feature extraction process may produce a higher accuracy. In addition, it should be pointed out that similar to the two feature extraction processes of BDPCA, the use of two-step MCPCA allows us to capture the major variations of original data (matrices) in the horizontal and vertical direction, respectively.

The two feature extraction processes have the following similarities: The first and the second feature extraction processes of MCPCA obtain the data components that most vary in, respectively, the horizontal and vertical directions. Capturing data components that most vary is the essence of the principal component analysis methodology. In addition, correlation not only exists in the vertical direction but also exists in the horizontal direction of the bimodal biometric sample. The first extraction process can thus eliminate correlation in the horizontal direction and the second feature extraction process can eliminate correlation in the vertical direction.

There are however, a number of differences between the first and second feature extraction processes. First, the sample for the first feature extraction process is an $m \times n$ complex matrix, whereas for the second feature extraction process it is an $m \times t$ complex matrix. Second, the second feature extraction process allows the result of the first feature extraction process to be further transformed into fewer data items. The main steps of two-step MCPCA are summarized in the flowchart in Fig. 1.

It should be pointed out that classification accuracy can also be affected if the second feature extraction process excessively reduces the number of dimensions, rendering the features no longer able to effectively represent the original bimodal biometric sample. This is similar difficulty to that faced by the traditional principal component analysis technique. While a suitable number of principal components that capture the major variations of the data can effectively represent the original data, too few principal components will leave major variations unrepresented.

5. Experiments and results

We tested MCPCA on a palmprint-based personal identification application (Section 5.1), with bimodal biometrics using

Step 1. Pair one kind of $m\times n$ biometric image with another kind of $m\times n$ biometric image to create a bimodal biometric sample. Suppose each image is $m\times n$.

Step 2. Denote the bimodal biometric samples using $m\times n$ complex matrices and calculate the generative matrix of all the complex matrices.

Step 3. Solve the t eigenvectors of the generative matrix corresponding to the first t largest eigenvalues.

Step 4. Perform feature extraction for each bimodal biometric sample by projecting it onto the obtained t eigenvectors. This process is the first feature extraction process of MCPCA. The features of each bimodal biometric sample form an $m\times t$ complex matrix.

Step 5. Perform the second feature extraction process for each bimodal biometric sample. This allows an original bimodal biometric sample represented by an $m\times n$ complex matrix to be finally transformed into a $q\times t$ complex matrix.

Step 6. Carry out the classification procedure for test samples.

Fig. 1. The main steps of two-step MCPCA.

palmprint and face images (Section 5.2), and with bimodal biometrics using ear and face images (Section 5.3). In each experiment, training samples (in the form of real or complex matrices) were centralized in advance. That is, the mean of all the samples denoted by real or complex matrices were subtracted from each of the samples and the samples were then newly denoted using the subtraction results. We should point out that since MCPCA is derived from the principal component analysis methodology, the focus of experimental comparison is mainly on the comparison between the proposed feature level fusion method i.e. MCPCA and other fusion schemes including the matching score level fusion scheme and other feature level fusion schemes on the basis of a principal component analysis technique. This confines us to making comparisons between the same kind of techniques integrated with different fusion schemes, producing a more persuasive comparison.

We implement PCA for bimodal biometric traits as follows: we first transform each image matrix into a one-dimensional vector. Suppose that each matrix is $m\times n$. The corresponding one-dimensional vector then has mn elements. We combine the two vectors associated with each set of paired images (bimodal biometric traits) to form a $2mn$-dimensional vector. Finally we calculate the generative matrix of these $2mn$-dimensional vectors denoting bimodal biometric traits and implement PCA. The generative matrix is a $2mn\times 2mn$ matrix, which could be very large. For example, if each image is denoted by a 128×128 matrix, the generative matrix of PCA will be a 32 768 by 32 768 matrix. In this section, we also compare MCPCA and complex vector PCA in [32] which exploits both left and right palmprint images as a complex one-dimensional vector.

5.1. Experiment on palmprint-based personal identification

We compared the performance of MCPCA with that of 2DPCA in four ways: using only left palmprint images (left 2DPCA); using only right palmprint images (right 2DPCA); using both left and right palmprint images as a combined real matrix (combination 2DPCA); and using the fusion information at the matching score level (score combination). Combination 2DPCA involves the direct combination of two $m\times n$ image matrices denoting bimodal biometric traits of one subject to form a new real $m\times 2n$ matrix and then implements 2DPCA using these $m\times 2n$ matrices. Score combination works as follows: 2DPCA is first applied to extract features from left and right palmprint images, respectively. Then the matching scores of the left and right palmprint features are calculated and combined by the simple sum rule to produce the final matching score to be used in personal identification. We note that the score fusion solution in [43] has some advantages. However, as pointed out by the authors of this reference, the proposed solution is only a principled and general approach that is optimal when the genuine and impostor matching score distributions are either known or can be estimated with high accuracy. As for our experiments, it seems that the limited sample number and the high-dimensional features are not advantageous for the accurate estimate of the matching score distribution at all, so we do not adopt the score fusion solution in the mentioned reference. These experiments used the nearest neighbor classifier to classify the feature extraction results generated from all the methods except for score combination. Score combination classifies a test sample as follows: It first identifies the training sample that has the largest final matching score with respect to this test sample and considers that the test sample is from the same class (identity) as the identified training sample.

The palmprint database is from the Hong Kong Polytechnic University and contains palmprint images from 189 subjects. Each subject provided 10 left palm images and 10 right palm images [44]. Palmprint subimages of 128×128 were extracted from the original images [45]. For high computational efficiency, we used palmprint images of all the subjects. Fig. 2 shows three original palmprint images and the reconstructed images obtained using naïve MCPCA.

The training samples for left 2DPCA are two arbitrarily selected left palmprint images. The eight remaining left palmprint images from each of the subjects are taken as test samples of left 2DPCA. Right 2DPCA is treated in the same way except that it uses the images of the right palm. To implement MCPCA, combination 2DPCA, and score combination, we pair the left palmprint images with right palmprint images of each subject in sequence, the first left palmprint image of every subject with the first right palmprint image of the same subject and so on. We refer to a left palmprint image and the paired right palmprint image as one bimodal sample. Two bimodal samples are arbitrarily selected as training samples for MCPCA, combination 2DPCA, and score combination. The remaining bimodal samples are used as test samples of these methods. Experiments are conducted on all the training sets and the corresponding testing sets.

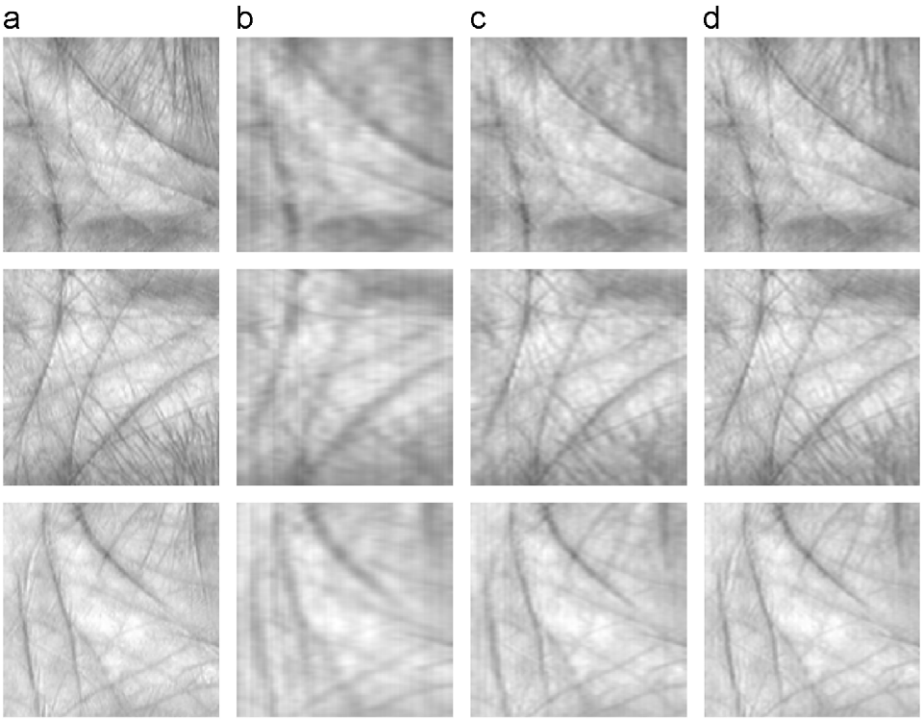

Fig. 2. Three palmprint images and the reconstructed images obtained using naïve MCPCA. The images in column (a) are the three original palmprint images. The first two are images from the left and right palmprints of the same person and the third image is from another individual. The images in columns (b), (c), and (d) are reconstructed images based on respectively the first 20, 40, and 60 transforming axes generated from naïve MCPCA.

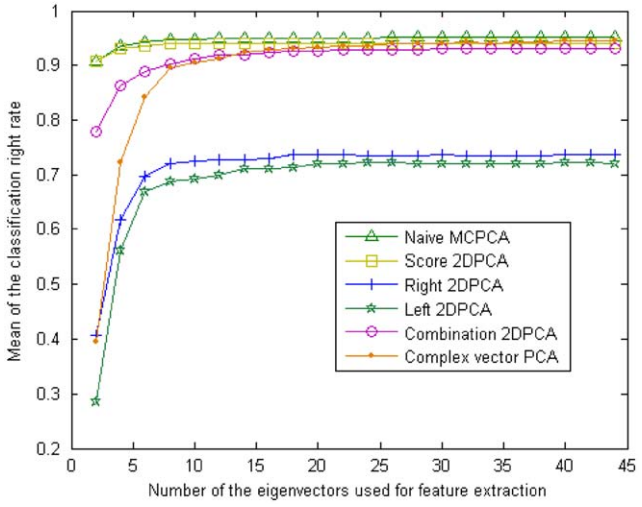

Fig. 3. Variation of the mean of the classification right rate of each of different feature extraction methods in relation to the number of the eigenvectors used in feature extraction. While the vertical coordinate axis shows the mean of the classification right rate, the horizontal coordinate axis shows the number of the eigenvectors used in feature extraction.

Table 1
The minimum values of the classification error rates on the palmprint-based personal identifications using different feature extraction methods and the dimension of the corresponding features extracted from one sample.

Methods	Minimum classification error rate (Best performance) (%)	Dimension of the corresponding features of one sample
Two-step MCPCA	4.0	1 by 16 complex matrix (a complex vector having 16 entries)
Naive MCPCA	4.1	128 by 15 complex matrix
Left 2DPCA	15.5	128 by 75 real matrix
Right 2DPCA	17.6	128 by 79 real matrix
Combination 2DPCA	5.9	128 by 21 real matrix
Score combination	5.4	One 128 by 11 matrix and another 128 by 19 matrix
Complex vector PCA	4.8	Complex vector with 60 data items
PCA	5.0	Vector with 31 data items

Since score combination simultaneously exploits the left and right palmprint image for authentication and the features of each image are obtained using 2DPCA, the features associated with score combination consist of two feature matrices. The features associated with any other matrix-based feature extraction method consist of one feature matrix.

Fig. 3 shows the variation of the mean of the classification right rate of each of the feature extraction methods in relation to the number of the eigenvectors used in feature extraction. Naïve MCPCA clearly outperforms combination 2DPCA, score combination and complex vector PCA notwithstanding combination 2DPCA, score combination and complex vector PCA also use information from both the left and right palmprints.

Table 1 shows the minimum classification error rate for each feature extraction method. Two-step MCPCA has a minimum error rate of 4.0%, smaller than any method. Table 1 also shows that the best classification result for two-step MCPCA corresponds to lower-dimensional features than left 2DPCA, right 2DPCA, combination 2DPCA, and score combination. The 5.0% minimum error rate of PCA occurs when the feature of one sample is in the form of a vector with 31 data items. It appears that when representing the original samples PCA can use fewer feature components than the other methods shown in the table. On the other hand, the transforming axes of PCA-based feature extraction should be achieved by solving a very high-dimensional eigenvalue problem. In this experiment, the generative matrix of PCA is a $32\,768 \times 32\,768$ matrix. Obviously this problem will involve a heavy computation and memory burden. In contrast, the

generative matrix of naïve MCPCA is just a 128 by 128 complex matrix and the generative matrix of the second feature extraction process in two-step MPCA is of a lower dimension. Thus, MCPCA is able to compute transforming axes more efficiently than PCA.

5.2. Experiment on bimodal biometrics using face and palmprint images

In the absence of an already existing database of bimodal face and palmprint biometrics, we simulated bimodal face and palmprint biometrics by using the AR face database and the left palmprint images described in Section 5.1. The AR face database includes more than 100 subjects and more than 4000 face images showing faces with different facial expressions, in varying lighting conditions and occluded in several ways. We selected the first ten non-occluded images of the first 120 individuals. The face portion of each image was cropped and normalized to form a new image of 50×40 pixels [27]. Each left palmprint image was then cropped to obtain a new palmprint image of the same size as the face images. Fig. 4 shows three cropped face images and the corresponding reconstructed images obtained using naïve MCPCA.

We created a virtual subject with bimodal biometric traits by pairing the face images selected from the AR database with left palmprint images in the palmprint database in sequence, the first face image with the first left palmprint image and so on. Each virtual subject thus possesses ten paired bimodal biometric traits, i.e. a face image and a left palmprint image. Creating virtual subjects in this way would appear to be a reasonable way to simulate a bimodal biometrics database. The all available 120 virtual subjects were used to test different methods using bimodal biometric traits.

MCPCA, combination 2DPCA, and score 2DPCA identify 120 virtual subjects using their respective procedures presented in Section 5.1 and the information of both the face images and left palmprint images, whereas "AR 2DPCA" uses only face images and "left palmprint 2DPCA" uses only left palmprint images to identify 120 real subjects. In the implementations of MCPCA, PCA, score 2DPCA, weighted score 2DPCA, and combination 2DPCA, the first two paired bimodal traits of each virtual subject are taken as training samples and the others are used as test samples. When "AR 2DPCA" is carried out, the first two face images of the corresponding 120 real subjects are used as training samples and the others are regarded as test samples. The implementation of "left palmprint 2DPCA" takes the first two left palmprint images of the corresponding 120 real subjects as training samples and the remainder as test samples. "Weighted score 2DPCA" is referred to as the score fusion method using the weighted sum rule rather than the simple sum rule in Score 2DPCA. Weighted score 2DPCA exploited the authentication accuracy of the biometric trait to set the weight. Actually, in this section the weight was set to the ratio of the accuracy obtained using "left palmprint 2DPCA" or "AR 2DPCA" to the sum of the two accuracies of "left palmprint 2DPCA" or "AR 2DPCA".

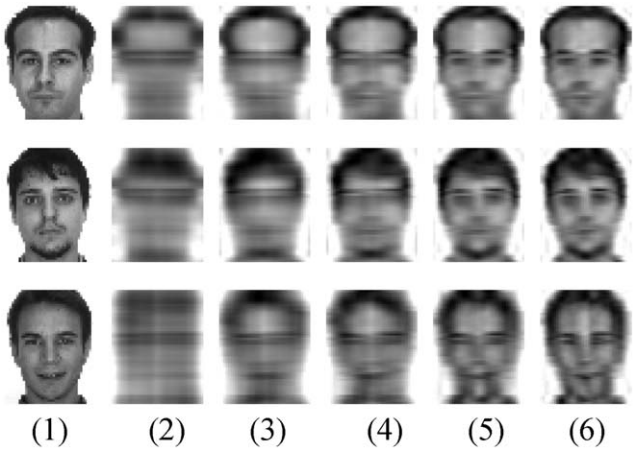

Fig. 4. Three original face images from three bimodal traits and the reconstructed face images obtained using naïve MCPCA. The first column shows the three original face images, and the second to the sixth columns show the reconstructed face images which are constructed respectively using 2, 4, 6, 8, 10 eigenvectors of the generative matrix of naïve MCPCA.

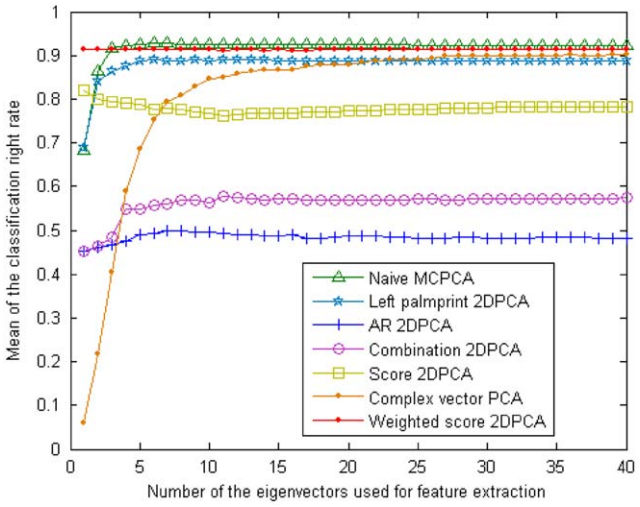

Fig. 5. Variation of the mean of the classification right rate of each of different feature extraction methods with the number of the eigenvectors used in feature extraction. While the vertical coordinate axis shows the mean of the classification right rate, the horizontal coordinate axis represents the number of the eigenvectors used in feature extraction.

Table 2
The minimum values of the classification error rates obtained using different methods and the dimension of the corresponding features extracted from one sample.

Methods	Minimum classification error rate (Best performance) (%)	Dimension of the corresponding features of one sample
Two-step MCPCA	7.5	A 8 by 12 complex matrix
Naive MCPCA	8.1	A 50 by 8 complex matrix
AR 2DPCA	40.3	A 50 by 7 real matrix
Left palmprint 2DPCA	10.4	A 50 by 9 real matrix
Combination 2DPCA	8.3	A 50 by 11real matrix
PCA	8.9	A vector with 30 data items

Fig. 5 shows the variation of the mean of the classification right rate of each of different feature extraction methods against the number of the eigenvectors used in feature extraction. Table 2 shows the minimum classification errors of different methods. MCPCA achieves a lower error rate than combination 2DPCA, score 2DPCA, weighted score 2DPCA, left palmprint 2DPCA and AR 2DPCA.

Table 2 shows that two-step MCPCA is able to obtain lower-dimensional features for original samples than naive MCPCA. Table 2 also tells us that the minimum error obtained using two-step MCPCA is lower than that obtained using the other methods.

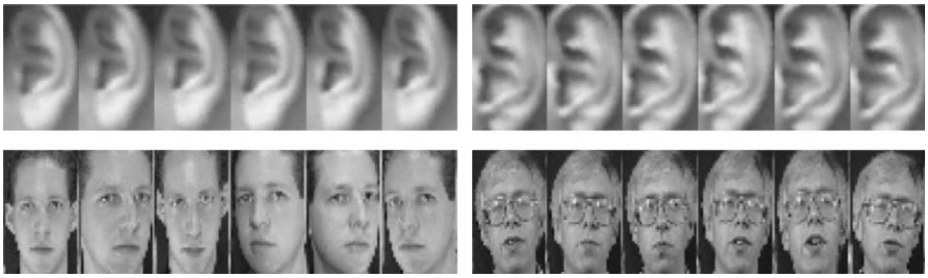

Fig. 6. The cropped ear images of two subjects and the used face images of two persons from ORL. Ear images and face images are shown in the first and second rows, respectively.

5.3. Bimodal biometric experiment based on ear and face images

We simulated a bimodal biometric database using an ear image database and the ORL face database. The ear database is made up of six ear images collected from each of 17 subjects for a total of 102 images [46]. This database was downloaded from the internet [47]. The downloaded images are of different sizes and so were all cropped to 53×32 pixels. The cropped ear images of two subjects are shown in the first row of Fig. 6.

The ORL database [48] includes 400 face images from 40 subjects. The images include variations in facial expression (smiling/not smiling, open/closed eyes) and facial detail. The subjects are in an upright, frontal position with tolerance for some tilting and rotation of up to 20°. Each of the face images contains 112×92 pixels. Only the first six face images of the first 17 subjects of the ORL database are involved in this experiment. We resized these face images to obtain new ones of the same size as the cropped ear images as follows: first, every image was cut to 106×92 pixels by preserving the central pixels in the original image and discarding the remaining 6×92 pixels located on the boundary of the original image. These images of 106×92 pixels were then reduced to 53×32 pixels by using the following procedure: The first and last columns of each image were recorded as the first column and the last columns (i.e. the 1st and 32th columns) of a new face image. We then resized the remaining image region (106×90 pixels) to 53×30 pixels and obtained final new face images of 53×32 pixels to be used in personal identification.

We also created virtual subjects to simulate bimodal biometrics using the ear and face images. We paired the cropped ear images to the final face images in sequence, the first ear image with the first face image, and so on. Each of 17 virtual subjects thus possesses 6 bimodal biometric traits, each consisting of one ear image paired with one face image. In this experiment, "ear 2DPCA" means the 2DPCA method using only ear images, while "ORL 2DPCA" means the 2DPCA method using only ORL face images. Combination 2DPCA involves the direct combination, as presented in Section 5.1, of ear and face images. Score 2DPCA identifies personal identities using the sum of the matching scores of "ear 2DPCA" and "ORL 2DPCA". In this section "weighted score 2DPCA" is also referred to as the score fusion method using the weighted sum rule and also exploited the authentication accuracy of the biometric trait to set the weight. Actually, in this section the weight was set to the ratio of the accuracy obtained using "ear 2DPCA" or "ORL 2DPCA" to the sum of the two accuracies of "ear 2DPCA" and "ORL 2DPCA". In implementing naïve MCPCA, two-step MCPCA, combination 2DPCA, score combination and weighted score 2DPCA, we arbitrarily selected two paired bimodal traits from each virtual subject as training samples and treated the others as test samples. As a result, there were 15 possible training sample sets and the corresponding 15 test sample sets. Experiments on all the 15 cases were conducted for all the methods.

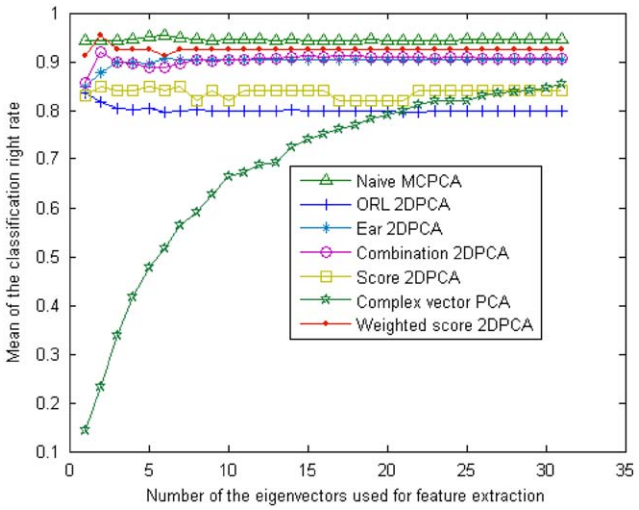

Fig. 7. Variation of the mean of the classification right rates of different feature extraction methods against the number of the eigenvectors used in feature extraction. The *x*-ordinate shows the number of the eigenvectors used for feature extraction. The *y*-ordinate shows the classification accuracies of different feature extraction methods.

Table 3
The minimum values of the classification error rates of different methods and the dimension of the corresponding features of one sample.

Methods	Minimum classification error rate (Best performance) (%)	Dimension of the corresponding features of one sample
Two-step MCPCA	0	A 6 by 6 complex matrix
Naive MCPCA	4.4	A 53 by 6 complex matrix
Ear 2DPCA	5.9	A 53 by 4 real matrix
ORL 2DPCA	10.3	A 53 by 2 real matrix
Combination 2DPCA	4.4	A 53 by 14 real matrix
PCA	2.9	A vector with 16 data items

Fig. 7 shows the mean of the classification right rates of all the methods. It is clear that naive MCPCA has the highest classification accuracy. From Fig. 7, we also know that weighted score 2DPCA can produce a higher accuracy than both "ear 2DPCA" and "ORL 2DPCA", whereas "score 2DPCA" obtains a lower accuracy than "ear 2DPCA". Table 3 shows the minimum values of the classification error rates of two-step MCPCA and other methods. Two-step MCPCA has a lowest classification error rate of 0%. Combination 2DPCA and PCA using the ear and face image have the respective lowest classification error rate of 4.4% and 2.9%. When obtaining its lowest classification error rate,

MCPCA uses the features that consist of just a few data items, which means that it makes only a small demand on memory. This also means that the classification of features obtained using two-step MCPCA will be very computationally efficient.

6. Conclusion and discussion

We have proposed MCPCA, a novel method that uses a complex matrix to denote bimodal biometrics traits. MCPCA is able to directly extract features from complex matrices and has been evaluated by theoretical analysis and experiments. MCPCA can also be seen as providing a unified framework for previously developed matrix-based PCA techniques. It is also mathematically tractable. MCPCA can not only identify more accurately than 2DPCA and PCA, it is particularly suited to extracting the features of image-based bimodal biometric data. The proposed two-step MCPCA usually obtains features that have fewer data items than naïve MCPCA and still well represents the original complex matrices. Moreover, two-step MCPCA can also lead to excellent classification performance. The fact that MCPCA (naïve MCPCA or two-step MCPCA), combination 2DPCA and score combination all perform feature extraction for personal identification using the same bimodal biometric traits but MCPCA can lead to better classification result is attracting. Indeed, when MCPCA is used to integrate two classes of biometric traits to form a complex space, the two parts (the real part and imaginary part) in the complex matrix can be cross-bred to produce new information through feature extraction. As a result, the authentication process exploits not only the bimodal biometric traits themselves but also the cross-bred information for personal authentication.

Since the proposed method should use two matrices with the same dimension to construct the complex matrix, it cannot be directly applied in real-world applications where the images of two biometric traits have different dimensions. However, we can address this problem as follows: In the phase of sample acquisition of the bimodal biometric traits, we may not control the dimension of the sample of each biometric trait. However, given bimodal biometric traits with different dimensions, we can bring the samples of the bimodal traits all to the same dimension by resizing one biometric trait sample or by cropping the trait sample with a larger dimension. For example, one resizing scheme has been proposed for the ear and face fusion experiment as shown in Section 5.3. Another example is shown in Section 5.2, where we crop the palmprint image and obtain a new palmprint image with the same size as the face image. This allows the proposed method to be applied since the cropping or resizing process has made the bimodal biometric traits have the same dimensions. In a specific real-world application, the cropping or resizing scheme might be different from the scheme used in Sections 5.2 and 5.3.

Acknowledgments

We wish to thank National Natural Science Foundations of China under Grants nos. 60602038, 60973098 and 60632050, 863 Program Project under Grant no. 2006AA01Z193, and Natural Science Foundation of Guangdong Province under Grant 06300862, the CERG fund from the HKSAR Government and the central fund from Hong Kong Polytechnic University for supporting.

References

[1] F. Alsaade, A.M. Ariyaeeinia, A.S. Malegaonkar, et al., Enhancement of multimodal biometric segregation using unconstrained cohort normalisation, Pattern Recognition 41 (3) (2008) 814–820.
[2] D. Bouchaffraa, A. Amira, Feature generation and machine learning for robust multimodal biometrics, Pattern Recognition 41 (3) (2008) 775–777.
[3] T. Sim, S. Zhang, R. Janakiraman, S. Kumar, Continuous verification using multimodal biometrics, IEEE Transactions on Pattern Analysis and Machine Intelligent 29 (4) (2007) 687–700.
[4] K. Chang, K.W. Bowyer, S. Sarkar, B. Victor, Comparison and combination of ear and face images in appearance-based biometrics, IEEE Transactions on Pattern Analysis and Machine Intelligent 25 (9) (2003) 1600–1615.
[5] F. Tsalakanidou, S. Malassiotis, M.G. Strintzis, A 3D face and hand biometric system for robust user-friendly authentication, Pattern Recognition Letters 28 (16) (2007) 2238–2249.
[6] A. Kumar, D. Zhang, Personal recognition using hand shape and texture, IEEE Transactions on Image Processing 15 (8) (2006) 2454–2461.
[7] D. Bouchaffraa, A. Amira, Structural hidden Markov models for biometrics: fusion of face and fingerprint, Pattern Recognition 41 (3) (2008) 852–867.
[8] M. Adán, A. Adán, R. Torres, A.S. Vázquez, Biometric verification/identification based on hands natural layout, Image Vision Computing 26 (2008) 451–465.
[9] R. Snelick, U. Uludag, A. Mink, M. Indovina, A.K. Jain, Large-scale evaluation of multimodal biometric authentication using state-of-the-art systems, IEEE Transactions on Pattern Analysis and Machine Intelligent 27 (3) (2005) 450–455.
[10] S. Dinerstein, J. Dinerstein, D. Ventura, Robust multi-modal biometric fusion via multiple SVMs, in: Proceedings of 2007 IEEE International Conference on Systems, Man and Cybernetics, 2007, pp. 1530–1535.
[11] G.L. Marcialis, F. Roli, Luca Didaci personal identity verification by serial fusion of fingerprint and face matchers, Pattern Recognition 42 (2009) 2807–2817.
[12] L. Nanni, A. Lumini, Fusion of color spaces for earauthentication, Pattern Recognition 42 (2009) 1906–1913.
[13] C. Xu, S. Li, T. Tan, L. Quan, Automatic 3D face recognition from depth and intensity Gabor features, Pattern Recognition 42 (2009) 1895–1905.
[14] A. Hadid, M. Pietikainen, Combining appearance and motion for face and gender recognition from videos, Pattern Recognition 42 (2009) 2818–2827.
[15] V. Kanhangad, N. Luo, A. Kumar, D. Zhang, Robust palmprint verification using 2D and 3D features, Pattern Recognition 43 (2010) 358–368.
[16] A. Kong, D. Zhang, M. Kamel, A survey of palmprint recognition, Pattern Recognition 42 (7) (2009) 1408–1418.
[17] D. Zhang, F. Song, Y. Xu, Z. Lang, Advanced pattern recognition technologies with applications to biometrics, Medical Information Science Reference, 2009.
[18] Q. Tao, R. Veldhuis, Threshold-optimized decision-level fusion and its application to biometrics, Pattern Recognition 42 (5) (2009) 823–836.
[19] K.-A. Toh, J. Kim, S. Lee, Maximizing area under ROC curve for biometric scores fusion, Pattern Recognition 41 (11) (2008) 3373–3392.
[20] K. Wu, K.-H. Yap, Fuzzy SVM for content-based image retrieval—a pseudo-label support vector machine framework, IEEE Computational Intelligence Magazine 1 (1) (2006) 10–16.
[21] K.-H. Yap, K. Wu, A soft relevance framework in content-based image retrieval systems, IEEE Transactions on Circuits Systems for Video Technology 15 (2005) 1557–1568.
[22] H. Fujisawa, Forty years of research in character and document recognition—an industrial perspective, Pattern Recognition 41 (8) (2008) 2435–2446.
[23] C.-F. Lin, S.-D. Wang, Training algorithms for fuzzy support vector machines with noisy data, Pattern Recognition Letters 14 (2004) 1647–1656.
[24] M. Kirby, L. Sirovich, Application of the KL procedure for the characterization of human faces, IEEE Transactions on Pattern Analysis and Machine Intelligence 12 (1) (1990) 103–108.
[25] M. Turk, A. Pentland, Eigenfaces for recognition, Journal of Cognitive Neuroscience 3 (1) (1991) 71–86.
[26] D. Zhang, X. Jing, J. Yang, Biometric Image Discrimination Technologies, Idea Group Publishing, Hershey, USA, 2006.
[27] J. Yang, D. Zhang, A.F. Frangi, J.Y. Yang, Two dimensional PCA: a new approach to appearance-based face representation and recognition, IEEE Transactions on Pattern Analysis and Machine Intelligence 24 (1) (2004) 131–137.
[28] Z. Jin, J.Y. Yang, Z.M. Tang, Z.S. Hu, A theorem on the uncorrelated optimal discriminant vectors, Pattern Recognition 34 (10) (2001) 2041–2047.
[29] J. Yang, J.Y. Yang, From image vector to matrix: a straightforward image projection technique—IMPCA vs. PCA, Pattern Recognition 35 (9) (2002) 1997–1999.
[30] D. Zhang, Z.H. Zhou, (2D)2PCA: two-directional two-dimensional PCA for efficient face representation and recognition, Neurocomputing 69 (1–3) (2005) 224–231.
[31] A. Xu, X. Jin, Y. Jiang, P. Guo, Complete two-dimensional PCA for face recognition, in: 18th International Conference on Pattern Recognition, 2006, pp. 481–484.
[32] J. Yang, J.Y. Yang, D. Zhang, J.F. Lu, Feature fusion: parallel strategy vs. serial strategy, Pattern Recognition 36 (6) (2003) 1369–1381.
[33] H. Yu, J. Yang, A direct LDA algorithm for high-dimensional data-with application to face recognition, Pattern Recognition 34 (10) (2001) 2067–2070.
[34] S. Mika, G. Rätsch, J. Weston, B. Schölkopf, K.-R. Müller, Fisher discriminant analysis with kernels, in: Proceedings of the IEEE International Workshop on Neural Networks for Signal Processing IX, 1999, pp. 41–48.

[35] S. Mika, G. Rätsch, B. Schölkopf, A. Smola, J. Weston, K.-R. Müller, Invariant feature extraction and classification in kernel spaces, Advances in Neural Information Processing Systems 12, MIT Press, Cambridge, Mass, 1999.
[36] G. Baudat, F. Anouar, Generalized discriminant analysis using a kernel approach, Neural Computation 12 (10) (2000) 2385–2404.
[37] Ch.-X. Ren, D.-Q. Dai, Incremental learning of bidirectional principal components for facerecognition, Pattern Recognition 43 (2010) 5–13.
[38] M.S. Park, J.Y. Choi, Theoretical analysis on feature extraction capability of class-augmented PCA, Pattern Recognition 42 (2009) 2353–2362.
[39] J. Ye, Generalized low rank approximations of matrices, in: Proceedings of the Twenty-first International Conference on Machine Learning, 2004, pp. 887–894.
[40] J. Yang, D. Zhang, Y. Xu, J.-Y. Yang, Two-dimensional discriminant transform for face recognition, Pattern Recognition 38 (7) (2005) 1125–1129.
[41] S. Chen, Y. Zhu, D. Zhang, J.-Y. Yang, Feature extraction approaches based on matrix pattern: MatPCA and MatFLDA, Pattern Recognition Letters 26 (8) (2005) 1157–1167.
[42] W. Zuo, D. Zhang, J. Yang, K. Wang, BDPCA plus LDA: a novel fast feature extraction technique for face recognition, IEEE Transactions on Systems, Man, and Cybernetics—Part B: Cybernetics 36 (4) (2006) 946–953.
[43] S.C. Dass, K. Nandakumar, A.K. Jain, A principled approach to score level fusion in multimodal biometric systems, AVBPA 2005, pp. 1049–1058.
[44] X.-Y. Jing, Y.-F. Yao, D. Zhang, et al., Face and palmprint pixel level fusion and Kernel DCV-RBF classifier for small sample biometric recognition, Pattern Recognition 40 (2007) 3209–3224.
[45] D. Zhang, W.K. Kong, J. You, M. Wong, On-line palmprint identification, IEEE Transactions on Pattern Analysis and Machine Intelligence 25 (9) (2003) 1041–1050.
[46] M.Á. Carreira-Perpiñán, Compression neural networks for feature extraction: application to human recognition from ear images (in Spanish), M.Sc. Thesis, Faculty of Informatics, Technical University of Madrid, Spain, 1995.
[47] 〈http://www.dcs.shef.ac.uk/∼miguel/papers/msc-thesis.html〉.
[48] 〈http://www.cl.cam.ac.uk/research/dtg/attarchive/facedatabase.html〉.

About the Author—YONG XU was born in Sichuan, China, in 1972. He received his B.S. degree, M.S. degree in 1994 and 1997, respectively. He received the Ph.D. degree in Pattern recognition and Intelligence System at NUST (China) in 2005. Now he works at Shenzhen graduate school, Harbin Institute of Technology. His current interests include feature extraction, biometric, face recognition, machine learning and image processing.

About the Author—DAVID ZHANG graduated in computer science from Peking University. He received his M.Sc. in computer science in 1982 and Ph.D. in 1985 from the Harbin Institute of Technology (HIT). From 1986 to 1988, he was a postdoctoral fellow at Tsinghua University and then an associate professor at the Academia Sinica, Beijing. In 1994, he received his second Ph.D. in electrical and computer engineering from the University of Waterloo, Ontario, Canada. Currently, he is a head, Department of Computing, and a chair professor at the Hong Kong Polytechnic University, where he is the founding director of the Biometrics Technology Centre (UGC/CRC) supported by the Hong Kong SAR government in 1998. He also serves as visiting chair professor in Tsinghua University, and Adjunct Professor in PeKing University, Shanghai Jiao Tong University, HIT and the University of Waterloo. He is the founder and editor-in-chief of the *International Journal of Image and Graphics* (IJIG); book editor of the Springer International Series on Biometrics (SISB); organizer of the International Conference on Biometrics Authentication (ICBA); Associate Editor of more than 10 international journals, including *IEEE Transactions and Pattern Recognition*; and the author of more than 10 books and 200 journal papers. Professor Zhang is a Croucher senior research fellow, distinguished speaker of the IEEE Computer Society, and fellow of both IEEE and IAPR.

About the Author—JING-YU YANG received the B.S. degree in Computer Science from NUST, Nanjing, China. From 1982 to 1984 he was a visiting scientist at the Coordinated Science Laboratory, University of Illinois at Urbana-Champaign. From 1993 to 1994 he was a visiting professor at the Department of Computer Science, Missuria University. And in 1998, he acted as a visiting professor at Concordia University in Canada. His current research interests are in the areas of pattern recognition, image processing and artificial intelligence, and expert system.