



Orthogonal discriminant vector for face recognition across pose

Jinghua Wang^a, Jane You^{a,*}, Qin Li^a, Yong Xu^b

^a Department of Computing, The Hong Kong Polytechnic University, Kowloon, Hong Kong

^b Harbin Institute of Technology, Shenzhen Graduate School, Shenzhen, Guangdong 518055, People's Republic of China

ARTICLE INFO

Article history:

Received 26 August 2010

Received in revised form

11 April 2012

Accepted 13 April 2012

Available online 24 April 2012

Keywords:

Pattern classification

Face recognition across pose

Face manifold

Orthogonal discriminant vector

ABSTRACT

Recognizing face images across pose is one of the challenging tasks for reliable face recognition. This paper presents a new method to tackle this challenge based on orthogonal discriminant vector (ODV). The result of our theoretical analysis shows that an individual's probe image captured with a new pose can be represented by a linear combination of his/her gallery images. Based on this observation, in contrast to the conventional methods which model face images of different individuals on a single manifold, we propose to model face images of different individuals on different linear manifolds. The contribution of our approach includes: (1) to prove that the orthogonality to ODVs is a pose-invariant feature.; (2) to categorize each person with a set of ODVs, where his/her face images possess zero projections while other persons' images are characterized by maximum projections; (3) to define a metric to measure the distance between a face image and an ODV, and classify the face images based on this metric. Our experimental results validate the feasibility of modeling the face images of different individuals on different linear manifolds. The proposed method achieves higher accuracy on face recognition and verification than the existing techniques.

© 2012 Published by Elsevier Ltd.

1. Introduction

Face recognition is an important topic in pattern recognition which has been extensively studied in the past decade [1–4]. One of the most popular methods for face recognition is the appearance-based method [5–10]. Appearance-based method represents a w -by- h face image by a vector in the wh -dimensional image space. It is concluded that the dimensionality of the face space is too high to allow robust face recognition [11], due to the variations of pose [12], occlusion [4], and illumination [13].

A number of algorithms are proposed [37–42] to recognize face images from different poses. Multidimensional scaling (MDS)-based approach performs well to recognize low resolution probe face images using high resolution gallery images [37]. Random Regression Forests [38] can estimate the head pose of different face images. To overcome large pose variation challenges, the ensemble based approach [39] boosts linear TFA models and achieve high accuracy. The method [42] stabilizes regressor against the pose difference and uses it to recognize face images under different poses. Both active appearance models-based landmarkings [41] and domain frequency based holistic features [40] are also proven to be helpful to solve the pose problem.

Though face images are high dimensional vectors, they are proven to reside on a low dimensional submanifold [7,14–17]. To understand the submanifold, many manifold learning techniques have been developed, which include marginal Fisher analysis (MFA) [18], neighborhood preserving embedding (NPE) [19], local discriminant embedding (LDE) [20], etc. Given face images in a high-dimensional space, these methods can extract the geometric properties of the images, such as intrinsic dimensionality, connected components, Euclidean embedding, etc [21].

Manifold learning methods [7,14–21] model face images of different individuals on a single manifold, which can efficiently identify the holistic structure of the original face images. However, these methods also eliminate very important discriminative information when learning the submanifold. Considering face images as points in a high dimensional space, we can regard the face images of each individual to span a linear manifold. The difference among these manifolds is a kind of discriminative information that is eliminated when modeling face images on a single manifold.

This paper presents a new appearance-based method for face recognition across pose by modeling face images of different individuals on different linear manifolds. Based on the idea that a linear manifold can be characterized by its norm vector, an orthogonal discriminant vector (ODV) for each manifold is defined and used to discriminate face images associated with this manifold from other images. (This new definition of ODV is different from what is described in [22–24], where orthogonal discriminant vectors are a

* Corresponding author. Tel.: +852 2766 7293; fax: +852 2774 0842.

E-mail addresses: csjihuwang@comp.polyu.edu.hk (J. Wang), csyjia@comp.polyu.edu.hk (J. You), csqinli@comp.polyu.edu.hk (Y. Xu).

set of mutually orthogonal vectors that maximize the Fisher criterion.) An ODV associated with one manifold is orthogonal to the face images on this manifold and not orthogonal to the rest face images. Association with different set of ODVs indicates the difference among linear manifolds.

The following lists the major contributions of this paper:

1. The introduction of a scheme to evaluate the intrapersonal relationship among face images with different poses by comparing an individual face image under a new pose with his/her gallery face images. The result of theoretical analysis shows that the face image under a new pose lies on the linear manifold spanned by his/her gallery face images despite of the change of poses.
2. The proposal of an identity-dependent and pose-invariant feature for face recognition across pose based on the intrapersonal relationship among face images. The new feature is the orthogonality to ODVs.
3. The development of a comprehensive procedure to examine the existence of ODV in face recognition and implementation of an effective two-step algorithm to calculate ODVs.

The remaining of this paper is organized as follows. Section 2 highlights the fundamental issues of the theoretical analysis to support the proposed algorithm while Section 3 describes the new ODV-based face recognition method. The experiments and performance evaluation are reported in Section 4. Finally, the conclusion and further discussion are presented in Section 5.

2. Intrapersonal relationship among face images across pose

We assume that, for each individual, only a few gallery 2-D face images under different poses are known and the probe face images are captured under novel poses. By predicting the face image under a novel pose using gallery face images, this section investigates the intrapersonal relationship among face images across pose. The prediction task takes the latent 3-D face object where the intrapersonal relationship among face images originates from as a medium. If the face images are well aligned, this section draws the conclusion that one's probe face image under a novel pose can be linearly expressed using his/her gallery face images.

Sections 2.1 and 2.2 present two procedures in reverse directions: 2-D image generation under a certain pose from the 3-D face and 3-D face reconstruction using the gallery face images. These two procedures are, respectively indicated by (a) and (b) in Fig. 1. With the help of these two procedures, Section 2.3 predicts a 2-D image under a novel pose, indicated by (c) in Fig. 1. Section 2.4 analyzes the prediction result and reveals the intrapersonal relationship among face images.

2.1. From 3-D face to 2-D image

Regarding the 3-D surface of a specific face as Lambertian [25], the intensity of the surface point (x,y,z) under a given lighting source \vec{s} can be computed as follows

$$F(x,y,z) = \rho(x,y,z) \cos \alpha \quad (1)$$

where α is the angle formed by lighting direction $\vec{s}(x,y,z)$ and normal $\vec{n}(x,y,z)$, $\rho(x,y,z)$ is the albedo of given point. After ordering the intensities of surface points lexicographically, we represent a 3-D face by a vector $F \in R^{N_3 \times 1}$, where the scalar N_3 denotes the number of pixels. For a fixed pose, we can derive the N_2 dimensional 2-D face image $f_0 \in R^{N_2 \times 1}$ from the 3-D face F by selecting the visible points (Fig. 1(a)). It is proved to be a linear orthogonal projection procedure from F to f_0 [25,26], and can be

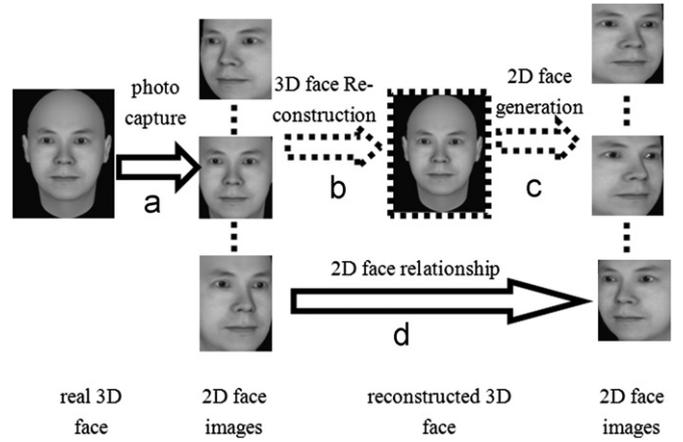


Fig. 1. Procedure for generating the intrapersonal relationship among 2-D face images. (a) 2-D face images generation; (b) 3-D face reconstruction; (c) 2-D face images generation from reconstructed 3-D face; (d) relationship generation among 2-D face images.

expressed as follows

$$f_0 = V_0 F \quad (2)$$

where $V_0 \in R^{N_2 \times N_3}$ is the pose-dependent projection operator. To achieve the goal of dropping all invisible points from F and only keeping all the visible ones in the face image f_0 , the elements of V_0 are set as follows: if the j th pixel in F is invisible, $(V_0)_{ij} = 0$ for $1 \leq i \leq N_2$; or else, if it is visible and projected as the i th pixel of f_0 , $(V_0)_{ij} = 1$ and $(V_0)_{kj} = 0$ ($1 \leq k \leq N_2, i \neq k$).

2.2. From 2-D image to 3-D face

Assume that n face images f_1, f_2, \dots, f_n under different poses V_1, V_2, \dots, V_n are generated from the same 3-D face F

$$f_i = V_i F (i = 1, 2, \dots, n) \quad (3)$$

If there is one point in F which is invisible in any of these images, the reconstruction of F from f_1, f_2, \dots, f_n will be theoretically ill-posed. However, based on the observation that

$$f = VF \quad (4)$$

where the matrix f consists of all the face images $f = [f_1^T \ f_2^T \ \dots \ f_n^T]^T \in R^{n \times N_2 \times 1}$ and the matrix V consists of all the pose-dependent projections $V = [V_1^T \ V_2^T \ \dots \ V_n^T]^T \in R^{n \times N_2 \times N_3}$, the 3-D face can be estimated as

$$\bar{F} = \bar{V}^+ (V^T f) = \bar{V}^+ (V^T VF) = \bar{V}^+ \left\{ \left(\sum_{i=1}^n V_i^T V_i \right) F \right\} = \bar{V}^+ \bar{V} F \quad (5)$$

where both $\bar{V} = \sum_{i=1}^n V_i^T V_i \in R^{N_3 \times N_3}$ and $\bar{V}^+ \in R^{N_3 \times N_3}$ are diagonal matrices. The diagonal elements of the matrix \bar{V}^+ are set as follows: $\bar{V}_{jj}^+ = 1/\bar{V}_{jj}$, if the j th ($1 \leq j \leq N_3$) diagonal element of \bar{V} is nonzero; and $\bar{V}_{jj}^+ = 1$, if $\bar{V}_{jj} = 0$.

As mentioned in Section 2.1, if and only if the k th pixel of f_i is the projection of the j th pixel of F , $(V_i)_{kj}$ equals one and, in turn, the j th diagonal element of $\bar{V}_i = V_i^T V_i$ equals one, i.e., $(\bar{V}_i)_{jj} = 1$. Thus, the j th diagonal element of $\bar{V} = \sum_{i=1}^n V_i^T V_i$ counts the number of face images in which the j th pixel of the 3-D face F is visible. If every pixel is visible in only one of the gallery face images, the task of 3-D face reconstruction has the unique solution $F = V^T f$. However, this case rarely occurs. Though recovering the pixels that are invisible in any of the face images is difficult, the unfavorable effects that some pixels are visible in

multiple face images can be eliminated by the multiplication of the diagonal matrix \bar{V}^+ . If a pixel is visible in m different images, its value in $V^T f$ is m times larger than that in the real face F . After the left multiplication of matrix \bar{V}^+ , the reconstructed face $\bar{F} = \bar{V}^+ V^T f$ is different from F only in the pixels that are visible in none of gallery face images.

2.3. 2-D image prediction

With the reconstructed 3-D face using Eq. (5), we can predict face images under any pose. Replacing the unknown real 3-D face with the reconstructed 3-D face \bar{F} , Eq. (2) for calculating the face image under a novel pose V_0 can be rewritten as

$$f_0 = V_0 \bar{V}^+ V^T f = W f$$

$$= [W_1 \quad W_2 \quad \dots \quad W_n] \begin{bmatrix} f_1 \\ f_2 \\ \vdots \\ f_n \end{bmatrix} = \sum_{i=1}^n W_i f_i \quad (6)$$

where $W = V_0 \bar{V}^+ V^T = [W_1 \quad W_2 \quad \dots \quad W_n]$. Eq. (6) states that one's face image f_0 under a novel pose can be expressed by his/her gallery face images. The underlying assumption of this prediction procedure is that the pixels of the probe face image are visible in at least one gallery face image. Note that, *the probe face image needs not to have a similar pose to one of the gallery face images.*

2.4. Discussion

As can be seen from Fig. 2, the well aligned face images of the same individual but under different poses are similar in terms of the main organs' spatial configuration and shape. Thus, if a novel face image f_0 can be expressed by a set of face images f_1, f_2, \dots, f_n , this globally linear expression should also hold in locally areas, such as different organs in the face. In other words, if the equation $f_0 = \sum_{i=1}^n \lambda_i f_i$ holds, we can linearly express the "eye" e_0 of the face f_0 by the "eyes" e_1, e_2, \dots, e_n of the gallery faces f_1, f_2, \dots, f_n , i.e., $e_0 = \sum_{i=1}^n \lambda_i e_i$. Intuitively, it does not make sense to express the "eye" of a face image by the combination of all the organs, such as eye, mouth, nose et al.

Under the assumption that the matrix W_i is a scalar matrix in Eq. (6), i.e., $W_i = \lambda_i I$ (I is the identity matrix), the unreasonable situation that kinds of organs are combined together to form a specific organ is avoided and all parts of the face images are linearly expressed only by their counterparts in other face images, as shown in Fig. 3. Then, we have the following theorem

Theorem 1. If the gallery and probe face images are well aligned, one's probe face image captured under a novel pose can be approximately expressed as a linear combination of his/her gallery face images, i.e.,

$$f_0 = \sum_{i=1}^n \lambda_i f_i \quad (7)$$



Fig. 2. Face images of the same individual have the similar configuration.

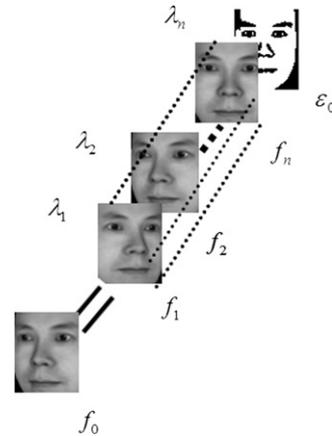


Fig. 3. Linear expression of a novel face image using gallery face images.

Theorem 1 shows that the probe face image under a novel pose lies on the linear manifold spanned by gallery face images. Thus, we can associate an individual with a linear manifold and model all of his/her face images on it.

In practice, both the gallery and probe face images may be polluted by noise and a novel face image may be only linearly expressible by the gallery face images with a nonzero residue ϵ_0 , as follows

$$f_0 = \sum_{i=1}^n \lambda_i f_i + \epsilon_0 \quad (8)$$

There are many ways to measure the residue in Eq. (8). For most of the measurements, it is time consuming to seek the optimal coefficient that minimizes the residue. This paper, however, just takes advantage of this intrapersonal relationship among face images and avoids seeking the coefficients for the linear expression in Eq. (8).

3. Orthogonal discriminant vectors

This section proposes a method for face recognition across pose based on orthogonal discriminant vector (ODV). Section 3.1 presents the basic idea and defines the ODV. Section 3.2 investigates the existence of ODV in face recognition. Section 3.3 develops an algorithm to calculate ODVs. Section 3.4 investigates the classification of face images using ODV. Section 3.5 analyzes the computational complexity of the proposed method.

3.1. Basic ideas

Geometrically, Theorem 1 indicates that one's novel face image lies on the linear manifold spanned by his/her gallery face images. Because of this, a vector v is orthogonal to the probe face image, if v is orthogonal to all the gallery face images. This is derived from the geometrical theorem that if a vector v is orthogonal to a set of vectors $x_i (i=1, 2, \dots, n)$, i.e., $v \perp x_i (i=1, 2, \dots, n)$, it is orthogonal to any linear combination of them, i.e., $v \perp \sum_{i=1}^n \alpha_i x_i$, where $\alpha_i (i=1, 2, \dots, n)$ are coefficients. Thus, we have the following Corollary 1 from Theorem 1:

Corollary 1. Being orthogonal to a certain vector is a pose-invariant feature for the face images of an individual.

Assume the face images of two individuals span linear manifolds S_1 and S_2 . The methods [5–10] overlook the difference among these two linear manifolds when modeling all the face images on the union manifold $S = S_1 \cup S_2$. Typically, the difference between face linear manifolds S_1 and S_2 is not null, i.e., $S_1 - S_2 \neq \emptyset$ and $S_2 - S_1 \neq \emptyset$. We can decompose S_1 into two orthogonal linear

manifolds $S_1=S_{11}+S_{12}$ and $S_{11}\cap S_{12}=\varphi$, where $S_{11}\perp S_{12}$ and $S_{12}\subset S_2$. While the face images of the second individual have nonzero projections onto the vectors in S_{11} , the face images of the first individual have zero projections. So, being orthogonal to any vector in S_{11} is a pose-invariant feature for the first individual.

We define the orthogonal discriminate vector (ODV) associating with the i th individual as follows:

Definition 1. *If a vector v is orthogonal to all the gallery face images of the i th individual and not orthogonal to any gallery images of the other individuals, this vector v is an ODV associating with the i th individual.*

As revealed by Theorem 1, one's probe face image captured under a novel pose can be linearly expressed by his/her gallery face images. So, the ODVs that are orthogonal to the gallery face images are also orthogonal to the face images under novel poses. Thus, being orthogonal to ODVs is a pose-invariant feature, which can be used in face recognition across pose.

3.2. The existence of the ODV

In this subsection, we study the existence of ODV in the task of face recognition. Let the N dimensional vectors $x_i \in \mathbb{R}^{N \times 1}$ ($1 \leq i \leq n_1$) be n_1 face images of one individual, and $y_j \in \mathbb{R}^{N \times 1}$ ($1 \leq j \leq n_2$) be n_2 face images of the others. We consider the face recognition as a two-class classification problem. The face image matrix of the first class is denoted as $X = [x_1 \ x_2 \ \dots \ x_{n_1}] \in \mathbb{R}^{N \times n_1}$ and the one of the second class is denoted as $Y = [y_1 \ y_2 \ \dots \ y_{n_2}] \in \mathbb{R}^{N \times n_2}$.

Let $Z = \{z_i \in \mathbb{R}^{N \times 1}, 1 \leq i \leq N\}$ be a set of mutually orthogonal unit vectors that span the N dimensional vector space. Then, the face images from these two classes can be linearly expressed by the bases, as follows

$$\begin{cases} x_i = Zd_i^1, & 1 \leq i \leq n_1 \\ y_j = Zd_j^2, & 1 \leq j \leq n_2 \end{cases} \quad (9)$$

and

$$\begin{cases} X = ZD_1 \\ Y = ZD_2 \end{cases} \quad (10)$$

where $Z = [z_1 \ z_2 \ \dots \ z_N]$ is the base matrix, $d_i^1, d_j^2 \in \mathbb{R}^N$ are coefficient vectors associating with x_i and y_j . The matrices $D_1 = [d_1^1 \ d_2^1 \ \dots \ d_{n_1}^1] \in \mathbb{R}^{N \times n_1}$ and $D_2 = [d_1^2 \ d_2^2 \ \dots \ d_{n_2}^2] \in \mathbb{R}^{N \times n_2}$ are two coefficient matrices.

The theorem of generalized singular value decomposition (GSVD) states as follows: for the given two matrices $D_1 \in \mathbb{R}^{N \times n_1}$,

$D_2 \in \mathbb{R}^{N \times n_2}$, $D = \begin{pmatrix} D_1^T \\ D_2^T \end{pmatrix}$ and $t = \text{rank}(D)$, there exist orthogonal matrices $L_1 \in \mathbb{R}^{n_1 \times n_1}$, $L_2 \in \mathbb{R}^{n_2 \times n_2}$, $W \in \mathbb{R}^{t \times t}$, and $Q \in \mathbb{R}^{N \times N}$ such that

$$L_1^T D_1^T Q = \Sigma_1 \begin{pmatrix} \underbrace{W^T K}_t & \underbrace{0}_{N-t} \end{pmatrix} \quad (11)$$

and

$$L_2^T D_2^T Q = \Sigma_2 \begin{pmatrix} \underbrace{W^T K}_t & \underbrace{0}_{N-t} \end{pmatrix} \quad (12)$$

where

$$\Sigma_1 = \begin{pmatrix} I_1 & & \\ & J_1 & \\ & & O_1 \end{pmatrix} \quad \text{and} \quad \Sigma_2 = \begin{pmatrix} O_2 & & \\ & J_2 & \\ & & I_2 \end{pmatrix} \quad (13)$$

and $K \in \mathbb{R}^{t \times t}$ is nonsingular with its singular values equal to the nonzero singular values of D . The matrices $I_1 \in \mathbb{R}^{t \times r}$ and $I_2 \in \mathbb{R}^{(t-r-s) \times (t-r-s)}$ are identity matrices, where

$$r = \text{rank} \begin{pmatrix} D_1^T \\ D_2^T \end{pmatrix} - \text{rank}(D_2^T) \quad \text{and} \quad s = \text{rank}(D_1^T) + \text{rank}(D_2^T) - \text{rank} \begin{pmatrix} D_1^T \\ D_2^T \end{pmatrix} \quad (14)$$

The matrices $O_1 \in \mathbb{R}^{(n_1-r-s) \times (t-r-s)}$ and $O_2 \in \mathbb{R}^{(n_2-t-s) \times r}$ are zero matrices with possible no rows or no columns. The matrices $J_1 = \text{diag}(a_{r+1}, \dots, a_{r+s})$ and $J_2 = \text{diag}(b_{r+1}, \dots, b_{r+s})$ satisfy

$$1 > a_{r+1} \geq \dots \geq a_{r+s} > 0 \quad \text{and} \quad 0 < b_{r+1} \leq \dots \leq b_{r+s} < 1$$

$$a_i^2 + b_i^2 = 1 \quad \text{for} \quad i = r+1, \dots, r+s \quad (15)$$

Based on (11) and (12), we have

$$\begin{cases} D_1 = Q \begin{bmatrix} K^T W \\ 0 \end{bmatrix} \Sigma_1 L_1^T \\ D_2 = Q \begin{bmatrix} K^T W \\ 0 \end{bmatrix} \Sigma_2 L_2^T \end{cases} \quad (16)$$

Using (16), we can rewrite (10) as follows

$$X = ZD_1 = ZQ \begin{bmatrix} K^T W \\ 0 \end{bmatrix} \Sigma_1 L_1^T = ZQ \begin{bmatrix} K^T W \\ 0 \end{bmatrix} \begin{bmatrix} I_1 & & \\ & J_1 & \\ & & 0 \end{bmatrix} L_1^T \quad (17)$$

and

$$Y = ZD_2 = ZQ \begin{bmatrix} K^T W \\ 0 \end{bmatrix} \begin{bmatrix} 0 & & \\ & J_2 & \\ & & I_2 \end{bmatrix} L_2^T \quad (18)$$

Using the above two equations, we can prove Theorem 2

Theorem 2. *The column vectors of V_\perp are ODVs of the second class, where*

$$V_\perp = ZQ \begin{bmatrix} K^T W \\ 0 \end{bmatrix} \begin{bmatrix} I_1 \\ 0 \\ 0 \end{bmatrix} \quad (19)$$

Proof. We only need to prove the columns of V_\perp are orthogonal to the images of the second class and are not orthogonal to the images of the first class.

$$V_\perp^T Y = [I_1 \ 0 \ 0] [W^T K \ 0] Q^T Z^T ZQ \begin{bmatrix} K^T W \\ 0 \end{bmatrix} \begin{bmatrix} 0 & & \\ & J_2 & \\ & & I_2 \end{bmatrix} L_2^T \quad (20)$$

Because both Z and Q are orthogonal matrices, we have

$$V_\perp^T Y = [I_1 \ 0 \ 0] W^T K K^T W \begin{bmatrix} 0 & & \\ & J_2 & \\ & & I_2 \end{bmatrix} L_2^T = 0 \quad (21)$$

Thus, the column vectors of V_\perp are orthogonal to all the images of the second class.

$$\begin{aligned} V_\perp^T X &= [I_1 \ 0 \ 0] [W^T K \ 0] Q^T Z^T ZQ \begin{bmatrix} K^T W \\ 0 \end{bmatrix} \begin{bmatrix} I_1 & & \\ & J_1 & \\ & & 0 \end{bmatrix} L_1^T \\ &= [I_1 \ 0 \ 0] W^T K K^T W \begin{bmatrix} I_1 & & \\ & J_1 & \\ & & 0 \end{bmatrix} L_1^T = W^T K K^T W L_1^T \end{aligned} \quad (22)$$

Because L_1 and W are orthogonal matrices, and K is a nonsingular matrix, we have

$$V_{\perp}^T X \neq 0 \tag{23}$$

Thus, the column vectors of V_{\perp} are not orthogonal to all the images of the first class.

In summary, the column vectors of V_{\perp} are orthogonal to all the images in the second class and not orthogonal to all the images in the first class. Thus, the columns of V_{\perp} are ODVs of the second class.

In a similar way, we can prove that the columns of the matrix V_{\perp}' in Eq. (24) are ODVs of the first class

$$V_{\perp}' = ZQ \begin{bmatrix} K^T W \\ 0 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ I_2 \end{bmatrix} \tag{24}$$

As $I_1 \in R^{r \times r}$ and $r = \text{rank}(D) - \text{rank}(D_2^T)$, we know that the ODV of the second class does not exist when $\text{rank}(D) = \text{rank}(D_2^T)$. Similarly, the ODV of the first class does not exist when $\text{rank}(D) = \text{rank}(D_1^T)$. Thus, when $\text{rank}(D) = \text{rank}(D_1^T) = \text{rank}(D_2^T)$, there is no ODV in this two-class classification problem. However, the equation $\text{rank}(D) = \text{rank}(D_1^T) = \text{rank}(D_2^T)$ holds only when all the samples from one class can be linearly expressed by the combinations of the samples from the other class. For high dimensional face images, such a situation hardly happens. This is because one individual must have a character that differentiates his/her face images from those of others. Our experimental results verify the existence of the ODV.

3.3. The calculation of the ODV

For simplicity in description, we first consider a two-class classification problem. We regard the face images $X = [x_1 \ x_2 \ \dots \ x_{n_1}] \in R^{N \times n_1}$ of one individual as the first class and the face images $Y = [y_1 \ y_2 \ \dots \ y_{n_2}] \in R^{N \times n_2}$ of the others as the second class. Also, we assume the first class has fewer images than the second class, i.e., $n_1 < n_2$.

We propose the following model to calculate the ODVs for the first class

$$\max \|Y^T v\|_2 \text{ s.t. } X^T v = 0 \tag{25}$$

By minimizing the projections of the first class to be zero and maximizing the projections of the second class, this model aims to generate the most discriminative ODVs. We propose a two-step procedure to solve the model in (25). Step 1 generates a set of candidate vectors for the ODV. From candidate set, step 2 chooses the most discriminative ODVs onto which the images of the second class have the maximum projections.

Step 1: generate candidate vectors for ODVs

To generate the candidate vectors, step 1 only needs to solve the following linear equation system

$$X^T Y \mu = 0 \tag{26}$$

In fact, we have a theorem as follows

Theorem 3. The nonzero vector $v = Y\mu \neq 0$ is an ODV of the first class, if μ is a nonzero solution vector of Eq. (26).

Proof. We only need to prove that the following two formulas regarding the vector v hold

$$\begin{cases} v^T X = (Y\mu)^T X = \mu^T Y^T X = (X^T Y \mu)^T = 0 \\ v^T Y = (Y\mu)^T Y = \mu^T Y^T Y \neq 0 \end{cases} \tag{27}$$

While the first formula in (27) is certain to be true and does not need further proof, the second formula can be proved by contradiction as follows.

Suppose $\mu^T Y^T Y = 0$, then

$$\mu^T Y^T Y = 0 \Rightarrow \mu^T Y^T Y \mu = 0 \Leftrightarrow (Y\mu)^T Y \mu = 0 \tag{28}$$

As we know $v^T v = (Y\mu)^T Y \mu = \sum_{i=1}^N v_i^2 \geq 0$, where N is the dimensionality of the face images. If and only if $v_i = 0$ for $1 \leq i \leq N$, i.e., $v = 0$, the equation $v^T v = 0$ holds. But it is impossible for v to be zero vector, which contradicts $v = Y\mu \neq 0$. This completes the proof.

Suppose $U = [\mu_1 \ \mu_2 \ \dots \ \mu_l] \in R^{n_2 \times l}$ is a set of linearly independent solutions of (26), we can denote the candidate ODV set as follows

$$S = \{v = YU\alpha \mid \alpha \in R^{l \times 1}\} \tag{29}$$

where α is a coefficient vector. In fact, solving (26) (obtaining the matrix U) implies finding the null space of the coefficient matrix $X^T Y$. This set includes all the candidate ODVs as $U\alpha$ is the general form for the solutions of (26). Among the infinite candidate vectors in S , step 2 picks out the most discriminative ones and takes them as ODVs.

Step 2: choose the most discriminative ODVs

Among all the vectors in S , step 2 chooses the ODVs corresponding to the longest second class projections for the further classification. We formulate it as follows

$$\max_{v \in S} \|Y^T v\|_2 \tag{30}$$

Substituting v by $YU\alpha$ in (30), we have

$$\max \|Y^T v\|_2^2 = \max \|v^T Y Y^T v\|_2 = \max \|\alpha^T U^T Y^T Y Y U \alpha\|_2 \tag{31}$$

It can be easily proved that the coefficient vectors α should be the eigenvectors of the matrix $M = U^T Y^T Y Y U \in R^{l \times l}$ associating with the leading eigenvalues. If the eigenvectors of the matrix M that associate with the k largest eigenvalues are $\alpha_i (i = 1, 2, \dots, k)$, the ODVs for the first class are

$$v_i = YU\alpha_i (i = 1, 2, \dots, k) \tag{32}$$

Eq. (32) shows that the ODVs for the first class are in fact linear combinations of the face images from the second class.

To sum up, the following two-step algorithm can generate the most discriminative ODVs:

Firstly, solve the equation $X^T Y \mu = 0$ and output a set of independent solutions $\mu_1, \mu_2, \dots, \mu_l$;

Secondly, perform eigendecomposition of matrix $M = U^T Y^T Y Y U$ to solve the maximization problem (31) and generate the leading eigenvectors α_i ; output the ODVs $v_i = YU\alpha_i (i = 1, 2, \dots, k)$;

3.4. ODV-based face classification

Based on the cosine metric, we define the distance between a face image x and an ODV v as follows

$$d(x, v) = 1 - |\cos(x, v)| = 1 - \frac{|x^T v|}{\|x\| \|v\|} \tag{33}$$

Such defined distance achieves its maximum value, i.e., one, if $x \perp v$; and achieves its minimum value, i.e., zero, if $x // v$.

3.4.1. Two-class classification problem

In a two-class classification problem, the ODV v for the first class is orthogonal to the first class image x_i and not orthogonal to the second class image y_j , i.e., $v^T x_i = 0$ and $v^T y_j \neq 0$. So, the distances between v and the first class face images equal one, and those between v and the second class face images are smaller than one. If we use the common center of the circles O to represent the ODV v and the metric defined in (33) to measure the distances between v

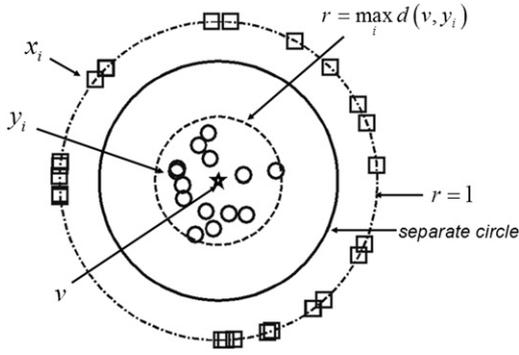


Fig. 4. The distance of the ODV v to the images from two classes.

and the face images, the first class images (represented by “□”) scatter on the unit circle centered at O for a 2-D case, whereas the second class images (represented by “○”) scatter inside of the unit circle, as shown in Fig. 4. Theoretically, a circle can correctly separate the images from these two classes, if its radius r satisfies $\max_{1 \leq i \leq n_2} d(v, y_i) < r < 1$.

3.4.2. Many-class classification problem

If the face images belong to more than two individuals, the face recognition task becomes a many-class classification problem. We take the one-to-many strategy and calculate a set of ODVs $v_j^i (1 \leq i \leq c; 1 \leq j \leq k_i)$ for every class, where $c (c > 2)$ is the number of individuals and k_i is the number of ODVs associating with the i th class. For a gallery face image x belonging to the $c(x)$ class, its distance to the ODV v_j^i satisfies

$$\begin{cases} d(x, v_j^i) = 1 & c(x) = i \\ d(x, v_j^i) < 1 & c(x) \neq i \end{cases} \quad (34)$$

Normally, the larger the distance between a face image and the ODV v_j^i , the more likely this face image belongs to the i th class. Here, we classify a probe face image x into the $c(x)$ th class, if

$$\frac{1}{k_{c(x)}} \sum_{j=1}^{k_{c(x)}} d(x, v_j^{c(x)}) = \max_{1 \leq i \leq c} \frac{1}{k_i} \sum_{j=1}^{k_i} d(x, v_j^i) \quad (35)$$

The complete face recognition procedure is summarized in the following algorithm.

Algorithm 1. ODV-based face recognition

Training stage: For $1 \leq i \leq c$, the following procedure is performed to calculate a set of ODVs for the i th class

Step 1. Take the face images in i th class as columns of X and those of other classes as columns of Y ;

Step 2. Solve the linear equation system $X^T Y \mu = 0$, and obtain a set of linearly independent solutions $U = [\mu_1 \ \mu_2 \ \dots \ \mu_l]$

$\in R^{n_2 \times l}$, where l is the number of linearly independent solutions; Step 3. Perform eigendecomposition of the matrix $M = U^T Y^T Y Y^T U$, and denote the eigenvectors associating with nonzero eigenvalues as $\alpha_j^i (j = 1, 2, \dots, k_i)$;

Step 4. Calculate the ODVs using $v_j^i = Y U \alpha_j^i (j = 1, 2, \dots, k_i)$;

Testing stage:

Step 1. Calculate the distances (defined in (33)) between probe face image x and the ODVs of each class;

Step 2. The probe face image x is classified into the $c(x)$ th class according to (35).

3.5. Computational complexity

Now, let us analyze the computational complexity of Algorithm 1 in Section 3.4.

In the training stage, step 1 only picks out the face images associating with the i th individual with computational complexity of $O(n)$, where n is the total number of images. As $X \in R^{N \times n_1}$ and $Y \in R^{N \times n_2}$, the construction of the linear equation system $X^T Y \mu = 0$ has computational complexity of $O(n_1 n_2 N)$ in step 2, where N is the dimensionality of the face images and n_1, n_2 are number of images in the two classes. Step 2 solves the linear equation system with computational complexity of $O(n^3)$.

Step 3 involves a series of matrix multiplication and an eigendecomposition procedure. To calculate the matrix M efficiently, we reformulate it as follows

$$M = (U^T (Y^T Y)) ((Y^T Y) U) = ((Y^T Y) U)^T ((Y^T Y) U) = M_1^T M_1 \quad (36)$$

Calculation of the matrix $M_1 = (Y^T Y) U \in R^{n_2 \times l}$ has the computational complexity of $O(n_2^2 N + n_2^2 l)$, and calculation of $M = M_1^T M_1$ has the computational complexity of $O(n_2 l^2)$, where $l (l < n)$ is the number of solutions generated in step 2. The eigendecomposition procedure of the matrix $M \in R^{l \times l}$ needs computational complexity of $O(l^3)$. Thus, the step 3 has computational complexity of

$$O(n_2^2 N + n_2^2 l) + O(n_2 l^2) + O(l^3) < O(n^3 + n^2 l + n^2 N) \quad (37)$$

Step 4 performs a series of matrix multiplication to generate ODV $v_j^i = Y U \alpha_j^i (j = 1, 2, \dots, k_i)$ with computational complexity of $O(N n l)$.

Thus, the calculation of ODVs of one class has computational complexity of

$$O(n_1 n_2 N) + O(n^3) + O(n^3 + n^2 l + n^2 N) + O(N n l) < O(N n^2) \quad (38)$$

These four steps are performed c times, where c is the number of classes. Totally, the computational complexity of the training stage is $O(c N n^2)$. This indicates that the computational complexity of the training stage is mainly determined by the number of face images other than the dimensionality of them which is usually larger.

In the testing stage, when classifying a probe face image, we only need to calculate its distances to $\sum_{i=1}^c k_i$ ODVs (where k_i is the number of ODVs associating with the i th class) with the computational complexity of $O(N(\sum_{i=1}^c k_i))$. From the calculation procedure of the ODV, we know that the number of ODVs associating with one class is strictly smaller than the number of gallery images, i.e., $k_i < n$. Thus, the computational complexity of classifying a probe face image satisfies

$$O(N \times \sum_{i=1}^c k_i) < O(N c n) \quad (39)$$

Note that the computational complexity of both the training and testing procedure is in direct proportion to the dimensionality of the face images.

4. Experiments

This section presents our experiments on the popular face databases. Section 4.1 validates the Theorem 1 by investigating the residues in the linear expressions of probe face images using gallery face images. Section 4.2 and 4.3, respectively show the experimental results of face verification and recognition.

One standard for evaluating the face recognition technologies is the Facial Recognition Technology (FERET) database [27], which was sponsored by US Department of Defense through the DARPA Program. A subset of the FERET database including 1 386 images of 198 individuals (7 images for each individual) is used in our experiments. An image is included in this subset, if its name is

marked with any of the following two character strings: “ba”, “bd”, “be”, “bf”, “bg”, “bj”, and “bk” [10]. The images in this subset have pose variations of $\pm 15^\circ$, $\pm 25^\circ$, and also the variations of illumination and the expression. The original images are cropped and resized to 80×80 pixels.

Another standard database for evaluating the face recognition across pose is the Carnegie Mellon University Pose, Illumination and Expression database (CMU PIE database) [28]. The CMU PIE database totally consists of more than 40 000 facial images of 68 people. In the construction of this database, the images of each individual are captured under 43 different illumination conditions, across 13 different poses, and with 4 different expressions. We use a subset contains five poses (C05, C07, C09, C27, C29) and all different illuminations and expressions. There are 170 images for each of the 50 individuals.

The Yale Face Database B [29] contains 5850 images, 585 images for each of 10 individuals. The images of one individual are captured under 9 different poses and illumination conditions. Also, an image with ambient illumination is captured under all of the 9 different poses for each individual.

The AR face database [30] contains 3120 images corresponding to the faces of 120 people. The images include frontal view faces with different facial expressions, conditions of illumination, and occlusions (sun glasses and scarf). Each person participates in two sessions, separated by intervals of two weeks. The same pictures are taken in both sessions.

4.1. Residue investigation

This subsection verifies Theorem 1 through investigating the residues in the linear expressions of probe face images using gallery images. Assume that a novel face image z can be linearly expressed by the gallery face images x_1, x_2, \dots, x_n of one individual with a residue ε as follows

$$z = \lambda_1 x_1 + \lambda_2 x_2 + \dots + \lambda_n x_n + \varepsilon \tag{40}$$

We calculate the minimum residue in terms of l_2 -norm using the following method. First, we calculate the covariance matrix of x_1, x_2, \dots, x_n and its eigenvectors g_1, g_2, \dots, g_n . Then, we can obtain the minimum residue as follows

$$\varepsilon = z - \sum_{i=1}^n x_i^T g_i \tag{41}$$

The reasoning behind doing so is that the two spaces, respectively spanned x_1, x_2, \dots, x_n and g_1, g_2, \dots, g_n are the same.

In this experiment, we divide the face images of each individual into two halves. The first half is the gallery set and the second half is the probe set. The residue in (41) is an intra-class residue if z and x_1, x_2, \dots, x_n associate with the same individual. The residue is an inter-class residue if they associate with different individuals.

Fig. 5 shows the probability distribution of the norm of intra- and inter-class residues. It can be seen that the intra-class residue is usually much smaller than the inter-class residue. One’s face image is more likely to be well expressed by his/her face images other than the images of others. The large inter-class residue indicates that one’s probe face image is far away from the linear manifold spanned by the face images of another person. So, the face images of different individuals span different linear manifolds, and the ODV that is orthogonal to only one of the linear manifolds exists.

4.2. Face verification

The goal of face verification is to determine whether a probe image belongs to a particular person. In the proposed ODV-based method, we work out a distance between the probe image and an

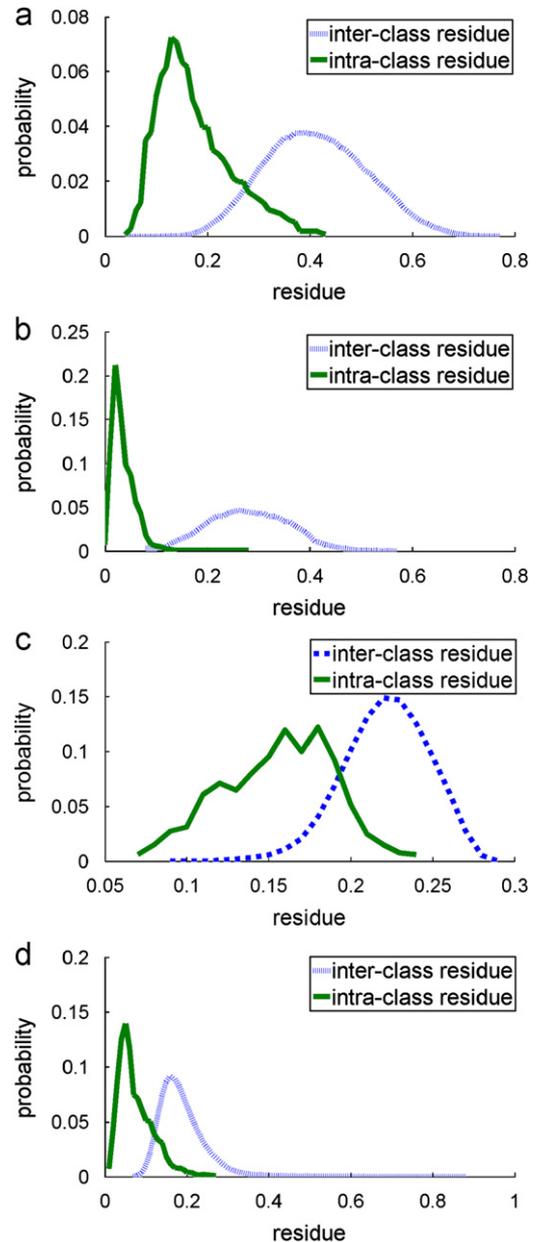


Fig. 5. Inter- and intra-class residue distribution on (a) AR face database; (b) YaleB face database; (c) FERET face database; (d) PIE database.

ODV, and compare it with a defined threshold. The probe image is classified into the class which the ODV associating with, only if the distance is larger than the threshold. We use the subset of CMU PIE containing 8500 face images of 50 different individuals and the whole YaleB databases. Both of them are randomly divided into two halves; one half is taken as gallery and the other half is taken as probe.

The distances between ODVs and face images can be classified into two groups:

1. Intra-class distance: the face image and ODV associate with the same class. The distance in this group is expected to be as large as possible.
2. Inter-class distance: the face image and ODV associate with other class. The distance in this group is expected to be as small as possible.

In Figs. 6 and 7, the original represents ODV and the dots represent the face images. If the distance between a face image

and the ODV is d_1 , this face image is represented by a random dot on the circle with radius equals to d_1 . Here, the distance is regularized using the following equation

$$d_{new} = \frac{d - \min(d)}{\max(d) - \min(d)} \quad (42)$$

In both Figs. 6 and 7, the circles in figures (a) and (b) have the same radius which is smaller than 95% intra-class distances, i.e., only 5% of the dots scatter inside of the circle in Fig. (a). As can be seen from Fig. (b), only a small portion of the inter-class distances are larger than the radius of the circle.

There are two kinds of misclassification in face verification: false acceptance (FA), where the system incorrectly classifies a face image into a class which it does not belong to; and false rejection (FR), where the system fails to classify a face image into the class which it belongs to. For each threshold, there will be one pair of false acceptance rate (FAR) and false rejection rate (FRR). By tuning the threshold, we can have a series of FARs and the corresponding FRRs. Then, we can plot FRR verse FAR and obtain a receiver operator curve (ROC) for each database. Fig. 8(a) and (b), respectively show the ROC for the subset of CMU PIE and the Yale B database. The Equal Error Rate (EER) is defined to be the FRR when it equals to FAR. The EER for the CMU PIE face database is 4.7%, which is lower than those of the bihashing algorithm (11.93%) [34], original Fisherface method (18.18%) [35] and its

improvements (larger than 5%) in [35]. The EER for the YaleB database is 3.5%, a little lower than the results reported in [36].

4.3. Face recognition

To test the proposed method, this subsection presents the face recognition experiments on the AR, FERET, CMU PIE, and YaleB face databases. We compare the performance of the proposed method with other six appearance-based methods: principal component analysis (PCA) [31], Fisher discriminant analysis (FDA) [5], (Maximum a posterior discriminant analysis) MLDA [40], Kernel principal component analysis (KPCA) [32,33], Kernel Fisher discriminant analysis (KFDA) [10], and local preserving projection (LPP) [7]. All of the above five methods and our method classify a face image using the nearest neighbor classifier. Both KFDA and KPCA adopt the Gaussian Kernel.

Seven-fold cross validation is used on the FERET face database. The subset of FERET are divided into seven portions based on the names of images, which are marked with “ba”, “bd”, “be”, “bf”, “bg”, “bj”, and “bk”. These seven portions are, respectively captured under different circumstances. Six portions are used for training and the rest portion is used for testing in our experiments. Thus, the sizes of the training and testing subset are, respectively 1 188 and 198. The proposed method generates 198 ODVs, and other methods generate 198 projection vectors. The classification results are listed in Table 1.

Table 1 shows that the proposed method outperforms the other five methods in most cases. When the “bg” portion is used for testing, the proposed method achieves the classification accuracy of 95.5%, more than 5% higher than the other methods. Though KFDA achieves a little higher accuracy (0.5%) on “bd” and “bf”, KFDA is very time consuming in both training and testing procedure. Moreover, it is very difficult to fix the parameter for kernel function in both KFDA and KPCA. All the methods (PCA, FDA, and LPP) make an implicit assumption that the face images of each individual cluster together. However, as it is widely recognized, the variations induced by pose change can be larger than those induced by identity change. Different from these three methods, our method models face images of different individuals on different linear manifolds and does not require the images of

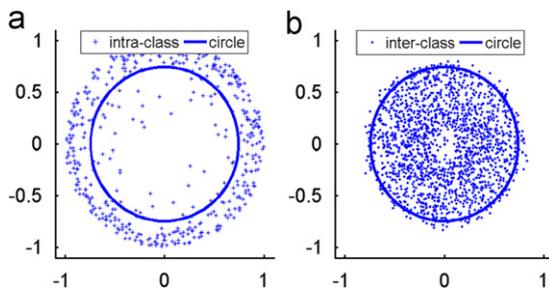


Fig. 6. The distribution of the face images in CMU PIE database.

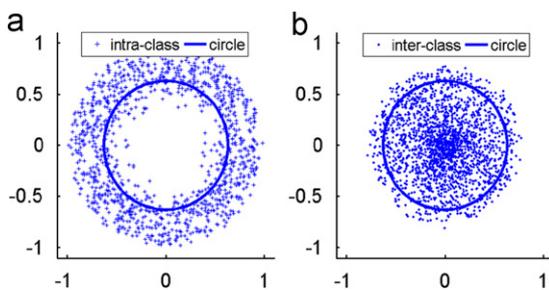


Fig. 7. The distribution of the face images in YaleB database.

Table 1

Classification accuracies (%) of methods on FERET face database.

Portion for test	Ba	bd	be	bf	bg	bj	bk
PCA	81.5	67.0	69.5	52.5	77.5	68.0	78.5
FDA	89.5	68.5	90.5	47.5	84.0	65.5	86.5
KPCA	86.0	76.0	75.5	58.0	89.5	68.5	82.5
KFDA	90.5	76.5	82.0	60.5	88.0	67.0	91.0
LPP	84.0	71.5	90.0	51.0	89.5	66.5	84.5
MLDA	87.5	72.5	89.0	52.0	81.5	66.0	78.0
Proposed method	92.0	76.0	92.5	60.0	95.5	69.0	94.0

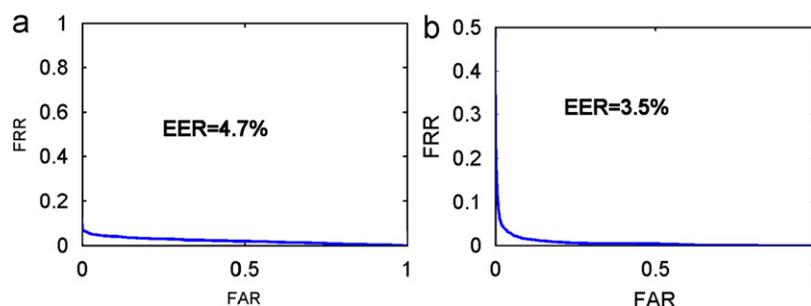


Fig. 8. The ROC curves. (a) The subset of CMU PIE database; (b) The YaleB database.

one individual cluster together. Because of this, our method achieves accuracies 2%–10% higher than PCA, LDA, and LPP.

The other three databases are randomly divided into two disjoint subsets, one for training (gallery images) and one for testing (probe images). For the YaleB database, the training subset contains 10 images for each individual (account for about 1.8%) and the testing subset contains 575 images for each individual (account for about 98.2%). The training subset and the testing subset in the CMU PIE face database, respectively have size of 3400 and 8160, and, respectively contain 50 and 120 images for each individual. The AR database is randomly divided into two halves, one for training and one for testing. These random divisions are repeated 20 times. The average classification accuracies are plotted versus the number of projection vectors (ODVs in the proposed method) in Fig. 9.

In these figures, the dimensionality means the number of ODVs in our method and number of projectors in other methods. As can be seen, the proposed method performs better than other methods. Note that, the proposed method can achieve much

Table 2

Training time (seconds) of different methods in three face databases.

Database	Size	PCA	FDA	LPP	KPCA	KFDA	Proposed method
AR	80 × 100	349	405	864	3447	4558	325
CMU PIE	32 × 32	703	891	1093	2420	1779	981
YaleB	160 × 120	54	69	206	1432	1030	48

higher classification accuracy than other methods when the dimensionality is low, especially on the CMU PIE and YaleB face databases. On the CMU PIE face database, the proposed method can achieve classification accuracy of 85.55% using two ODVs, and 91.16% using three ODVs. On the contrast, the other methods cannot achieve classification accuracy higher than 65% using two projection vectors and 80% using three projection vectors. On the YaleB face database, the classification accuracies of the proposed method are 20% higher than other methods when dimensionality is less than three. None of appearance-based face recognition methods [4,5,7,21,22,32,33] has achieved accuracy as high as ours with such low dimensionalities.

In AR and YaleB databases, some intra-class distances are large in the image space as well as the kernel feature space. The KLDA-based feature extraction method reduces these intra-class distances and improves the clustering for classification. As a result, KLDA achieves higher accuracy than KPCA. In CMU, however, the intra-class distance is relatively smaller than the inter-class distance. Therefore, the explicit minimization of intra-class distance in KLDA has little effect on the accuracy.

Table 2 lists the training time of different methods on the three face databases. In the AR and YaleB face databases, the dimensionality of the face images is much larger than the number of training images. While the computational complexity of our method largely depends on the number of training images, the ones of [5,7,31] largely depend on the dimensionality. This explains why our method is much faster than PCA, FDA and LPP. The nonlinear methods [10,32,33] are time consuming because it (1) has an additional procedure to fix the kernel parameter and (2) must perform an implicit nonlinear transformation instead of working with the original images. Our method is slower than PCA and FDA when tested on CMU PIE database mainly because the number of training images is larger than the dimensionality of the face images in this database. However, this is not the general case, because our training face images are obtained by windowing and scaling the original 640 by 486 images. It is widely accepted that in the task of face recognition, the dimensionality of the images is normally larger than the number of images available for training. Though the computational complexity of our method grows proportionally as the number of classes grows, this number is much smaller compared with dimensionality of the face images. With computational complexity grows proportional to the dimensionality of images, our method is faster than the other methods whose computational complexity grows quadratically as the dimensionality grows.

5. Conclusion and future work

We conclude that our new approach to face recognition across pose is effective in comparison with the existing methods. Unlike the traditional methods that investigate the face images on a single manifold, our algorithm explores the differences among linear manifolds spanned by face images of different individuals. Based on the comprehensive theoretical analysis of intrapersonal relationship among face images across pose, it is found that a

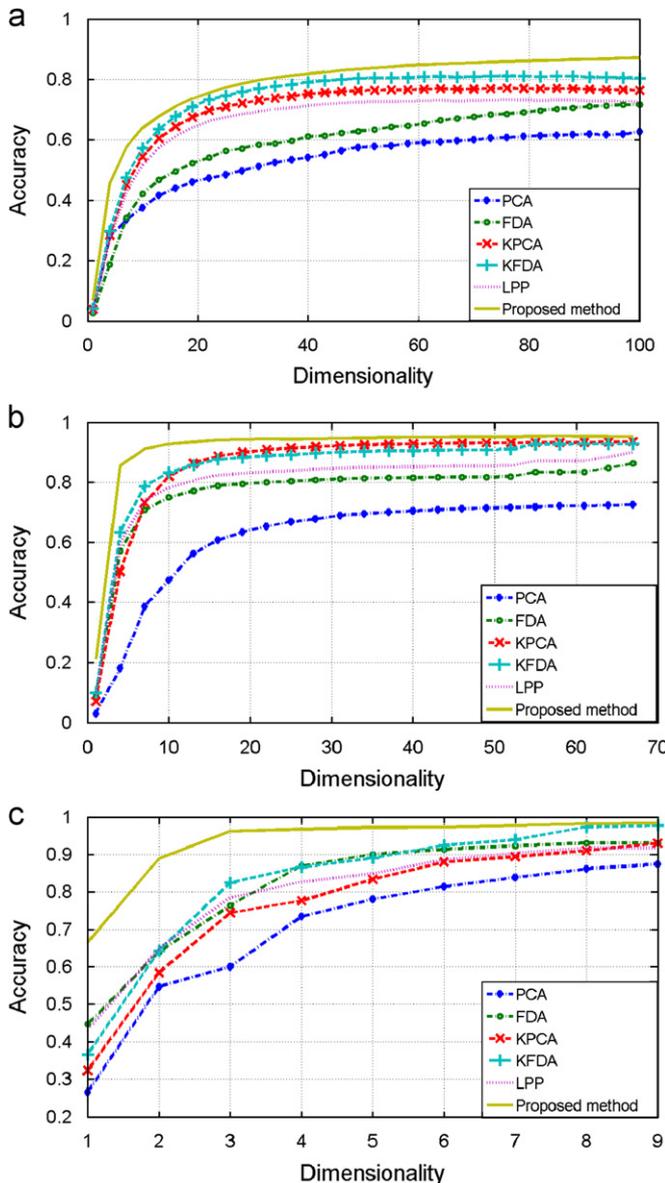


Fig. 9. Classification accuracy comparisons of different methods on different databases. (a) AR face database, (b) CMU PIE face database and (c) YaleB face database.

person's face image with a new pose can be linearly expressed by his/her gallery face images. The experimental results reported in Section 4.1 validate our observation presented in Section 2. As a result, images of one individual can be characterized by the orthogonality to certain vectors. By introducing the concept of orthogonal discriminant vector (ODV) which is the vector orthogonal to the images of the same person, we can discriminate images of one person from others. Our experimental results show that the intra-class residue is much smaller than the inter-class residue. This means that the probe images can be linearly approximated by the gallery images, which confirms the feasibility to represent the intrapersonal relationship among face images as derived in Section 2.

The existence of ODV for face recognition is further proven theoretically by introducing a two-step algorithm to calculate ODVs via solving a linear equation system. The distance between a face image and an ODV is measured by a novel distance metric to categorize face images for classification. Our experimental results demonstrate that the new measurement is more effective than the existing methods and achieves lower EERs for face verification and higher accuracy for face recognition.

It is noted that all of face images are generated from the 3D face. Therefore, both the probe and gallery face images contain pixels of the 3D face. If a probe face image is formed by pixels which are contained in the gallery face images, it can be linearly expressed by the gallery images. This is tested by the experiment in Section 4.1, where the intra-class residue is much smaller than the inter-class residue. Consequently, the intrapersonal relationship among face images is validated and our method performs well. Furthermore, as stated in Section 2.3, our method is easier to implement because it does not require the pose of the probe face image to approximate that of any gallery images.

However, if the probe face image has many pixels that are not contained in gallery face images, the intrapersonal relationship among face images across pose becomes inaccurate. The probe face image cannot be correctly represented by the linear combination of the gallery images. Thus, the distances between the probe image and ODVs do not necessarily approximate one, even if the ODVs are orthogonal to the gallery images.

As part of our future work, quantitative estimation of the reliability of an individual's pose manifold under various conditions including scare pose situations will be studied. Occlusion is a very difficult issue in face recognition. The theoretical analysis of our proposed algorithm for intrapersonal relationship among face images under different poses without occlusion would laid the foundation for further development of a powerful approach to occluded face recognition. In addition, a general model of 3-D face will be developed based on the gallery images, where the gallery images can be used for face modeling with different parameters to achieve robustness with flexibility.

Acknowledgement

The authors are most grateful for the constructive advice on the revision of the manuscript from the anonymous reviewers. The funding support from Hong Kong Government under its GRF scheme (5341/08E and 5366/09E) and the research grant from Hong Kong Polytechnic University (1-ZV5U) are greatly appreciated.

References

- [1] C. Kotropoulos, A. Tefas, I. Pitas, Frontal face authentication using morphological elastic graph matching, *IEEE Transactions on Image Processing* 9 (4) (2000) 555–560.
- [2] S.-W. Lee, J. Park, S.-W. Lee, Low resolution face recognition based on support vector data description, *Pattern Recognition* 39 (9) (2006) 1809–1812.
- [3] L. Chen, H. Man, A.V. Nefian, Face recognition based on multi-class mapping of fisher scores, *Pattern Recognition* 38 (6) (2005) 799–811.
- [4] J. Wright, A.Y. Yang, A. Ganesh, S.S. Sastry, Y. Ma, Robust face recognition via sparse representation, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 31 (2) (2009) 210–227.
- [5] P.N. Belhumeur, J.P. Hespanha, D.J. Kriegman, Eigenfaces vs. fisherfaces: recognition using class specific linear projection, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 19 (7) (1997) 711–720.
- [6] R. Gross, I. Matthews, S. Baker, Appearance-based face recognition and light-fields, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 26 (4) (2004) 449–465.
- [7] X. He, S. Yan, Y. Hu, P. Niyogi, H.-J. Zhang, Face recognition using laplacian-faces, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 27 (3) (2005) 328–340.
- [8] F.S. Samaria, A.C. Harter, Parameterization of a stochastic model for human face identification, in: *Proceedings of the Second IEEE Workshop on Applications of Computer Vision*, 1994, pp. 138–142.
- [9] Y. Xu, F. Song, G. Feng, Y. Zhao, A novel local preserving projection scheme for use with face recognition, *Expert Systems with Applications* 37 (9) (2010) 6718–6721.
- [10] J. Yang, A.F. Frangi, J.-Y. Yang, D. Zhang, Z. Jin, KPCA plus LDA: a complete kernel fisher discriminant framework for feature extraction and recognition, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 27 (2) (2005) 230–244.
- [11] D. Cai, X. He, Y. Hu, J. Han, T. Huang, Learning a spatially smooth subspace for face recognition, *IEEE Conference on Computer Vision and Pattern Recognition* (2007) 1–7.
- [12] X. Zhang, Y. Gao, Face recognition across pose: a review, *Pattern Recognition* 42 (11) (2009) 2876–2896.
- [13] S.-W. Lee, S.-H. Moon, S.-W. Lee, Face recognition under arbitrary illumination using illuminated exemplars, *Pattern Recognition* 40 (5) (2007) 1605–1620.
- [14] Y. Chang, C. Hu, M. Turk, Manifold of facial expression, in: *IEEE International Workshop on Analysis and Modeling of Faces and Gestures*, 2003, pp. 28–35.
- [15] K.-C. Lee, J. Ho, M.-H. Yang, D. Kriegman, Video-based face recognition using probabilistic appearance manifolds, in: *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, vol. 1, 2003, pp. 313–320.
- [16] S.T. Roweis, L.K. Saul, Nonlinear dimensionality reduction by local linear embedding, *Science* 290 (5500) (2000) 2323–2326.
- [17] A. Shashua, A. Levin, S. Avidan, Manifold pursuit: a new approach to appearance based recognition, in: *Proceedings of 16th International Conference on Pattern Recognition*, vol. 3, 2002, pp. 590–594.
- [18] S. Yan, D. Xu, B. Zhang, H.-J. Zhang, Q. Yang, S. Lin, Graph embedding and extension: a general framework for dimensionality reduction, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 29 (1) (2007) 40–51.
- [19] X. He, D. Cai, S. Yan, H.-J. Zhang, Neighborhood preserving embedding, in: *Tenth IEEE International Conference on Computer Vision*, vol. 2, 2005, pp. 1208–1213.
- [20] H.-T. Chen, H.-W. Chang, T.-L. Liu, Local discriminant embedding and its variants, in: *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, vol. 2, 2005, pp. 846–853.
- [21] D. Cai, X. He, J. Han, H.-J. Zhang, Orthogonal laplacianfaces for face recognition, *IEEE Transactions on Image Processing* 15 (11) (2006) 3608–3614.
- [22] J. Wang, Y. Xu, D. Zhang, J. You, An efficient method for computing orthogonal discriminant vectors, *Neurocomputing* 73 (10–12) (2010) 2168–2176.
- [23] T. Xiong, J. Ye, and V. Cherkassky, kernel uncorrelated and orthogonal discriminant analysis: a unified approach, in: *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, vol. 1, 2006, pp. 125–131.
- [24] Y. Hamamoto, A. Ohama, T. Kanaoka, S. Tomita, Orthogonal discriminant analysis based on a modified Fisher criterion [Feature extraction], in: *11th IAPR International Conference on Pattern Recognition*, vol. 2, 1992, pp. 363–366.
- [25] X. Chai, S. Shan, X. Chen, W. Gao, Locally linear regression for pose-invariant face recognition, *IEEE Transactions on Image Processing* 16 (7) (2007) 1716–1725.
- [26] D. Beymer, T. Poggio, Face recognition from one example view, in: *5th International Conference on Computer Vision*, 1995, pp. 500–507.
- [27] P.J. Phillips, H. Moon, S.A. Rizvi, P.J. Rauss, The FERET Evaluation methodology for face-recognition algorithms, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 22 (10) (2000) 1090–1104.
- [28] T. Sim, S. Baker, M. Bsat, The CMU pose, illumination, and expression database, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 25 (2) (2003) 1615–1618.
- [29] A.S. Georghiades, P.N. Belhumeur, D.J. Kriegman, From few to many: illumination cone models for face recognition under variable lighting and pose, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 23 (6) (2001) 643–660.
- [30] A.M. Martinez, R. Benavente, The AR-face database, CVC Technical Report, 24, June, 1998.
- [31] M. Kirby, L. Sirovich, Application of karhunen-loeve procedure for the characterization of human faces, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 12 (1) (1990) 103–108.
- [32] B. Scholkopf, A.J. Smola, K.R. Muller, Nonlinear component analysis as a kernel eigenvalue problem, *Neural Computation* 10 (5) (1998) 1299–1319.
- [33] Y. Xu, D. Zhang, F. Song, J.-Y. Yang, J. Zhong, M. Li, A method for speeding up feature extraction based on KPCA, *Neurocomputing* 70 (4–6) (2007) 1056–1061.

- [34] A. Teoh, A. Goh, D. Ngo, Random multispace quantization as an analytic mechanism for bihashing of biometric and random identity inputs, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 28 (2) (2006) 1892–1901.
- [35] Y.C. Feng, P.C. Yuen, A.K. Jain, A hybrid approach for generating secure and discriminating face template, *IEEE Transactions on Information Forensics and Security* 5 (1) (2010) 103–117.
- [36] Y.Z. Goh, A.B.J. Teoh, K.O.M. Goh, Wavelet-based illumination invariant pre-processing in face recognition, *Journal of Electronic Imaging* 18 (2) (2009) 1–12.
- [37] S. Biswas, G. Aggarwal, P.J. Flynn, Pose-robust recognition of low-resolution face images, 2011 IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2011), 2011, pp. 601–608.
- [38] G. Fanelli, J. Gall, L. Van Gool, Real time head pose estimation with random regression forests, 2011 IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2011), 2011, pp. 617–624.
- [39] S. Khaleghian, H.R. Rabiee, M.H. Rohban, Face recognition across large pose variations via boosted tied factor analysis, 2011 IEEE Workshop on Applications of Computer Vision (WACV), 2011, pp. 190–195.
- [40] I. Wijaya, Pose invariant face recognition using dominant frequency based holistic features and statistical classifier, Thesis, Kumamoto University, 2010 (available via Google).
- [41] L. Teijeiro-Mosquera, J.L. Alba-Castro, D. Gonzalo 'lez-Jime' nez, Face recognition across pose with automatic estimation of pose parameters through AAM-based landmarking, 2010 20th International Conference on Pattern Recognition (ICPR), 2010, pp. 1339–1342.
- [42] A. Li, S.G. Shan, W. Gao, Coupled Bias-Variance trade off for cross-pose face recognition, *IEEE Transactions on Image Processing* 21 (1) (2012) 305–315.

Jinghua WANG received his B.S. degree in Computer Science from Shandong University and his M.S. degree from Harbin Institute of Technology. He is currently a Ph.D. candidate in the Department of Computing, The Hong Kong Polytechnic University. His current research interests are in the areas of pattern recognition and image processing.

Jane YOU obtained her B.Eng. in Electronic Engineering from Xi'an Jiaotong University in 1986 and Ph.D. in Computer Science from La Trobe University, Australia in 1992. She was a lecturer at the University of South Australia and senior lecturer at Griffith University from 1993 till 2002. Currently she is an professor at the Hong Kong Polytechnic University. Her research interests include image processing, pattern recognition, medical imaging, biometrics computing, multimedia systems and data mining.

Qin LI received his B.Eng. degree in computer science from China University of Geoscience, the M.Sc. degree (with distinction) in computing from the University of North-Umbria at Newcastle, and the Ph.D. degree from the Hong Kong Polytechnic University. His current research interests include medical image analysis, biometrics, image processing, and pattern recognition.

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.