# Fractional order Singular Value Decomposition Representation for Face Recognition

Jun Liu, Songcan Chen * and Xiaoyang Tan

*Department of Computer Science and Engineering*
*Nanjing University of Aeronautics and Astronautics*
*Nanjing, 210016, P.R. China*

---

**Abstract**

Face Representation (FR) plays a typically important role in face recognition and methods such as Principal Component Analysis (PCA) and Linear Discriminant Analysis (LDA) have been received wide attention recently. However, despite of the achieved successes, these FR methods will inevitably lead to poor classification performance in case of great facial variations such as expression, lighting, occlusion and so on, due to the fact that the image gray value matrices on which they manipulate are very sensitive to these facial variations. In this paper, we take notice of the facts that every image matrix can always have the well-known Singular Value Decomposition (SVD) and can be regarded as a composition of a set of base images generated by SVD, and we further point out that the leading base images (those corresponding to large singular values) on one hand are sensitive to the aforementioned facial variations and on the other hand dominate the composition of the face image. Then based on these observations, we subtly deflate the weights of the facial variation sensitive base images by a parameter $\alpha$ and propose a novel Fractional order Singular Value Decomposition Representation (FSVDR) to alleviate facial variations for face recognition. Finally, our experimental results show that FSVDR can: 1) effectively alleviate facial variations; and 2) form an intermediate representation for many FR methods such as PCA and LDA to significantly improve their classification performance in case of great facial variations.

*Key words:* Singular Value Decomposition (SVD), Fractional order Singular Value Decomposition Representation (FSVDR), Face Representation (FR), Intermediate Representation (IR), Face Recognition.

## 1 Introduction

Although we human beings can easily detect and identify faces in a scene, it is very challenging for an automated system to achieve such objectives. The challenges become more profound when large variations exist in the face images at hand, e.g., variations in illumination conditions, viewing directions or poses, facial expression, aging, and disguises such as facial hair, glasses, cosmetics and scarves. Despite of these challenges, face recognition has drawn wide attention from researchers in areas of machine learning, computer vision, pattern recognition, neural networks, and so on, thanks to the great need for face recognition in areas of access control, information security, law enforcement and surveillance, smart cards and so on [3,6,21,45].

Over the past decades, geometric feature-based methods [4,8,14,15,18,17,39] and appearance-based methods [31,38,44,46] are the two mainly employed face recognition methods. Geometric feature-based methods extract the relative position and other parameters of distinctive features such as eyes, mouth, nose, and chin as features, while appearance-based methods directly manipulate on the gray level values of the image pixels (e.g., a face image pattern is represented as an $r \times c$ matrix $A$, where $r$ and $c$ are respectively the numbers of rows and columns) and employ statistical tools to extract features for subsequent classification. Recently, it has been witnessed a strong trend away from geometry towards statistical and appearance-based models for face recognition [22], and this trend is supported by psychological, physiological, and biological studies dealing with vision in humans and animals [11,22,36]. A typical appearance-based face recognition scheme is given in Fig. 1, where the input is a face gray value matrix and the output is the given face's class label. Generally speaking, this recognition scheme can be divided into two sequential stages: Face Representation (FR) and face classification. FR (illustrated between the two vertical dashed lines in Fig. 1) plays a key role in face

* Corresponding author, Tel: +86-25-8489-6481 Ext 12106, Fax: +86-25-8489-8069, E-mails: {s.chen, j.liu, x.tan}@nuaa.edu.cn

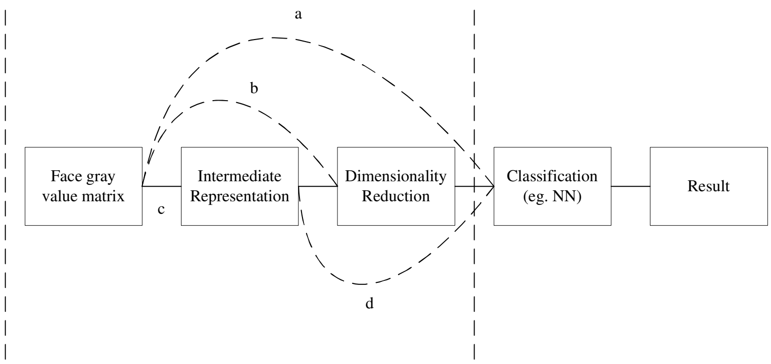


Fig. 1. Illustration of the appearance-based face recognition.

recognition, impacts the classification performance to a large extent and as a result has always been a hot topic in face recognition.

A natural choice for FR is the Original Gray Value Matrix (OGVM), whose classification scheme is depicted in route a in Fig. 1. Although OGVM is simple, it will encounter the following two problems: 1) the dimensionality of OGVM is much higher than the number of training samples (e.g., the dimensionality of a $100 \times 100$ image matrix is 10000, while the number of training samples is often quite less) which leads to the so-called curse of dimensionality and consequently depresses the generalization ability of the correspondingly trained classifier; and 2) OGVM is very sensitive to facial variations such as expression, lighting, occlusion, etc, and thus will obtain poor classification performance in case of the aforementioned facial variations.

To obtain a good representation for face images, researchers have proposed many renowned Dimensionality Reduction (DR) methods that directly manipulate on OGVM (illustrated in route b in Fig. 1). The two most well-known DR methods are Principal Component Analysis (PCA) [20,38] and Linear Discriminant Analysis (LDA) [10,13].

Kirby and Sirovich [20] showed that any particular face can be 1) economically represented along the eigenpictures coordinate space, and 2) approximately reconstructed using just a small collection of eigenpictures and their corresponding projections ('coefficients'). Turk and Pentland [38] applied PCA technique to face recognition, and proposed the well-known Eigenfaces method. A recent major improvement on PCA is to directly manipulate on two-dimensional matrices (not one-dimensional vectors as in traditional PCA), e.g., Two-dimensional PCA (2DPCA) [40], Generalized Low Rank Approximation of Matrices [41], Non-iterative Generalized Low Rank Approximation of Matrices (NIGLRAM) [28] and so on. The advantages of manipulating on two-dimensional matrices rather than one-dimensional vectors are:[40] 1) it is simpler and straightforward to use for image feature extraction; 2) it is better in terms of classification performance; and 3) it is computationally more efficient. Based on the viewpoint of minimizing reconstruction error, the above PCA-based methods [28,38,40,41] are unsupervised methods that do not take the class labels into consideration.

Taking the class labels into consideration, LDA aims at projecting face samples to a subspace where the samples belonging to the same class are compact while those belonging to different classes are far away from each other. The major problem in applying LDA to face recognition is the so-called Small Sample Size (SSS) problem (namely, the number of samples is far less than sample dimensionality), which leads to the singularities of the within-class and between-class scatter matrices. Recently, researchers have exerted great endeavor to deal with this problem. In [1,34], a PCA procedure was applied prior to the LDA procedure, which led to the well-known PCA+LDA or Fisherfaces method. In [5,27], samples were first projected to the null space of the within-class scatter matrix and then LDA was applied in this null space to yield the optimal (infinite) value of the Fisher's Linear Discriminant criterion, which led to the so-called Discriminant Common Vectors (DCV) method. In [26,42], LDA was applied in the range space of the between-class scatter matrix to deal with the SSS problem, which led to the Linear Discriminant Analysis via QR decomposition (LDA/QR) method.

These DR methods have been proven to effectively lower the dimensionality of OGVM. Furthermore, in face recognition, PCA and LDA have become defacto baseline approaches. However, despite of the achieved successes, these FR methods will inevitably lead to poor classification performance in case of great facial variations such as expression, lighting, occlusion and so on, due to the fact that the OGVM $A$ on which they manipulate is very sensitive to these facial variations. To mitigate this problem, in this paper, we propose a novel Fractional order Singular Value Decomposition Representation (FSVDR), which acts as an Intermediate Representation (IR) between OGVM and DR (see route c in Fig. 1) for face recognition.

In literature, there have been a number of approaches that form an IR between OGVM and DR for face recognition. In [24,25], the authors first obtained the Gabor wavelet representation for face images and then utilized it as an IR for subsequent DR utilizing the enhanced Fisher Linear Discriminant Model [23] or Independent Component Analysis [9]. The Gabor wavelet representation was reported to capture local structure corresponding to spatial frequency (scale), spatial localization, and orientation selectivity. Besides Gabor wavelet, Er et al.[12] employed Discrete Cosine Transform (DCT) as an IR and then applied LDA subsequently for DR. It was reported that the low-frequency DCT coefficients accounted for a large area nonuniform illumination variations, and consequently the nonuniform illumination effect can be reduced by discarding several low-frequency DCT coefficients. Our FSVDR distinguishes from Gabor wavelet and DCT in that 1) FSVDR employs the well-known SVD, not Gabor wavelet or DCT and 2) FSVDR needs to tune a parameter $\alpha$ to yield an IR and the choice of the parameter is both database and DR method dependent, while neither Gabor wavelet nor DCT has such parameter to be tuned. Besides, it is worthwhile to point out that, since FSVDR is still human face like (see Section 2.2), it can be applied prior to Gobor wavelet or DCT for forming an new IR for face recognition.

Furthermore, in face recognition, there have been studies of employing SVD to obtain representation for face images. Hong [19] proposed to apply SVD to each OGVM to obtain Singular Values(SVs) to represent this face image, and then to perform classification based on these SVs. Cheng et al.[7] made use of the SVs as an IR, and then employed an optimal discriminant transformation to transform the SVs into a new space for subsequent classification. Although good classification performance was reported in [7], Tian et al.[37] pointed out that the SVs contained little useful information for face recognition and attributed the good performance reported in [7] to the small testing database. Comparing FSVDR with SVs [7,19], it is clear that the two methods are quite different. More specifically, the representation by SVs only employs the singular values, while our FSVDR utilizes not only the singular values, the left and right transformation matrices but also a parameter $\alpha$ to yield the so-needed Intermediate Representation (IR). Besides SVs, SVD was utilized to generate multiple virtual samples for face recognition with one sample per class in [43], which is a non-parametric method and is different from FSVDR.

Now, it is worthwhile to summarize our contributions in this paper as follows:

1) We take notice of the fact that every image matrix has the well-known Singular Value Decomposition (SVD) and point out both theoretically and experimentally that, for any given single face matrix $A$, its leading base images (those corresponding to large singular values) are sensitive to facial variations such as illumination, occlusions, etc. Meanwhile, we show experimentally that these leading base images in fact dominate the composition of $A$.

2) Based on the observations made in 1), our proposed FSVDR subtly deflates the weights of the facial variation sensitive base images by introducing a parameter $\alpha$ and effectively alleviates the influence of the facial variations on face recognition. Furthermore, the relationship between FSVDR and OGVM is theoretically studied.

3) We employ FSVDR as an Intermediate Representation (IR) for many well-known DR methods such as PCA, 2DPCA, NIGLRAM, PCA+LDA, DCV and LDA/QR. And the experimental results show that, by employing FSVDR as an IR, the classification performance of these methods can be significantly improved in case of great facial variations.

4) Our FSVDR offers a both database and DR method dependent IR, namely for different methods and different databases, the optimal value for parameter $\alpha$ should be different. This is a very important characteristic of FSVDR, and is in accord with the main argument made in [31,32] that the DR methods can only obtain good performance under certain conditions. Further, based on our previous work in [27], we offer a heuristic criterion for choosing the parameter $\alpha$ for the LDA-based methods.

In what follows, we present our proposed FSVDR in detail in Section 2, carry out extensive experiments to verify the effectiveness of the proposed FSVDR in Section 3, and draw an conclusion to this paper in Section 4.

## 2 Fractional order Singular Value Decomposition Representation

We will first analyze the SVD on each face image matrix pattern $A$ in Section 2.1, and propose FSVDR in Section 2.2.

Mathematically, every $r \times c$ ($r \geq c$ without loss of generality) OGVM $A$ can always have the Singular Value Decomposition (SVD) [16] as

$$A = \tilde{U}\tilde{S}\tilde{V}^T, \tag{1}$$

where $\tilde{U} = [u_1, u_2, ..., u_r]$, $\tilde{V} = [v_1, v_2, ..., v_c]$, $\tilde{S} = (D\ 0)^T$, $D = diag(\lambda_1, \lambda_2, ..., \lambda_c)$, 0 is a $c \times (r - c)$ zero matrix and $\lambda_i$'s are the singular values in a non-increasing order.

Further, assuming the rank of $A$ to be $k$, we have

$$A = USV^T = \sum_{i=1}^{k} \lambda_i u_i v_i^T, \tag{2}$$

where $S = diag(\lambda_1, \lambda_2, ..., \lambda_k)$, $U = [u_1, u_2, ..., u_k]$ and $V = [v_1, v_2, ..., v_k]$.

We have the following five properties for the SVD applied to $A$.

**Property 1** *Denote $A$ as*

$$A = [a_1, a_2, \ldots, a_r]^T, \tag{3}$$

*where $a_i^T$ is a $1 \times c$ row vector that represents the $i$-th row of image matrix $A$, and we have*

$$V^T C_{row} V = S^2, \tag{4}$$

*where*

$$C_{row} = A^T A = \sum_{i=1}^{r} a_i a_i^T. \tag{5}$$

*That is to say, $v_j$ is the eigenvector of the covariance matrix $C_{row}$ corresponding to eigenvalue $\lambda_j^2$, $j = 1, 2, \ldots, k$.*

**Property 2** *Denote $A$ as*

$$A = [a_1, a_2, \ldots, a_c], \tag{6}$$

*where $a_i$ is an $r \times 1$ column vector that represents the $i$-th column of image matrix $A$, and we have*

$$U^T C_{col} U = S^2, \tag{7}$$

*where*

$$C_{col} = AA^T = \sum_{i=1}^{c} a_i a_i^T, \tag{8}$$

*namely, $u_j$ is the eigenvector of the covariance matrix $C_{col}$ corresponding to eigenvalue $\lambda_j^2$, $j = 1, 2, \ldots, k$.*

**Property 3** *$u_i v_j^T$, $i = 1, 2, \ldots, r$, $j = 1, 2, \ldots, c$ form a set of orthonormal base matrices for the space $R^{r\times c}$ and can be termed as base images.*

**Property 4** *When projecting $A$ to the base images $u_i v_j^T$, $i = 1, 2, \ldots, r$, $j = 1, 2, \ldots, c$, we have: 1) the coefficients are all nonnegative; 2) the coefficients are $\lambda_i$'s in the base images $u_i v_i^T, i = 1, 2, \ldots, k$; and 3) the coefficients are zeros in the base images complement to those in 2).*

**Property 5** *When $A$ is regarded as the composition of $uv^T$ like bases, where $u$ and $v$ are $r \times 1$ and $c \times 1$ column vectors respectively, then the least number of such bases is $k$ and $u_i v_i^T$, $i = 1, 2, \ldots, k$ provide a solution.*

The proofs of these properties are simple and omitted here. Property 1 and 2 reveal that: 1) $u_i$'s corresponding to large $\lambda_i$'s capture the great horizontal variances among the column vectors of $A$, and 2) $v_i$'s corresponding to large $\lambda_i$'s capture the great vertical variances among the row vectors of $A$. Property 3 defines the concept of base images, and it is easy to conclude from property 1-3 that the leading base images (those corresponding to leading singular values) capture the great variances in the face image $A$ itself. Property 4 and 5 further offer some characteristics of these base images. Now, we experimentally demonstrate two important properties of the leading base images as follows:

### *2.1.1 Leading base images being sensitive to facial variations*

We choose six images (under distinct facial variations) from the well-known AR face database [30] and show the original face images and the corresponding three leading base images in Fig. 2, from which we can clearly observe that the first base image $u_1 v_1^T$ well captures: 1) the light condition (left light on or right light on) of the face images; and 2) the horizontal position of eyebrows, mouth (if not masked) and occlusions such as wearing glasses and wearing scarves. The reason behind these phenomena is as described in property 1 and 2, namely $u_1$ and $v_1$ respectively capture the greatest variance of the column

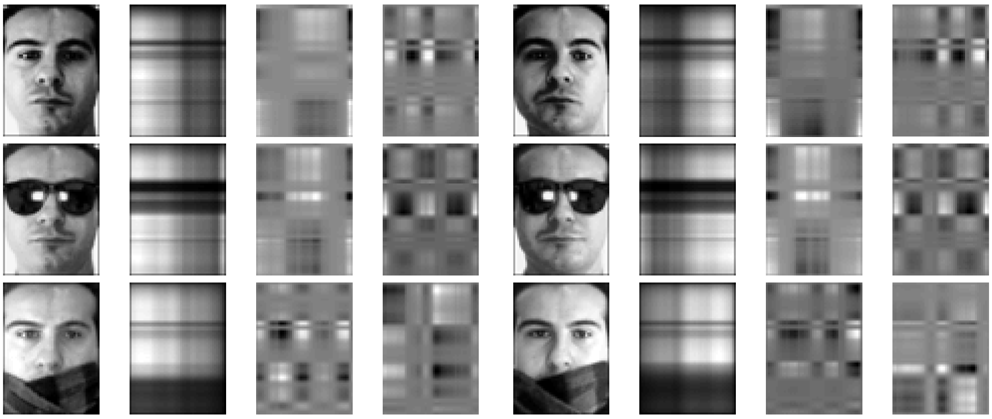

Fig. 2. An illustration of three leading base images of six face images under different facial conditions. The first row: left light on, right light on; The second row: left light on & wearing glasses, right light on & wearing glasses; The third row: left light on & wearing scarves, right light on & wearing scarves.

and row vectors of the face images and $u_1v_1^T$ captures the greatest variance among $A$ itself. Theoretically, $u_2v_2^T$ and $u_3v_3^T$ also capture the great variances (although less than those by $u_1v_1^T$) among the face image itself, but they are less obvious than $u_1v_1^T$ in vision which may attribute to the fact that our eyes are good at detecting some larger variations (as in $u_1v_1^T$) while not for some smaller variations (as in $u_2v_2^T$ and $u_3v_3^T$). Moreover, a thorough look at the second row of Fig. 2 show that $u_2v_2^T$ does reveal the two reflected light spots on the black glasses. Finally, due to the fact that the leading base images capture the great variances within the face image itself, they are naturally more sensitive to facial variations such as expression, lighting, occlusion and so on, which can be typically read from $u_1v_1^T$ in Fig. 2.

### *2.1.2 Leading base images dominating the composition of face image*

The leading base images contribute a great deal to the composition of $A$, for which sake we define the cumulative energy contained in these leading images as

$$e_i = \frac{\sum_{j=1}^{i} ||\lambda_i u_j v_j^T||_F^2}{||A||_F^2}, \tag{9}$$

where $||.||_F^2$ is the squared *Frobenius* norm. We plot the cumulative energy $e_i$ in Fig. 3, where the six face images are just the corresponding six face images shown in Fig. 2. From Fig. 3, one can easily read: 1) the first base images of the six face images here all possess an energy of over 85%; 2) the

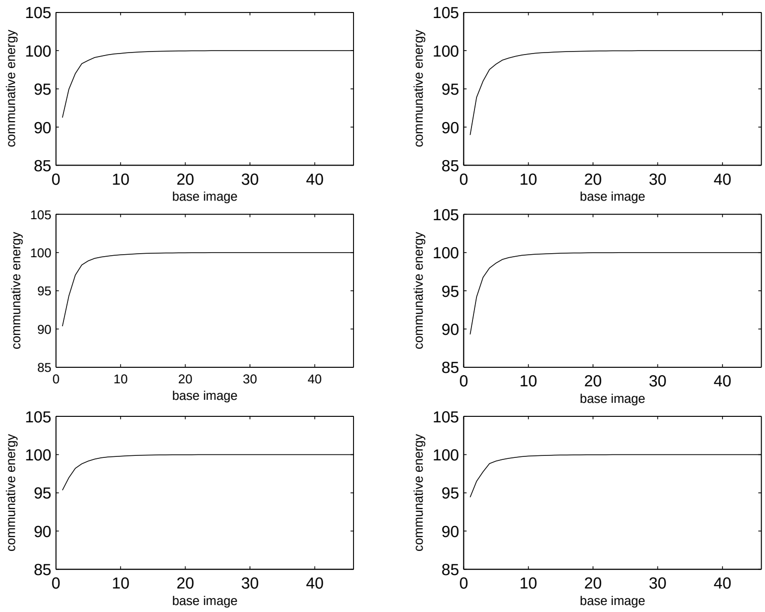


Fig. 3. Cumulative energy of the base images for the six images in Fig. 2.

first few leading base images (e.g., 10) almost occupy all the energy (e.g., the cumulative energy $e_{10} > 99\%$) contained in $A$. As a result, the leading base images dominate the composition of face image.

From the discussion in Section 2.1.1 and 2.1.2, we know that if we handle the facial variation sensitive leading base images nicely, the great facial variations within each face image can be effectively alleviated, which will be discussed in our proposed FSVDR in the next subsection.

### *2.2 Fractional order Singular Value Decomposition Representation*

To alleviate the facial variations on face images, we propose a novel Fractional order Singular Value Decomposition Representation (RSVDR), whose underlying ideas are that 1) the weights of the leading base images $u_i v_i^T$ should be deflated, since they are very sensitive to the great facial variations within the image matrix $A$ itself; 2) the weights of base images $u_i v_i^T$ corresponding to relatively small $\lambda_i$'s should be inflated, since they may be less sensitive to the facial variations within $A$, which can be read from Property 1 and 2; and 3) the order of the weights of the base images $u_i v_i^T$ in formulating the new representation $B$ should be retained. More specifically, for each face image matrix

$A$ which has the SVD in (2), its FSVDR $B$ is defined as:

$$B = US^{\alpha}V^{T} \tag{10}$$

where $U$, $S$ and $V$ are the corresponding matrices in equation (2), and in order to achieve the above underlying ideas, $\alpha$ is a fractional parameter that satisfies:

$$0 \leq \alpha \leq 1 \tag{11}$$

It is easy to obtain the following properties with regard to FSVDR:

**Property 6** *The rank of FSVDR $B$ is $k$, i.e., identical to the rank of $A$.*

**Property 7** *$u_i v_i^T, i = 1, 2, \ldots, k$ form a set of $uv^T$ like base images for the FSVDR $B$.*

**Property 8** *When $0 < \alpha \leq 1$, the transformation of $A$ to $B$ is a bijection; when $\alpha = 0$, the transformation of $A$ to $B$ is not a bijection, but a surjection.*

**Property 9** *Let $\beta = (\alpha - 1)/2$ and define $(V^T S V)^{\beta} = V^T S^{\beta} V$, then we have $B = A(A^T A)^{\beta}$, namely the FSVDR $B$ utilizes some cross product information contained in $A$, which has somewhat flavor of employing high-order information.*

Property 6 shows that an intrinsic characteristic of $A$, the rank, is retained in the FSVDR $B$. From Property 7 and Property 5, we know that the FSVDR $B$ in fact has the same $uv^T$ like base images as $A$, and considering the fact that these base images are the components to compose $A$ and $B$, we can say that the information in $A$ is passed to $B$ nicely. Property 8 further shows the close relationship between the FSVDR $B$ and the original representation $A$. Property 9 says that the FSVDR $B$ utilizes some cross product information contained in OGVM $A$.

Now, we illustrate some FSVDR face images under different $\alpha$ in Fig. 4, from which we can observe that:

1) The FSVDR $B$ is still like human face;

2) The FSVDR deflates the lighting condition in vision. Taking the two face images in the first row for example, when $\alpha$ is set to 0.4 and 0.1, from the

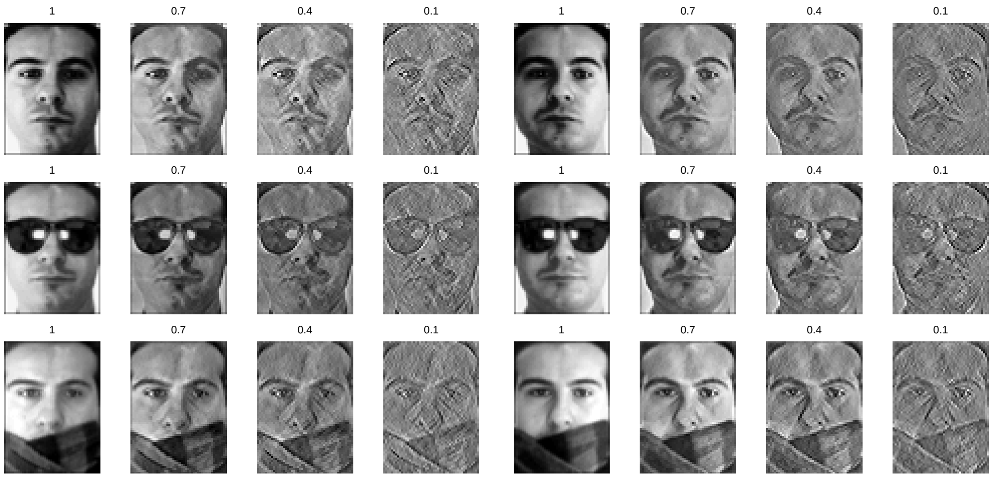


Fig. 4. An illustration of six face images under different facial conditions employing FSVDR, where $\alpha = 1$ is just the OGVM.

FSVDR $B$ alone, one can hardly tell whether the original face image matrix $A$ is of left light on or right light on;

3) The FSVDR $B$ reveals some facial details. In the original face images $A$ presented in the first row, neither the right eyeball of the left face image nor the left eyeball of the right face image is visible, however, when setting $\alpha$ to 0.4 and 0.1 in FSVDR, the eyeballs become visible. Moreover, the jowls become more obvious in the FSVDR, as revealed in the face images of the first and the second rows;

4) The FSVDR $B$ can not remove the occlusions such as glasses and scarves, due to the fact that it shares the same base images as the OGVM $A$. However, we can see from the second and third rows that the contrast between the face images and the occlusions decreases. As a result, in the FSVDR, the occluded images will be nearer to the faces without occlusions compared to those in the OGVM $A$.

The most important characteristic of the FSVDR lies in that it can form an Intermediate Representation (IR) for many DR methods such as PCA, LDA and so on (illustrated in route c of Fig. 1). Furthermore, when employing FSVDR as an IR for DR methods, the time complexities in training and testing are almost the same as the original DR methods. To elaborate this, we take DCV as an example in the following analysis. According to [5,27], the time complexity for training $N$ samples with dimensionality $d = rc$ is $O(N^2 d)$, and the time complexity in testing any given unknown sample is $O(dC)$, where $C$ is the number of classes. For DCV based on FSVDR, on one hand, it consumes additional $O(Nd \max(r, c))$ in computing the FSVDR for $N$ samples, where $\max(r, c)$ is usually smaller than $N$, and thus the time complexity in training is still $O(N^2 d)$, the same as original DCV; on the other hand, for any unknown sample, it takes additional $O(\max(r, c)d)$ in computing its FSVDR, and thus the time complexity in testing is also $O(dC)$ since $\max(r, c)$ is usually comparable to or less than $C$. And we will verify the effectiveness of FSVDR in the next section.

## 3 Experiments

In this section, we carry out extensive experiments to show that: 1) when directly applied to face recognition, FSVDR can yield significantly better classification performance than OGVM and the SVs (see Section 3.2); and 2) as an Intermediate Representation (IR), the FSVDR can significantly improve the classification performance of quite a few DR methods such as PCA, PCA+LDA, LDA/QR, DCV, 2DPCA and NIGLRAM (see Section 3.3). Furthermore, we will experimentally visualize the samples by the FSVDR and OGVM to show the benefit brought by FSVDR in Section 3.4 and dwell on the problem of parameter choice in Section 3.5. Before reporting the experimental results, we first describe the database and the experimental setting in Section 3.1.

### *3.1 Database and experimental setting*

#### *3.1.1 Database description*

We carry out experiments on three renowned face databases: AR [30], FERET [33] and YALE [1].

The AR database consists of over 4000 color images of 126 person's faces (70 men and 56 women). Each person has 26 different images which were grabbed in two different sessions separated by two weeks, and 13 images in each session were recorded. The 13 images are respectively of neutral expression, smile,

Table 1
Data partition on AR face database

| *Category* | *Training* | *Testing* |
|---|---|---|
| AR1 | a, b, c, d, e, f, g | n, o, p, q, r, s, t |
| AR2 | a, b, c, d, e, f, g | h, i, j |
| AR3 | a, b, c, d, e, f, g | k, l, m |
| AR4 | a, b, c, d, e, f, g, h, i, j, k, l, m | n, o, p, q, r, s, t, u, v, w, x, y, z |

anger, scream, left light on, right light on, both light on, occlusion by glasses & left light on, occlusions by glasses & right light on, occlusions by glasses & both light on, occlusion by scarves & left light on, occlusions by scarves & right light on, occlusions by scarves & both light on. Fig. 5 illustrates the 26 image faces under different facial variations from one subject in AR face database. In our experiments here, we use a subset of the AR face database provided and preprocessed by Martinez [30]. This subset contains 2600 image faces corresponding to 100 person (50 men and 50 women), where each person has 26 different images under the aforesaid conditions. The original resolution of these image faces is $165 \times 120$. Here, for computational convenience, we resize them to $66 \times 48$, and the gray level values are rescaled to [0 1]. As can been seen from Fig. 5, the AR face database is very challenging. Here, we carry out four independent experiments, AR1, AR2, AR3 and AR4, where the training and testing samples are listed in Table 1. From Table 1 and Fig. 5, we can observe that: AR1 evaluates the classification performance over time with variations in expression and lighting conditions; AR2 tests the classification performance in case of occlusions by wearing glasses; AR3 evaluates the the classification performance in case of occlusions by wearing scarves; AR4 tests the classification performance over time with great variations in expression, lighting conditions and occlusions such as wearing glasses and scarves.

The FERET database is one of the most well-known face recognition benchmarks. The Colour FERET database contains a total of 11338 facial images corresponding to 994 subjects, and the Grey FERET contains a total of 14051 greyscale images corresponding to 1209 subjects. Here, we carry out experiments on the hardest subset of FERET Tests September 1996, whose testing samples have great facial variations in illumination. More specifically, we employ the gallery set that contains 1196 face images as training set and the fafc

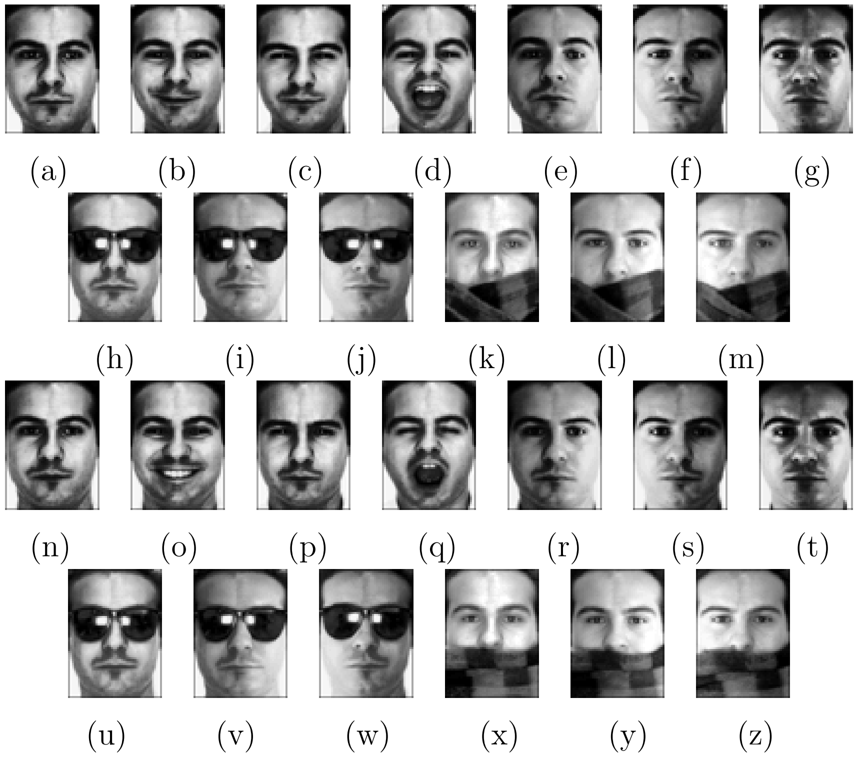


Fig. 5. An illustration of 26 images of one subject from AR face database.

set that has 194 face images as testing set. The face images are preprocessed according to the CSU Face Identification Evaluation System [2] with a resolution of $75 \times 65$. The challenges of this FERET subset are: 1) a large number of subjects (1196) in the training set, 2) one training sample per class and 3) great illumination variations in the testing set. Due to difficulty of this subset, we follow the CSU Face Identification Evaluation System [2] to report the Rank $k$ classification, where the testing sample is considered to be correctly classified so long as it belongs to the same class as one of its $k$ nearest neighbor samples in the training set.

The YALE face database contains 165 gray level face images of 15 persons. There are 11 images per subject, and these 11 images are respectively under the following different facial expression or configuration: center-light, wearing glasses, happy, left-light, wearing no glasses, normal, right-light, sad, sleepy, surprised, and wink. In our experiment, the images are cropped to a size of $50 \times 50$, and the gray level values of all images are rescaled to [0 1]. Fig. 6 shows the eleven images of one person from this database. On YALE face database, we perform two different experiments, YALE1 and YALE2, where the training and testing samples are given in Table 2. From Table 2 and Fig.

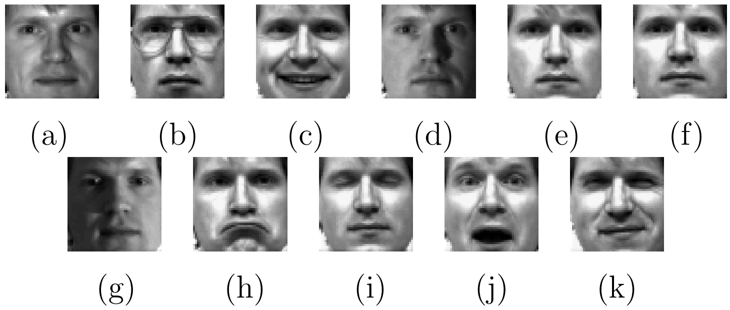


Fig. 6. An illustration of 11 images of one subject from YALE face database

Table 2
Data partition on YALE face database

| *Category* | *Training* | *Testing* |
|---|---|---|
| YALE1 | c, f, h, i, j, k | a, b, d, e, g |
| YALE2 | a, b, e, f, i, k | c, d, g, h, j |

6, we know that: YALE1 evaluates the classification performance in case of different lighting conditions and wearing glasses; and YALE2 evaluates the classification performance in case of some distinct expressions and lighting conditions.

### *3.1.2 Experimental setting*

In DR methods such as PCA+LDA, DCV and LDA/QR, they can achieve at most $C-1$ projection vectors, thus the reduced dimensionality is usually set to $C-1$. For the unsupervised method such as PCA, 2DPCA, NIGLRAM, how to choose an adequate number of projection vectors in order to yield the optimal generalization ability (or classification performance on unknown samples) is still an open problem. Here, without loss of generality, we set the reduced dimensionality for PCA employing two ways : 1) PCA1, where the dimensionality is set to $C-1$; and 2) PCA2, where the dimensionality is chosen such that 90% information in the sense of reconstruction are retained. We set the parameter $l$ to 10 in 2DPCA, and set the parameters $l_1$ and $l_2$ both to 20 in NIGLRAM.

When the features are extracted, we employ a Nearest Neighbor (NN) classifier for reporting the classification performance, where the underlying distance metric is the standard *Euclidian* distance.

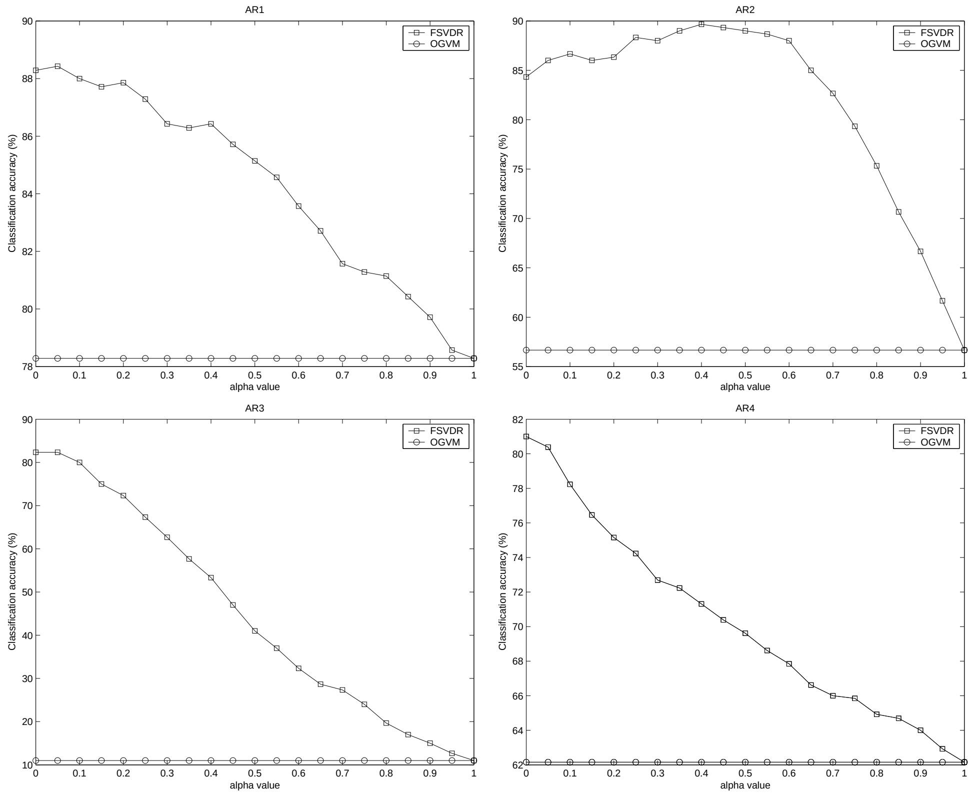


Fig. 7. Classification performance comparison between employing FSVDR and OGVM directly for face recognition on AR face database.

In our proposed FSVDR, one key problem is to choose the adequate value for $\alpha$. Instead of tackling this problem from the very beginning, we just let $\alpha$ change from 0 to 1 incremented by 0.05, and as a result, we will get a total of 21 classification results for a given method and a given experiment.

### *3.2 Comparison with OGVM and SVs for classification*

In this subsection, we first compare the classification performance of FSVDR (illustrated in route d of Fig. 1, where the IR is FSVDR) and OGVM (illustrated in route a of Fig. 1) and report the experimental results in Fig. 7 and Fig. 8. Moreover, for illustration convenience, we list the best classification accuracies of FSVDR as a column in Table 3 with the name NN.

From Fig. 7, Fig. 8 and Table 3 we can observe that the classification accuracies under FSVDR are significantly higher than those by OGVM. More specifically: 1) on AR1, the classification accuracy based on OGVM is 78.3%,

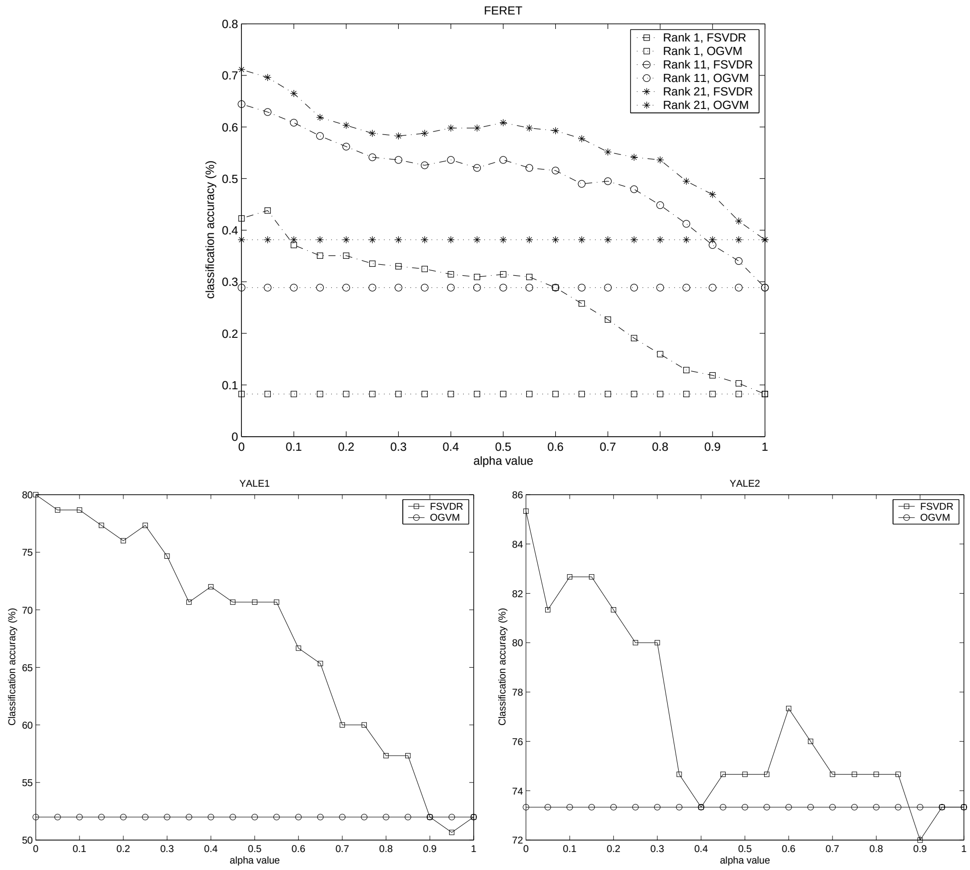


Fig. 8. Classification performance comparison between employing FSVDR and OGVM directly for face recognition on FERET and YALE face database.

and FSVDR achieves a best classification accuracy of 88.4%, an improvement of 10.1%, which attributes to that FSVDR can alleviate the variations caused by expression, lighting and time duration; 2) on AR2, OGVM can only achieve a classification accuracy of 56.7% which attributes to the occlusions by wearing glasses. However, the best classification accuracy for FSVDR is 89.7%, 33% higher than that of OGVM, which clearly shows that FSVDR can alleviate the occlusions caused by wearing glasses; 3) on AR3, OGVM can only achieve a classification accuracy of 11.0% which attributes to the occlusions by wearing scarves. For FSVDR, it can achieve a best classification accuracy of 82.3%, an improvement of 71.3%. And the reason behind the classification improvements is due to that FSVDR can effectively alleviate the occlusions caused by wearing scarves. Furthermore, comparing the classification performance (of either OGVM or FSVDR) on AR2 and AR3, we can see that in face recognition, the occlusion caused by wearing scarves is much more harder than that by wearing glasses, which is in accord with the argument made in [29,35];

4) on AR4, FSVDR achieves a best classification accuracy of 81.0%, a significant improvement of 18.8% compared to OGVM's 62.2%, which attributes to FSVDR's alleviation of the facial variations in expression, lighting conditions, and occlusions on wearing glasses and scarves; 5) on FERET, OGVM operates very poorly, with Rank 1, 11 and 21 classification accuracies only being 8.2%, 28.9% and 38.1%. This is due to a large number (1196) of distinct subjects in the training set and the great illumination variation in the testing samples. Favored by the proposed FSVDR, the best Rank 1 classification is 42.3%, which is higher than the Rank 21 classification of OGVM. Meanwhile, the best Rank 11 and 21 classification accuracies of FSVDR are significantly higher than those of OGVM; 6) on YALE1, OGVM achieves a classification accuracy of 52.0%, 28% lower than FSVDR's best classification accuracy of 80.0%, which again verifies that FSVDR can alleviate such facial variations caused by lighting and wearing glasses; 7) on YALE2, FSVDR achieves a best accuracy of 85.3%, 12% higher than OGVM's 73.3%, which attributes to FSVDR's ability to alleviate the distinct expression variations and lighting changes.

Finally, the classification accuracies by directly applying SVs for classification on AR1, AR2, AR3, AR4, FERET, YALE1 and YALE2 are respectively 11.6%, 1.7%, 1.7%, 9.6%, 0.5%, 27.3% and 44.0%, which is in accord with the argument in [37] that the SVs contain little information for classification. Obviously, our FSVDR can achieve significantly better classification performance than FR by SVs.

### *3.3 An intermediate stage for several renowned DR methods*

After verifying that FSVDR can yield significantly better classification performance compared to OGVM and SVs, we now move on to show that the FSVDR can improve the classification performance of two categories of renowned DR methods: LDA-based methods and PCA-based methods. More specifically, we will carry out experiments on the LDA-based methods such as PCA+LDA, DCV and LDA/QR in Section 3.3.1 and the PCA-based methods such as PCA, 2DPCA and NIGLRAM in Section 3.3.2.

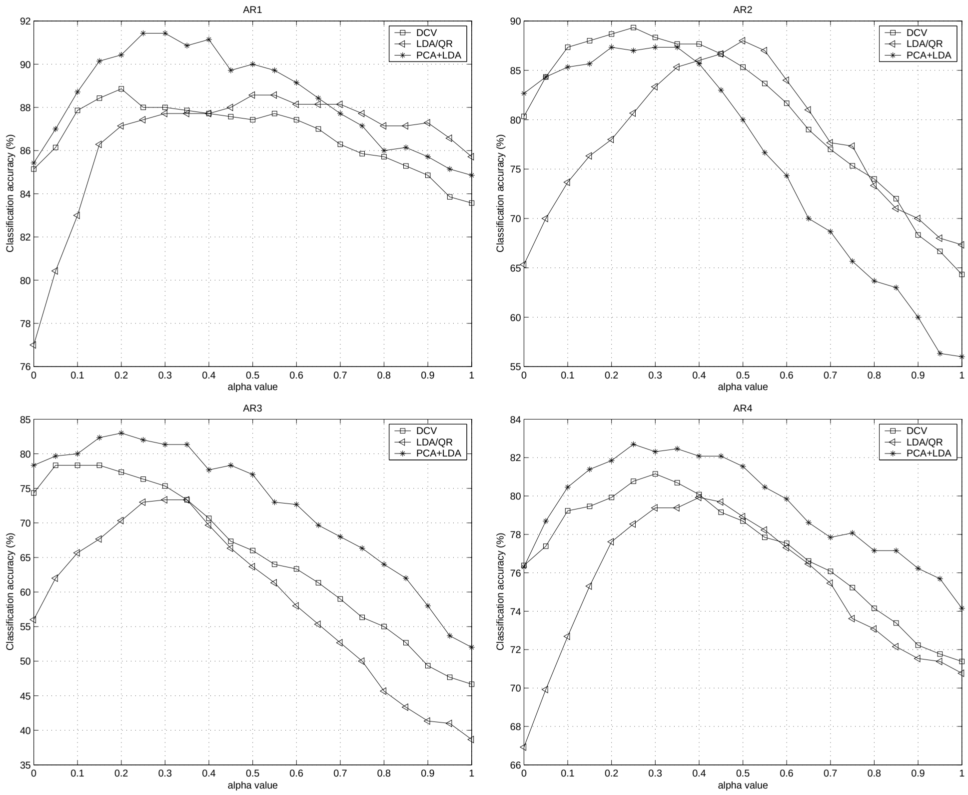


Fig. 9. FSVDR as an IR for LDA-based methods on AR face database.

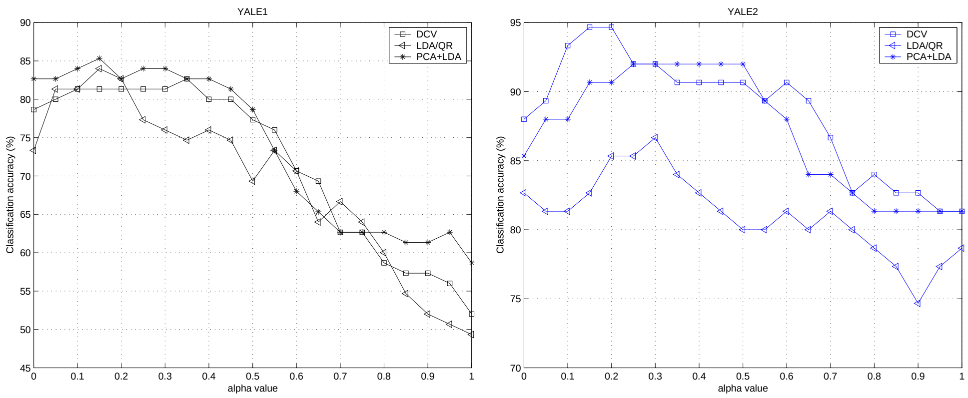


Fig. 10. FSVDR as an IR for LDA-based methods on YALE face database.

### *3.3.1 LDA-based methods*

We report the experimental results of employing FSVDR as an IR for LDA-based methods in Fig. 9 and Fig. 10. For illustration convenience, we list the best classification accuracies of FSVDR under optimal parameter $\alpha$ in Table 3, Fig. 13 and Fig. 14. Firstly, from Table 3, Fig. 13 and Fig. 14, we can observe that, when performing the LDA-based methods based on OGVM for face recognition, they can generally yield significantly better classification accuracies compared to those obtained on OGVM. For example, on AR3, the classification accuracy based on OGVM is only 11.0%, but DCV, LDA/QR and PCA+LDA-based on OGVM respectively yield classification accuracies of 46.7%, 38.7% and 52.0%, or improvements of 35.7%, 27.7% and 41.0%. However, due to the great facial variations, the LDA-based methods based on OGVM can not achieve a satisfactory classification performance on AR3. Secondly, from Fig. 9, Fig. 10, Table 3, Fig. 13 and Fig. 14, we can observe that when employing FSVDR as an intermediate representation, DCV, LDA/QR and PCA+LDA can achieve significantly higher classification accuracies than those based on OGVM. More specifically, 1) on AR1, AR4, and YALE2, the LDA-based methods utilizing FSVDR as an IR almost achieve 5.0% improvement in classification accuracies than the LDA-based methods based on OGVM; 2) on AR2 and YALE1, the LDA-based methods employing FSVDR as IR witness over 20.0% improvement in classification accuracies compared to those based on OGVM; 3) on AR3, the classification accuracies of the LDA-based methods using FSVDR as IR are more than 30.0% higher than those based on OGVM.

As a result, the LDA-based methods can benefit from FSVDR, which acts as an IR to effectively alleviate facial variations such as expression, lighting, occlusions on wearing glasses and scarves and so on.

### *3.3.2 PCA-based methods*

After verifying that the FSVDR can improve the classification performance of the LDA-based methods, we move on to show that the PCA-based methods can benefit from FSVDR, which acts as an IR. The experimental results under different $\alpha$'s are reported in Fig. 11 and Fig. 12, and the best classification accuracies based on the FSVDR are reported in Table 3, Fig. 13 and Fig. 14. From these results, we can clearly see that the classification performance of the PCA-based methods employing FSVDR as an IR is significantly better than that of the PCA-based methods based on OGVM. More specifically, on one hand, from the results reported on AR2, we can get that, the best classification accuracies based on FSVDR for PCA1, PCA2, 2DPCA and NIGLRAM are respectively 88.0%, 90.0%, 91.3% and 91.0%, which are 37.3%, 43.7%, 42.3% and 38.3% higher than those based on OGVM. On the other hand, the most

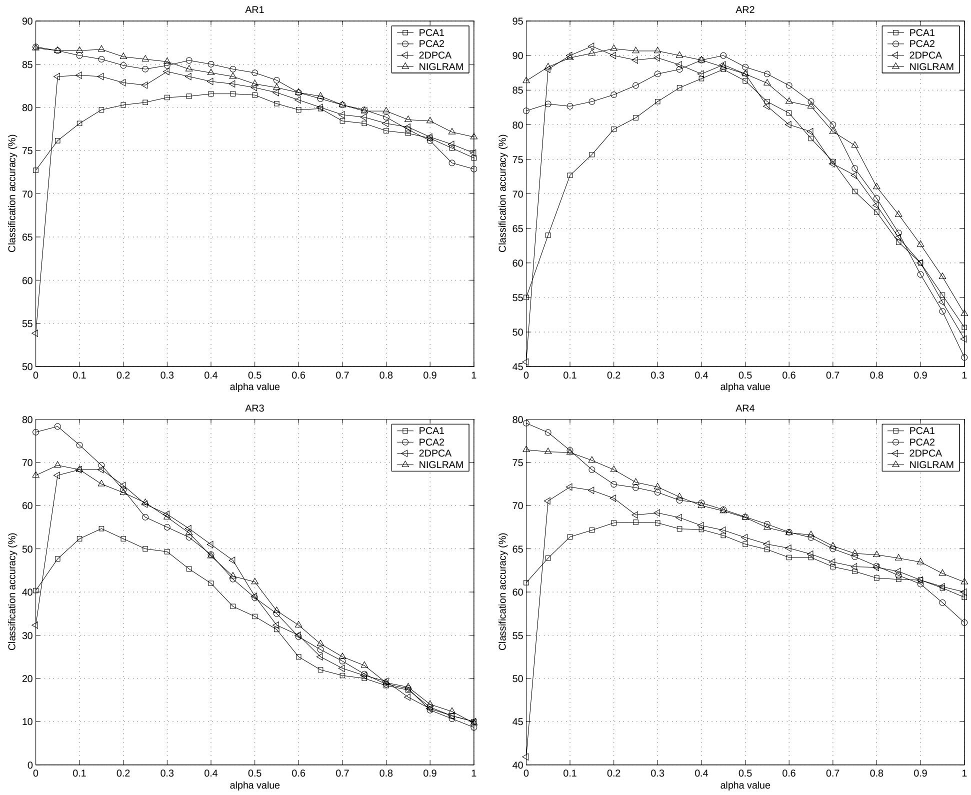


Fig. 11. FSVDR as an IR for PCA-based methods on AR face database.

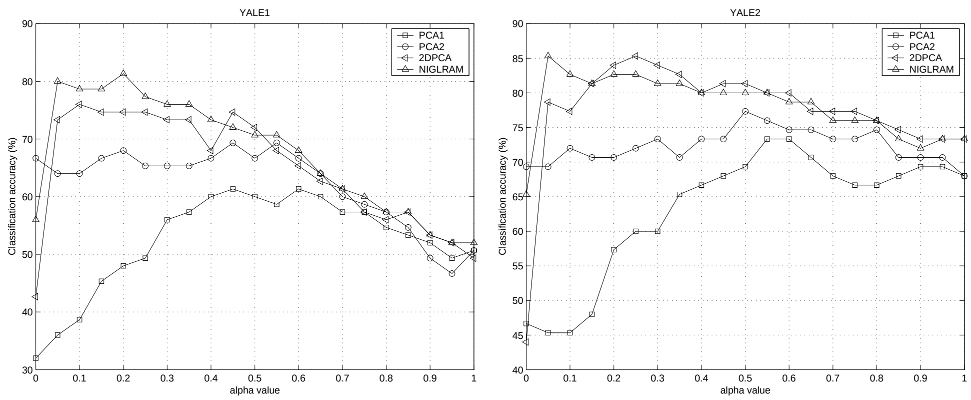


Fig. 12. FSVDR as an IR for PCA-based methods on YALE face database.

significant improvement of classification performance can be typically drawn from the experiments on AR3, where PCA1, PCA2, 2DPCA and NIGLRAM based on FSVDR are respectively 44.7%, 60.6%, 58.3% and 59.6% higher than those based on OGVM.

As a result, from the results presented above, it is easy to conclude that when employing FSVDR as IR, the classification performance of the DR methods

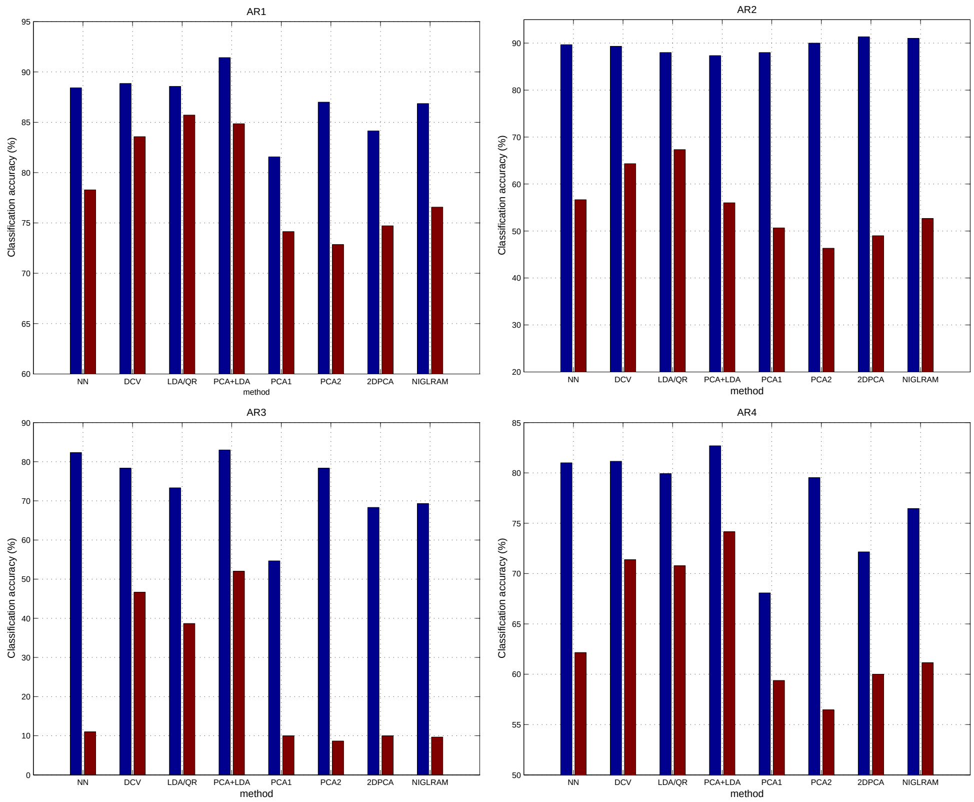


Fig. 13. A comparison of the best classification performance of the proposed FSVDR and OGVM utilizing different methods on AR face database. Left bar: the best classification accuracy by FSVDR, right bar: the classification accuracy by OGVM.

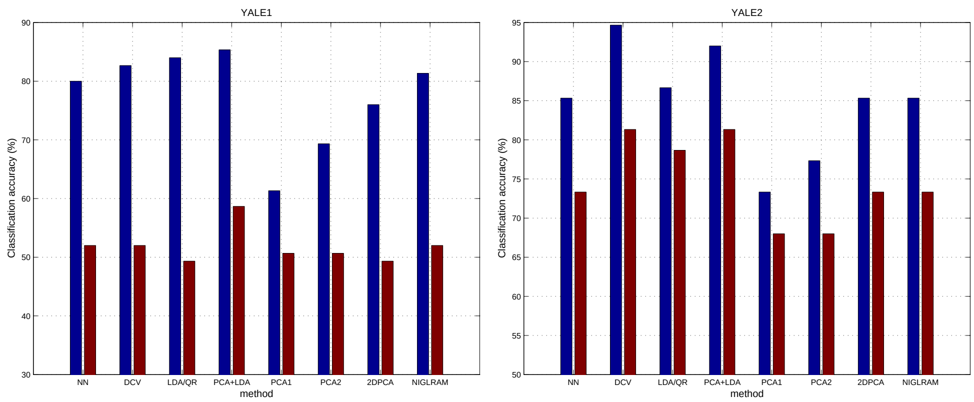


Fig. 14. A comparison of the best classification performance of the proposed FSVDR and OGVM utilizing different methods on YALE face database. Left bar: the best classification accuracy by FSVDR, right bar: the classification accuracy by OGVM.

Table 3
The best classification accuracy (%) of the proposed FSVDR on the six independent experiments and by all the methods. B: FSVDR and A: OGVM

| *Dataset* | | *NN* | *DCV* | *LDA/QR* | *PCA+LDA* | *PCA1* | *PCA2* | *2DPCA* | *NIGLRAM* |
|---|---|---|---|---|---|---|---|---|---|
| AR1 | B | **88.4** | **88.9** | **88.6** | **91.4** | **81.6** | **87.0** | **84.1** | **86.9** |
| | A | 78.3 | 83.6 | 85.7 | 84.9 | 74.1 | 72.9 | 74.7 | 76.6 |
| AR2 | B | **89.7** | **89.3** | **88.0** | **87.3** | **88.0** | **90.0** | **91.3** | **91.0** |
| | A | 56.7 | 64.3 | 67.3 | 56.0 | 50.7 | 46.3 | 49.0 | 52.7 |
| AR3 | B | **82.3** | **78.3** | **73.3** | **83.0** | **54.7** | **78.3** | **68.3** | **69.3** |
| | A | 11.0 | 46.7 | 38.7 | 52.0 | 10.0 | 8.7 | 10.0 | 9.7 |
| AR4 | B | **81.0** | **81.2** | **79.9** | **82.7** | **68.1** | **79.5** | **72.2** | **76.5** |
| | A | 62.2 | 71.4 | 70.8 | 74.2 | 59.4 | 56.5 | 60.0 | 61.2 |
| YALE1 | B | **80.0** | **82.7** | **84.0** | **85.3** | **61.3** | **69.3** | **76.0** | **81.3** |
| | A | 52.0 | 52.0 | 49.3 | 58.7 | 50.7 | 50.7 | 49.3 | 52.0 |
| YALE2 | B | **85.3** | **94.7** | **86.7** | **92.0** | **73.3** | **77.3** | **85.3** | **85.3** |
| | A | 73.3 | 81.3 | 78.7 | 81.3 | 68.0 | 68.0 | 73.3 | 73.3 |

can be significantly improved.

### *3.4 Visualization of samples under the FSVDR*

In this subsection, we will experimentally visualize the distribution of the training and testing samples in the reduced space to show the benefits brought by FSVDR. For this sake, we choose PCA+LDA as feature extraction method which can project the data from $C$ classes to a reduced $C-1$ dimensional space. We carry out two independent experiments: 1) in Fig. 15 (a) and (b), the 30 samples corresponding to 3 randomly chosen classes are from AR2, namely, samples (a-g) (see Fig. 5) are used for training and samples (h-j) are utilized for testing; and 2) in Fig. 15 (c) and (d), the 30 samples corresponding to 3 randomly chosen classes are from AR3, namely, samples (a-g) are used for training and samples (k-m) are utilized for testing.

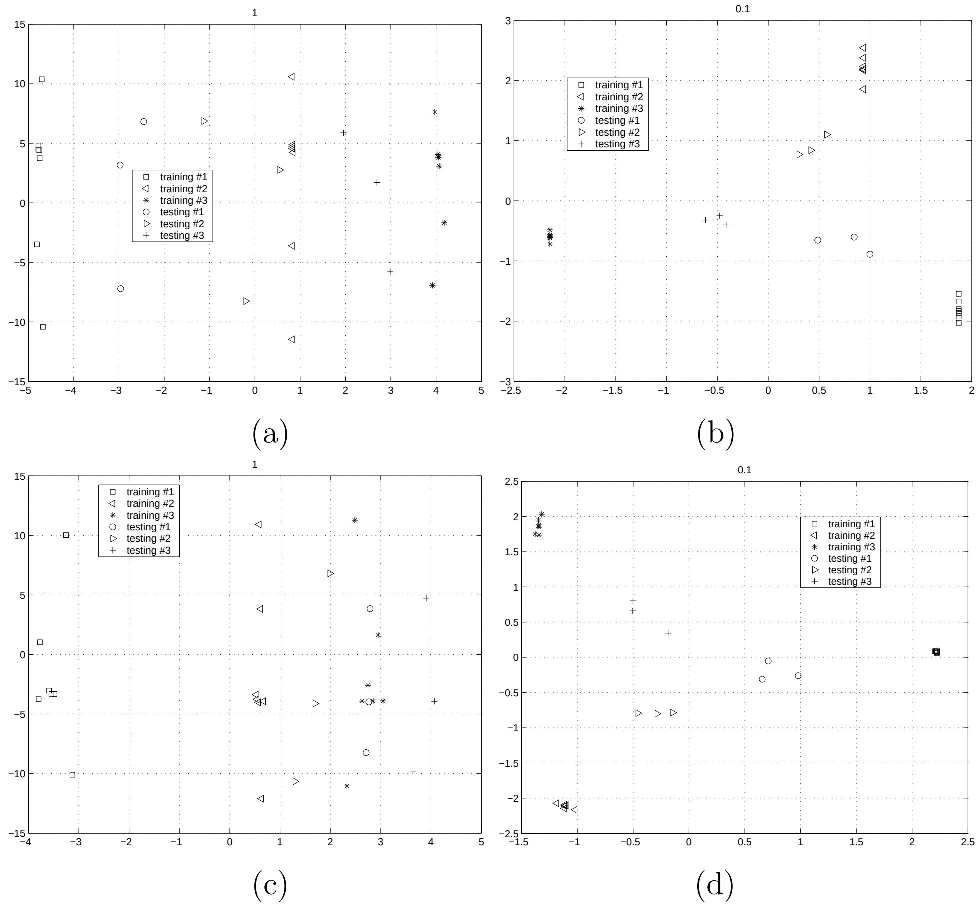


Fig. 15. Visualization of samples from three classes projected by PCA+LDA. In (a) and (c), OGVM is utilized; in (b) and (d), FSVDR is employed, with $\alpha = 0.1$. (a) and (b) use the same training and testing samples, and so for (c) and (d).

Firstly, we look at the training samples presented in Fig. 15 (a), from which we can see that, when employing OGVM (or setting $\alpha$ to 1 in the FSVDR), the samples from the same class are not compact, and the samples from different class are not far away. More specifically, on on hand, the four training samples from # 1 (denoted by square) are around (-4.8,4.4), but the other three training samples from # 1 are far away from (-4.8,4.4) and locate at (-4.7,10.4), (-4.7,-10.4) and (-4.8,-3.5) respectively. Tracing the experimental results, we know that the four training samples from # 1 that locate compactly around (-4.8,4.4) are samples (a-d), while the other three samples are samples (e-g) respectively. Similar observations can also be obtained from the training samples from # 2 and # 3. On the other hand, the *Euclidean* distance between (-4.7,10.4) from # 1 to (0.8, 10.5) from # 2 is about 5.5, quite less than 20.8, the Euclidean distance between (-4.7,10.4) from # 1 and (-4.7,-10.4) from the same class # 1. Obviously, the above given experimental results show that, when performing PCA+LDA on OGVM, we can not come to the objective that the samples from the same class are compact and meanwhile the samples from different classes are far away in case of great facial variations.

Secondly, we look at the testing samples presented in Fig. 15 (a), from which we can clearly see that the testing samples from the same class are not compact too and meanwhile the samples from different classes are not far away. Moreover, the testing samples (of the three classes) corresponding to (h) are around a horizontal line $y = 2.7$, those corresponding to (i) around a horizontal line $y = 6.0$ and those corresponding to (j) around a horizontal line $y = -7.0$. The *Euclidean* distance between (-2.5,6.8) from # 1 and (-1.1,6.8) from # 2 is 1.4, quite less than 14, the *Euclidean* distance between (-2.5,6.8) from # 1 and (-2.9,-7.2) from the same class # 1.

The results presented in Fig. 15 (c) witness similar phenomenon as Fig. 15 (a). The reason behind the phenomena in Fig. 15 (a) and (c) is: when based on OGVM, the PCA+LDA is unable to compactly cluster the same class samples, which are under severe facial variations such as lighting, expression and occlusions.

Finally, we turn to the experimental results with FSVDR as an intermediate representation ($\alpha$ is set to 0.1) in Fig. 15 (b) and (d), from which we can clearly see that: 1) the training samples belonging to the same class become very compact; 2) the training samples belonging to different classes are well separated from each other; 3) the testing samples from the same class are compact with each other; 4) generally speaking, these testing samples do not locate very near to the training with the same classes, which attributes to the fact that our proposed FSVDR does not remove the occlusions such as glasses and scarves, and on the contrary it just alleviates the influence of such occlusions; 5) despite that the testing samples do not locate very near to the training samples in the same classes, the testing samples can all be correctly classified, which attributes to the fact that the testing samples are farther away from the training samples of different classes compared to those of the same class.

Comparing the results presented in Fig. 15 (a) and (c) with those in (b) and (d), we can clearly see that, FSVDR's alleviation of the facial variations can help PCA+LDA achieve its objective (namely, samples from the same class are compact and samples from different classes are far away) and as a result can help improve the classification performance.

### *3.5 The parameter $\alpha$*

In FSVDR, $\alpha$ is a key parameter that should be tuned. From Fig. 7, 8, 9, 10, 11 and 12, we can observe that the classification accuracy curves are approximately unimodal and that there are many $\alpha$'s that can achieve superior performance to OGVM. Further, in our experiments, $\alpha = 0.1$ seems to be a good choice for NN and the PCA-based methods, and $\alpha = 0.4$ seems to a nice choice for the LDA-based methods.

Generally speaking, in designing automatic criterion for choosing adequate parameter $\alpha$, one should consider the following factors: 1) the smaller $\alpha$ is, the more the leading base images (which are sensitive to facial variations) are deflated but meanwhile the discriminant information contained in the leading base images may be deflated too; 2) some face images have great facial variations (e.g., Fig. 5 (h-m, u-z)) and are perhaps in favor of smaller $\alpha$'s, while some face images have slight facial variations (e.g., Fig. 5 (a, n)) and might be in favor of larger $\alpha$'s; 3) the $\alpha$ learned from the training set is a tradeoff among all the training samples and thus is only applicable to the unknown sample from the similar distribution; and 4) each DR method has its specific application scope, which leads to the difficulty in designing a unique $\alpha$ selection criterion for all the DR methods. As a result, the criterion for automatic choosing $\alpha$ should be dependent on the training samples, the given testing sample and the specific DR method.

In the following, instead of aiming at finding an automatic criterion that is applicable to all methods and databases, we try to look for an automatic criterion for the LDA-based methods. Furthermore, in order to ensure that the learned $\alpha$ on the training set is applicable to the testing set, we only consider the case that the testing set has similar distribution as the training set (In fact, this is also almost a basic and common assumption of statistical learning theory, and one can easily get that AR1 and AR4 satisfy this condition, while others not).

Firstly, we introduce a Mean Square Variance (MSV) criterion defined in [27]:

$$MSV = \frac{1}{C}\sum_{i=1}^{C} SV_i, \tag{12}$$

where $SV_i$ is the standard variance of the $i$-th class defined as

$$SV_i = \frac{1}{d}\sum_{k=1}^{d}\sqrt{\frac{1}{N_i - 1}\sum_{j=1}^{N_i}(x_{jk}^i - m_{ik})^2}, \tag{13}$$

where $x_{jk}^i$ and $m_{ik}$ respectively denote the $k$-th element of the $d$-dimensional samples $x_j^i$ and class mean $m_i$, $C$ is the number of classes, and $N_i$ is the number of training samples contained in the $i$-th class. In [27], we argued that when $MSV$ is relatively small, LDA-based methods such as DCV will operate well, and on the contrary, when the MSV value is relatively high, LDA-based methods such as DCV will operate poorly. A justification for this criterion is given as follows: 1) the smaller $MSV$ is, the compact the same class samples are, and on the contrary, the bigger $MSV$ is, the looser the same class samples are; 2) when the same class samples are very loose, these samples will lead to biased estimation of the class mean, within-class and between-class scatter matrices, while on the contrary, when the same class samples are compact, the estimation of the class mean, within-class and between-class variance matrices may be much more reliable; 3) when the same class samples are compact, it is more likely that these samples can nicely depict the Gaussian distribution from which they are generated; and 4) considering the fact that LDA is a special case of the Bayesian decision theory [10] under the assumption that the $C$ classes samples are respectively from $C$ Gaussian distributions with equal covariance, then it is essential for the same class samples to be compact, namely $MSV$ to be small in the LDA-based methods.

Based on the above argument, we give a heuristic criterion to automatically choose an adequate $\alpha$ for the LDA-based methods

$$\alpha_{opt} = arg\min_{\alpha} MSV(\alpha) \tag{14}$$

where $MSV(\alpha)$ is the Mean Square Variance of the training samples $x_j^i = vec(B_j^i), j = 1, 2, \ldots, N_i, i = 1, 2, \ldots, C$, $B_j^i$ is the FSVDR of the face image samples $A_j^i$ under parameter $\alpha$ and $A_j^i$ denotes the $j$-th face image from the $i$-th class.

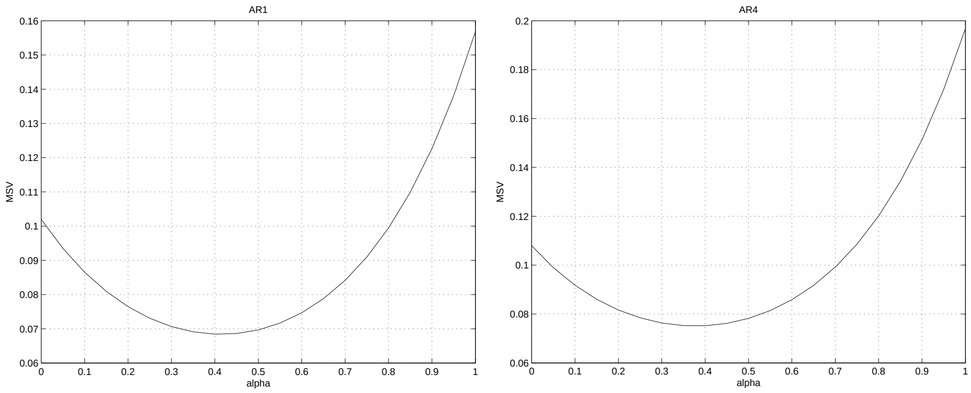


Fig. 16. Values of MSV under different $\alpha$'s on AR1 and AR4.

We report the experimental results on AR1 and AR4 in Fig. 16, from which we can see that, on AR1 and AR4, when $\alpha$ is set to 0.4, $MSV(0.4)$ achieves the optimal value. Turning to the results presented in Fig. 9, we can clearly see that, when $\alpha$ is set to 0.4, the LDA methods (such as DCV, LDA/QR and PCA+LDA) based on FSVDR can yield near optimal classification performance that is significantly higher than those based on the OGVM. To further verify that the learned $\alpha$ from a distribution is applicable to the testing samples from the similar distribution, we carry out the following five experiments by utilizing the same testing samples as AR1 and the first $p$ face images of each person for training, where $p$ changes from 3 to 7 incremented by 1. We employ the learned $\alpha = 0.4$ on AR1 for FSVDR in the five experiments and present the results in Fig. 17, from which we can see that: 1) the learned $\alpha = 0.4$ on AR1 is effective on the five experiments and the LDA methods based on FSVDR can obtain significantly superior performance to those based on OGVM. The underlying reason might be that the testing samples in the five experiments are from the similar distribution as the training samples in AR1 and thus the learned $\alpha = 0.4$ on AR1 is applicable to them; and 2) both FSVDR and OGVM can benefit from more training samples per class.

Finally, the experimental results on YALE1 and YALE2 show that, $MSV(0.6)$ achieves the optimal value in terms of (14). However, from Fig. 10, we can observe that, when $\alpha$ is set to 0.6, although the LDA methods (such as DCV, LDA/QR and PCA+LDA) based on FSVDR can yield better performance than the LDA methods based on OGVM, they are inferior to the optimal ones, which might attribute to the fact that the testing samples are quite different from the training samples, and thus the $\alpha$ learned on the training

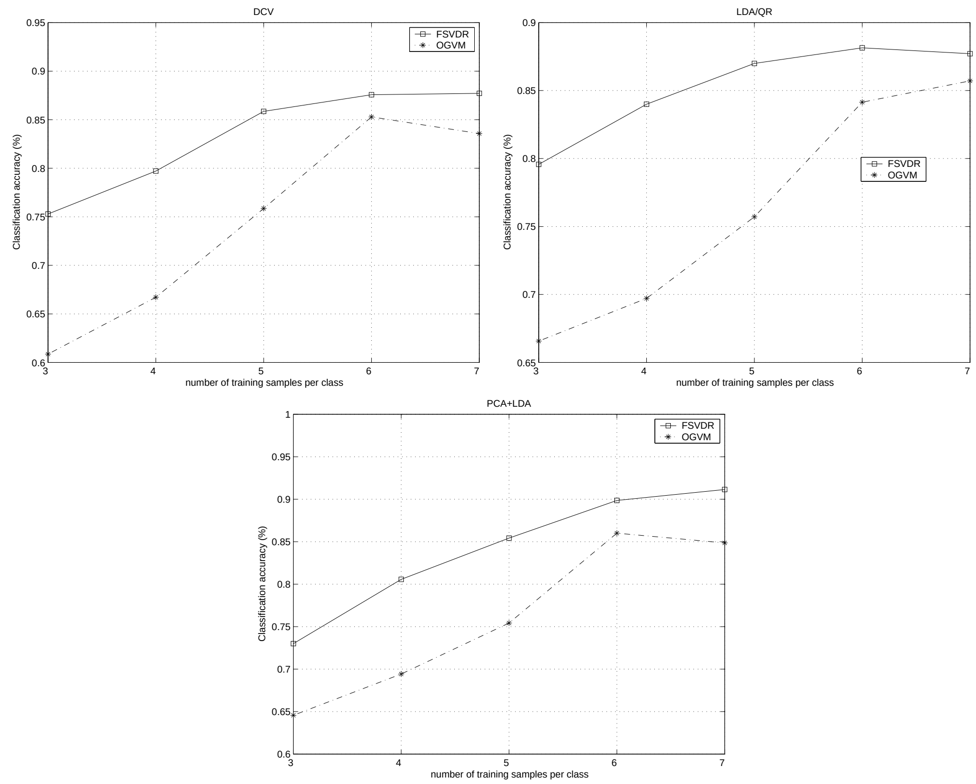


Fig. 17. Performance under different number of training samples per class with $\alpha = 0.4$ on AR1

set is not well-tuned for the testing samples. To deal with this problem, the testing sample should be taken into consideration in designing criterion for $\alpha$, which is worthy of our future study.

## 4 Conclusion

In this paper, we show that the face image matrix $A$ can be viewed as a composition of a set of base images generated by their SVD per se, where the leading base images on one hand dominate the composition of $A$ and on the other hand are sensitive to the great facial variations within the image matrix $A$. Based on these observations, we propose a novel FSVDR $B$, which is a transformed version of OGVM $A$ by SVD and the fractional parameter $\alpha$ and can alleviate facial variations for face recognition. The effectiveness of the proposed FSVDR is verified by extensive experiments conducted on AR, FERET and YALE face databases. More specifically, 1) when directly employing FSVDR for classification, it can yield significantly higher classification accuracies than both OGVM and SVs; and 2) as an intermediate representation, the FSVDR can significantly improve the classification performance of two important categories of DR methods: a) the LDA-based methods such as PCA+LDA, DCV and LDA/QR, and b) the PCA-based methods such as PCA, 2DPCA and NIGLRAM.

In order to show the benefits brought by FSVDR, we select PCA+LDA as a representative, visualize the distribution of both the training and testing samples in the space projected by PCA+LDA to show that, based on the FSVDR, PCA+LDA can cluster the same class samples compact and meanwhile can make the samples from different class far away from each other while PCA+LDA based on OGVM can not. And as a result, FSVDR can improve the classification performance of the LDA-based methods such as PCA+LDA.

Generally speaking, it is both database and method dependent to choose the appropriate value for the parameter $\alpha$ in the FSVDR. Based on our previous work in [27], we give a heuristic criterion in (14) for choosing the appropriate $\alpha$ value for the LDA-based methods such as DCV, LDA/QR and PCA+LDA. And the experimental results show that this criterion can operate well when the testing sample are from the similar distribution as the training samples.

In our viewpoint, there are the following aspects that are worthy of further studies: 1) come up with criteria that can automatically set $\alpha$ adequate values for PCA-based methods such as PCA, 2DPCA and NIGLRAM; 2) modify the MSV criterion to take the testing sample into consideration so that it can be applicable to the case that the testing sample differs from the training samples to a large extent (e.g., AR2, AR3, FERET, YALE1 and YALE2); 3) carry out research to set sample dependent parameter $\alpha$, since different face samples are affected by facial variations differently (note that, in our FSVDR, we set a universal value for all the face samples); 4) set different $\alpha$ values for different singular values in order to better suppress noise and meanwhile retain discriminant information; 5) compare FSVDR with intermediate representations such as Gabor wavelet and DCT for face recognition; and 6) combine FSVDR and Gabor wavelet or DCT to form a new intermediate representation for face recognition.

## Acknowledgment

We would like to thank the anonymous reviewers for their constructive comments that enhance this paper, Martinetz for providing the AR database, the FERET program for the FERET database, and Belhumeur et al. for the YALE database. This work is supported by the National Natural Science Foundation of China under Grant No. 60473035, the Natural Science Foundation of Jiangsu Province under Grant No. BK2006187, and the Innovation Foundation of NUAA under Grand No. Y0603-042.

## References

[1] P. Belhumeur, J. Hespanha, and D.J. Kriegman. Eigenfaces vs. fisherfaces: Recognition using class specific linear projection. *IEEE Trans. Pattern Anal. Mach. Intelligence*, 19(7):711–720, 1997.

[2] D. Bolme, J. Beveridge, M. Teixeira, and B. Draper. The csu face identification evaluation system: Its purpose, features and structure. In *Third International Conference on Computer Vision Systems*, pages 303–311, 2003.

[3] V. Bruce, Hancock, and A.M. Burton. *Face Recognition: From Theory to Applications*. Springer-Verlag, 1998.

[4] R. Brunell and T. Poggio. Face recognition: Features vs templates. *IEEE Trans. Pattern Anal. Mach. Intelligence*, 15(10):1042–1053, 1993.

[5] H. Cevikalp, M. Neamtu, Wilke, and A. Barkana. Discriminative common vectors for face recognition. *IEEE Trans. Pattern Anal. Mach. Intelligence*, 27(1):4–13, 2005.

[6] R. Chellappa, C.L. Wilson, and S. Sirohey. Human and machine recognition of faces: A survey. *Proceedings of the IEEE*, 83(5):705–740, 1995.

[7] Y. Cheng, K. Liu, J. Yang, Y. Zhuang, and N. Gu. Human face recognition method based on the statistical model of small sample size. In *Intelligent Robots and Computer Vision X: Algorithms and Techniques*, pages 85–95, 1991.

[8] I. Cox and P. Yianilos. Feature-based face recognition using mixture distance. In *IEEE Conf. on Comp. Vision and Patt. Recog.*, pages 209–216, 1996.

[9] G. Donato, M. S. Bartlett, Hager J.C., P. Ekman, and T. J. Sejnowski. Classifying facial actions. *IEEE Trans. Pattern Anal. Mach. Intelligence*, 21:974–989, 1999.

[10] R.O. Duda and Stork D. G. Hart, P. E. *Pattern Classification.* Wiley-Interscience, second edition, 2000.

[11] S. Edelman. *Representation and Recognition in Vision, A Bradford Book.* MIT Press, Cambridge, MA, 1999.

[12] M.J. Er, W. Chen, and S. Wu. High-speed face recognition based on discrete cosine transform and rbf neural networks. *IEEE Trans. on Neural Networks*, 16(3):679–691, 2005.

[13] R.A. Fisher. The use of multiple measures in taxonomic problems. *Ann. Eugenics*, 7:179–188, 1936.

[14] A. Goldstein, L. Harmon, and A. Lesk. Identification of human faces. *Proceedings of the IEEE*, 59(5):748–760, 1971.

[15] A. Goldstein and P. Iyengar. Automatic recognition and analysis of human faces and facial expressions: A survey. *Pattern Recognition*, 25:65–77, 1992.

[16] G.H. Golub and C.F. Van Loan. *Matrix Computations.* The Johns Hopkins University Press, 1995.

[17] L. Harmon, M. Kaun, R. Lasch, and P. Ramig. Machine identification of human faces. *Pattern Recognition*, 13(2):97–110, 1981.

[18] L. Harmon, S. Kuo, P. Ramig, and U. Raudkivi. Identification of human face profiles by computer. *Pattern Recognition*, (10):301–312, 1978.

[19] Z. Hong. Algebraic feature extraction of image for recognition. *Pattern Recognition*, 24:211–219, 1991.

[20] M. Kirby and L. Sirovich. Application of the karhunen-loeve procedure for the characterization of human faces. *IEEE Trans. Pattern Anal. Mach. Intelligence*, 12:103–108, 1990.

[21] S.G. Kong, J. Heo, B.R. Abidi, J. Paik, and M. A. Abidi. Recent advances in visual and infrared face recognition-a review. *Computer Vision and Image Understanding*, 97(1):103–135, 2005.

[22] A. Leonardis and H. Bischof. Kernel and subspace for computer vision. *Pattern Recognition*, 36(9):1925, 2003.

[23] C. Liu and H. Wechsler. Robust coding schemes for indexing and retrieval from large facedatabases. *IEEE Trans. Image Processing*, 9:132–137, 2000.

[24] C. Liu and H. Wechsler. Gabor feature based classification using the enhanced fisher linear discriminant model for face recognition. *IEEE Trans. Image Processing*, 11(4):467–476, 2002.

[25] C. Liu and H. Wechsler. Independent component analysis of gabor features for face recognition. *IEEE Trans. Neural Networks*, 14(4):919–928, 2003.

[26] J. Liu and S. Chen. Resampling lda/qr and pca+lda for face recognition. In *The 18th Australian Joint Conference on Artificial Intelligence*, pages 1221–1224, 2005.

[27] J. Liu and S. Chen. Discriminant common vecotors versus neighbourhood components analysis and laplacianfaces: A comparative study in small sample size problem. *Image and Vision Computing*, 24(3):249–262, 2006.

[28] J. Liu and S. Chen. Non-iterative generalized low rank approximation of matrices. *Pattern Recognition Letters*, 27(9):1002–1008, 2006.

[29] A. M. Martinez. Recognizing imprecisely localized, partially occluded, and expression variant faces from a single sample per class. *IEEE Trans. Pattern Anal. Mach. Intelligence*, 25(6):748–763, 2002.

[30] A. M. Martinez and R. Benavente. The ar face database. Technical report, CVC, 1998.

[31] A. M. Martinez and A.C. Kak. Pca versus lda. *IEEE Trans. Pattern Anal. Mach. Intelligence*, 23(2):228–233, 2001.

[32] A. M. Martinez and M. Zhu. Where are linear feature extraction methods applicable. *IEEE Trans. Pattern Anal. Mach. Intelligence*, 27(12):1934–1944, 2005.

[33] Phillips, H. Wechsler, J. Huang, and P. Rauss. The feret database and evaluation procedure for face recognition algorithms. *Image and Vision Computing*, 16(5):295–306, 1998.

[34] D. Swets and J. Weng. Using discriminant eigenfeatures for image retrieval. *IEEE Trans. Pattern Anal. Mach. Intelligence*, 18:831–836, 1996.

[35] X. Tan, S. Chen, Z. Zhou, and F. Zhang. Recognizing partially occluded, expression variant faces from single training image per person with som and soft knn ensemble. *IEEE Trans. on Neural Networks*, 16(4):875–886, 2005.

[36] M. Tarr and H.H. Bulthoff. *Object Recognition in Object Recognition in Man, Monkey, and Machine.* MIT Press, Cambridge, MA, 1999.

[37] Y. Tian, T. Tan, Y. Wang, and Y. Fang. Do singular values contain adequate information for face recognition. *Pattern Recognition*, 36(3):649–655, 2003.

[38] M. Turk and A. Pentland. Eigenfaces for recognition. *J. of Cognitive Neuroscience*, 3(1):71–96, 1991.

[39] L. Wiskott, J. Fellous, N. Kruger, and C. von der Malsburg. Face recognition by elastic bunch graph matching. *IEEE Trans. Pattern Anal. Mach. Intelligence*, 19(7):775–779, 1997.

[40] J. Yang, D. Zhang, and et al. Two-dimensional pca: A new approach to appearance-based face representation and recognition. *IEEE Trans. Pattern Anal. Mach. Intelligence*, 26(1):131–137, 2004.

[41] J. Ye. Generalized low rank approximation of matrices. In *International Conference on Machine Learning*, pages 887–894, 2004.

[42] J. Ye and Q. Li. A two-stage linear discriminant analysis via qr-decomposition. *IEEE Trans. Pattern Anal. Mach. Intelligence*, 27(6):929–941, 2005.

[43] D. Zhang, S. Chen, and Z. Zhou. A new face recognition method based on svd perturbation for single example image per person. *APPLIED MATHEMATICS AND COMPUTATION*, 163(2):895–907, 2005.

[44] W. Zhao, R. Chellappa, and Krishnaswamy. A. Discriminant analysis of principal components for face recognition. In *International Conference on Automatic Face and Gesture Recognition*, pages 336–341, 1998.

[45] W. Zhao, R. Chellappa, Phillips, and A. Rosenfeld. Face recognition: A literature survey. *ACM Computing Survey*, 35(4):399–458, 2003.

[46] W. Zhao, R. Chellappa, and P.J. Phillips. Subspace linear discriminant analysis for face recognition. Technical report, Center for Automation Research, University of Maryland, 1999.