Variational Feature Representation-based Classification for Face Recognition with Single Sample Per Person

Ru-Xi Ding[a,b], Daniel K. Du[a,c], Zheng-Hai Huang[a,b], Zhi-Ming Li[a,b,*], Kun Shang[a]

[a] *Center for Applied Mathematics of Tianjin University, Tianjin 300072, P.R. China*
[b] *Department of Mathematics, School of Science, Tianjin University, Tianjin 300072, P.R. China*
[c] *School of Electronic Information Engineering, Tianjin University, Tianjin 300072, P.R. China*

Abstract

The single sample per person (SSPP) problem is of great importance for real-world face recognition systems. In SSPP scenario, there is always a large gap between a normal sample enrolled in the gallery set and the non-ideal probe sample. It is a crucial step for face recognition with SSPP to bridge the gap between the ideal and non-ideal samples. For this purpose, we propose a Variational Feature Representation-based Classification (VFRC) method, which employs the linear regression model to fit the variational information of a non-ideal probe sample with respect to an ideal gallery sample. Thus, a corresponding normal feature, which reserve the identity information of the probe sample, is obtained. A combination of the normal feature and the probe sample is used, which makes VFRC method more robust and effective for SSPP scenario. The experimental results show that VFRC method possesses higher recognition rate than other related face recognition methods.

Keywords: Face recognition, Single sample per person, Generic learning, Variational feature representation

*Corresponding author. Tel.: 86-22-27403615, fax: 86-22-27403615.
Email addresses: dingruxi@tju.edu.cn (Ru-Xi Ding), daniel@tju.edu.cn (Daniel K. Du), huangzhenghai@tju.edu.cn (Zheng-Hai Huang), lizm@tju.edu.cn (Zhi-Ming Li), skun@tju.edu.cn (Kun Shang)

1. Introduction

In modern society, face recognition (FR) technology has become more and more popular in many application fields [1, 2], such as information security, law enforcement and surveillance, smart cards, access control, and so on. With the increased attention from researchers, many methods have been proposed in the literature [3, 4, 5, 6, 7, 8, 9, 10]. However, there are still many challenges need to be faced in FR field. One of the challenges is single sample per person (SSPP) problem [11, 12], especially for the situation that the probe samples contain large appearance variations caused by illumination, expression, age, pose, and so on. For some FR applications, such as law enhancement, e-passport and ID card, there is usually only a single sample per person recorded in the systems. The main reasons lie in two aspects: on the one hand, it is difficult for those FR systems to collect additional samples under many scenarios; and on the other hand, it is need to be considered to reduce the cost of storage space. As the probe samples may contain various varying information, numerous appearance-based multi-sample methods (e.g., Eigenfaces [13], Fisherfaces [14], Local Binary Pattern [15]) may be noneffective or may fail to work for SSPP issue. Since SSPP problem with non-ideal conditions is a big challenge in FR field, many multi-sample methods have been developed to address SSPP problem. At the same time, a variety of novel specific methods have been created by many researchers [11]. The methods for SSPP problem proposed in the literature can be classified the following three categories.

The first category of methods is based on virtual sample generation. In order to make the discriminative subspace learning methods adjustable for extracting feature with SSPP problem, some additional training samples, which are virtually generated, are added into each class of the gallery set. In [16, 17], two singular value decomposition (SVD)-based perturbation algorithms were proposed to make the conventional Linear Discriminant Analysis (LDA) [18] available for SSPP issue by generating multiple images for each person. However, the distinguish information of the virtual samples in each class was not increased with

the increasing number of virtual samples. Such an evident shortcoming comes from high correlation among the virtual samples and the corresponding original sample.

The second category of methods is based on image partitioning, as Modular Principal Component Analysis (Block-PCA) [19], Block based Linear Discriminant Analysis (Block-LDA) [20], Discriminative Multimanifold Analysis (DMMA) [21] and Multi-feature Multi-Manifold Learning (M^3L)[22]. In Block-PCA and Block-LDA, patch skill is used to divided face images into small local patches, on which the discriminant learning techniques are applied. In DMMA [21], Lu et. al. used the manifold theory to divided one person manifold to multi-manifold and compared the similarities between the local patch-manifold per person. Yan et. al. proposed M^3L by extending DMMA with respect to the feature representation. In M^3L, not only the raw intensity feature, but also LBP and Gabor features were extracted within each small patch. Nevertheless, the form and size of the patch have big influences on the recognition results.

The last category of methods is based on generic learning [23, 24, 25]. The discriminative features, which are learned from an additional generic training set with multi-samples per person, are used to recognize the probe person. The methods with generic learning are based on the assumption that the generic training set and SSPP gallery set share similar variation information of both inter-class and intra-class. Su et al. [25] proposed an adaptive generic learning (AGL) method to successfully apply LDA to solve SSPP problem. Instead of directly employing the discriminatory information (e.g., the mean and covariance of each class) learned from the generic set, AGL method adapts it to predict intra-personal variations and mean for each subject enrolled in the gallery through least square regression. Then, the next step is to estimate the total intra-class and inter-class scatter matrix of all subjects in the gallery set. The classical sparse representation based classification (SRC) [26, 27, 28] can represent the probe samples well, when there are enough training samples per person. In addition, the representation ability of SRC is much more limited when it faces SSPP problem. Deng et. al. [29, 30] presented an effective

method, named Extended SRC (ESRC), to overcome the disadvantage of SRC for SSPP. ESRC applies an auxiliary generic variant dictionary to represent the possible intra-class variation between the gallery and probe samples. Though ESRC performs well for SSPP problem, both the extraction of variational feature and the computational time are wildly discussed by researchers. Recently, an effective method, called Sparse Variation Dictionary Learning (SVDL) [31] method, was proposed for FR with SSPP. Instead of learning from the auxiliary generic set independently, SVDL method obtains a projection by learning from both auxiliary generic and gallery sets. The learned sparse variation dictionary plays an important role in handling all kinds of variations in face samples, including illumination, expression, occlusion, and pose. SVDL requires a sufficiently large generic training set, in which each subject should contain multiple face images with all type of variations.

In this paper, we propose a new generic learning method, named Variational Feature Representation-based Classification (briefly noted by VFRC), to solve SSPP problem with various non-ideal conditions (e.g., illumination, expression, occlusion, pose, age and comprehensive situations). Unlike the conventional generic learning methods, such as AGL, ESRC, and SVDL, the proposed VFRC does not directly employ the discriminatory information learned from the generic set. Alternatively, in the proposed VFRC model, the variational feature of probe sample is represented by the joint information of the generic set and the gallery set. As the gallery set can be regarded as a basic reference substance in the representation for variational feature, the rest normal feature of probe sample can be obtained more precisely. For the reason that the normal feature keeps less variational information, it is beneficial to enhance the identity information of the corresponding probe sample and makes a great contribution to the accuracy and robustness of VFRC with complicated, broad changing variations in SSPP scenario.

In order to well verify the effectiveness of the proposed VFRC, two popular image-to-image and image-to-set experiments are implemented. For image-to-image experiments, we compare the proposed VFRC with seven related methods

(SRC [26], CRC [32], Block-LDA [20], DMMA [21], AGL [25], ESRC [29], and SVDL [31]) on three public face databases (AR [33], Extended Yale B [34, 35], CMU-PIE [36] databases) with various variations, including illumination, expression, pose, and occlusion. For the image-to-set experiments, we choose one of the most challenging face database LFW [37] to verify the effectiveness of our method. The seven methods above and Locality Repulsion Projections and Sparse Reconstruction based Similarity Measure (LRP-SRSM) [38] are chose to compare with VFRC for image-to-set experiments. Experimental results demonstrate that the proposed VFRC achieves state-of-the-art performance for FR under the SSPP scenario with multiple non-ideal conditions.

The remainder of this paper is organized as follows. Section 2 presents our method for SSPP problem. In Section 3, the experimental results on three face databases are presented. The final section is the conclusion of this paper.

2. Proposed Approach

It is well known that FR with SSPP problem has two natural characters: the first is that the images enrolled in the gallery set are usually frontal pose with ideal conditions (such as natural expression and illumination without occlusion), and the second is that the images from the probe set can possess kinds of non-ideal conditions (such as varying expression, illumination, pose, occlusion, and so on). Hence, how to build a bridge between the normal gallery image and the non-ideal probe image is a key point for improving the performance of FR methods for solving SSPP problem. In the following subsections, we provide a feasible scheme for this problem.

2.1. Normal Feature and Variational Feature of a Face Image

The face images we get from real world may contain many variational information, such as illumination, expression, occlusion, pose, and so on. A face image can be represented by the normal face and the variational information, which is written as the following form:

$$I = N + V,$$

where $I \in \mathbb{R}^d$ $(d = m \cdot n)$ is the vectorization of face image with size $m \times n$; $N \in \mathbb{R}^d$ denotes the normal face feature; and $V \in \mathbb{R}^d$ denotes the variational feature. Thus, we divide a probe sample y into two subparts: the normal feature y_n and the variational feature y_v, formally denoted as

$$y = y_n + y_v, \tag{1}$$

where y_v contains some interference factors, which may mislead the identity recognition of the non-ideal probe sample; and y_n possesses the inherent identity information of the probe sample.

In SSPP scenario, $A = [x_1, x_2, \cdots, x_C]$ denotes the gallery set, which contains C different individuals with only one ideal sample per subject, $x_i \in \mathbb{R}^d$ is the vectorization of the i-th enrolled subject. As the most interference factors consist in the variational feature y_v, the normal feature y_n is closer to the corresponding ideal sample in the gallery set than the non-ideal probe sample y. And there is a strong correlation between two samples from the same subject, the normal feature y_n of the probe sample y could be represented as

$$y_n = \alpha_i x_i, \tag{2}$$

where x_i (the i-th sample in the gallery set) possesses the same identity as the probe sample y, and α_i is the representation coefficient of y_n over x_i. Because the identity of the probe sample is initially unknown, the representation of y_n can be rewritten as

$$y_n = A\alpha$$

in terms of all gallery samples, where $\alpha \in \mathbb{R}^C$ (C is the number of classes in the gallery set A) is coefficient vector.

Under the same variational conditions, different individuals share the similar variational feature, which makes the variational feature universal. However, the identity information existed in the normal feature is individual. Thus, it is not easy to get the individual normal feature directly. As showed in Eq. (1), if we get the variational feature y_v, the normal feature can be got indirectly.

2.2. Representation of Variational Feature

To obtain the variational feature, a generic set $G = \{G_1, G_2, \cdots, G_K\}$ is constructed, in which K is the number of individuals in the generic set, $G_k = [g_{k0}, g_{k1}, \cdots, g_{kn_k}] \in \mathbb{R}^{d\times(n_k+1)}$ $(k = 1, 2, \cdots, K)$ stacks the available samples of the k-th class, g_{k0} is the ideal normal sample, and g_{kl} $(l = 1, \cdots, n_k)$ is the sample with various variations including illumination, expression, pose and occlusion. The variational feature in each class can be obtained by subtracting the normal sample from other samples of the same class, denoted by

$$V_k = \bar{G}_k - g_{k0} * e_k^\top, \tag{3}$$

where $\bar{G}_k = [g_{k1}, \cdots, g_{kn_k}] \in \mathbb{R}^{d\times n_k}$, $e_k = [1, 1, \cdots, 1]^\top \in \mathbb{R}^{n_k}$. So we can get the variation matrix V via the generic set:

$$V = [V_1\ V_2\ \cdots\ V_K]. \tag{4}$$

Given a probe sample y, its variational feature y_v is obtained by the linear regression model:

$$\begin{bmatrix} \hat{\alpha} \\ \hat{\beta} \end{bmatrix} = \arg\min_{\alpha,\beta} \left\| y - [A\ V] \begin{bmatrix} \alpha \\ \beta \end{bmatrix} \right\|_2^2 + \lambda_1 \left\| \begin{bmatrix} \alpha \\ \beta \end{bmatrix} \right\|_2^2, \tag{5}$$

where $\left[\alpha^\top\ \beta^\top\right]^\top$ is the representation coefficient, and λ_1 is the penalty parameter.

Obviously, the objective function of the above model is convex and differentiable. Let $G\left(\left[\alpha^\top\ \beta^\top\right]^\top\right)$ denote the gradient of the objective function, i.e.,

$$\begin{aligned} G\left(\begin{bmatrix} \alpha \\ \beta \end{bmatrix}\right) &= 2\begin{bmatrix} A^\top \\ V^\top \end{bmatrix}\left\{[A\ V]\begin{bmatrix} \alpha \\ \beta \end{bmatrix} - y\right\} + 2\lambda_1 \begin{bmatrix} \alpha \\ \beta \end{bmatrix} \\ &= 2\left\{\left(\begin{bmatrix} A^\top \\ V^\top \end{bmatrix}[A\ V] + \lambda_1 I\right)\begin{bmatrix} \alpha \\ \beta \end{bmatrix} - \begin{bmatrix} A^\top \\ V^\top \end{bmatrix} y\right\}. \end{aligned}$$

By the first-order optimality criterion, $G\left(\left[\alpha^\top\ \beta^\top\right]^\top\right) = 0$, we get

$$\left(\begin{bmatrix} A^\top \\ V^\top \end{bmatrix}[A\ V] + \lambda_1 I\right)\begin{bmatrix} \alpha \\ \beta \end{bmatrix} - \begin{bmatrix} A^\top \\ V^\top \end{bmatrix} y = 0 \tag{6}$$

The solution of Eq. (6) is the optimal solution of (5):

$$\left[\hat{\alpha}^\top \ \hat{\beta}^\top\right]^\top = \left\{[A\ V]^\top [A\ V] + \lambda_1 I\right\}^{-1} [A\ V]^\top y.$$

Then, we can get the variational feature of the probe sample, denoted by

$$y_v = V\hat{\beta}, \tag{7}$$

where the coefficient $\hat{\beta}$ of variational feature is

$$\begin{aligned} \hat{\beta} \ = \ & (F - V^\top A E^{-1} A^\top V)^{-1} V^\top y \\ & -F^{-1} V^\top A (E - A^\top V F^{-1} V^\top A)^{-1} A^\top y, \end{aligned} \tag{8}$$

here $E = A^\top A + \lambda_1 I$ and $F = V^\top V + \lambda_1 I$.

Different from those existing popular generic learning methods [25, 29, 31], as showed in Eq. (8), the variational feature y_v is got via the gallery set A and the variation matrix V. In other words, the variational feature is represented by the variation matrix with the reference of the gallery set. Furthermore, the variational feature describes the difference information between the ideal samples in the gallery set and the non-ideal probe sample, which helps to get more precise normal feature y_n. Fig. 1 gives a visual example, in which two groups of variational information and corresponding normal face are displayed. One can see that the variational feature is well represented, and the corresponding normal feature is well reconstructed, simultaneously.

2.3. Variational Feature Representation-based Classification Model

The normal feature y_n can be obtained from the analysis in previous subsection, which is given as follows:

$$y_n = y - y_v. \tag{9}$$

As mentioned above, y_n is regarded as a normal feature of the non-ideal probe sample y. Since the normal feature reserves few interference factors, it is beneficial to enhance the identity information of the corresponding probe

sample. Thus, we use a combination of the probe sample y and its normal feature y_n in our model of VFRC, which is given by

$$\hat{\eta} = \arg\min_{\eta} \|(1-\gamma)y + \gamma y_n - A\eta\|_2^2 + \lambda_2\|\eta\|_2^2, \tag{10}$$

where y is the probe sample; y_n is the normal feature of y; $\gamma \in (0,1)$ is the weighted factor between y and y_n; and λ_2 is the penalty parameter. It is easy to see that the objective function in (10) is convex with respect to η. So, we can further get the global optimal solution of (10):

$$\hat{\eta} = \left(A^\top A + \lambda_2 I\right)^{-1}[(1-\gamma)y + \gamma y_n].$$

Now, the classification of the probe sample y is conducted via

$$\text{identity}(y) = \arg\min_i \{e_i\},$$

where $e_i = \|y - \gamma y_v - x_i\hat{\eta}_i\|_2^2 / \|\hat{\eta}\|_2^2$.

In the combination used in (10), the probe sample part makes the proposed VFRC utilize original image information and the normal feature part helps to avoid the influence of the interference information. As mentioned in the previous subsection, the normal feature part possesses the identity information of the corresponding probe sample, which is of great importance in the recognition process.

2.4. Algorithm for VFRC

The complete recognition procedure of our proposed VFRC is summarized in following Algorithm 1.

3. Experimental Results

In this section, we focus on three standard face databases (AR database [33], Extended Yale B database [34, 35] and CMU-PIE database [36]) and one challenging database (LFW database [37]) to demonstrate the performance of VFRC for image-to-image and image-to-set SSPP problem with complicated variational situations, respectively. We compare VFRC with seven popular methods:

Algorithm 1 VFRC method

Input: A gallery set $A = [x_1, x_2, \cdots, x_C] \in \mathbb{R}^{d\times C}$, a probe sample $y \in \mathbb{R}^d$, a generic set $G = \{G_1, G_2, \cdots, G_K\}$, a weighted factor γ, two penalty parameters λ_1 and λ_2.

Compute:

1. The variation matrix $V = [V_1\ V_2\ \cdots\ V_K]$, where V_k is got by Eq. (3);
2. the variational feature $y_v = V\hat{\beta}$ of the probe sample y, where $\hat{\beta}$ is obtained by solving the optimal problem (5);
3. solving the optimal problem (10);
4. the residuals $e_i = \|y - \gamma y_v - x_i\hat{\eta}_i\|_2^2/\|\hat{\eta}\|_2^2$, $i = 1, 2, \cdots, C$.

Output: The identity of the probe sample y: identity$(y) = \arg\min_i\{e_i\}$.

SRC [26], CRC [32], Block-LDA [20], DMMA [21], AGL [25], ESRC [29], and SVDL [31] for image-to-image experiments. We also compare the LRP-SRSM [38] and these seven methods with VFRC for the image-to-set experiments. All experiments are finished in Matlab 64bit on a Lenovo Think Center M9201z platform with four Intel Core i5-3550S 2.9GHz CPUs and 8GB of RAM.

In the following experiments, the l_1-regularized minimization in SRC, ESRC and LRP-SRSM is solved by $l_1\_l_s$ method [39]; all the methods have been implemented without dimensionality reduction except for Block-LDA (the dimension after reduction is the number of the samples in the gallery set), AGL (according to [25], the energy is reserved with 96%), ESRC (according to [29], the dimension after reduction is 90), SVDL (according to [31], the dimension after reduction is 90) and LRP-SRSM (according to [38], the energy is reserved with 99.9%); the experiments of parameters setting are carried on the four databases for VFRC and other compared methods, except for SRC (due to the expensive computing cost), AGL and SVDL (all the parameters are fixed according to [25] and [31], respectively); the parameter λ is set 0.001 in SRC throughout all experiments.

3.1. Parameters Setting

We make experiments for parameters setting with CRC, Block-LDA, DMMA, ESRC and the proposed method VFRC on each database. Cross validation is used to find the available combination schemes of relevant parameters for each method. In order to present the complete experimental process of parameters setting, VFRC is implemented on ARs (which is a subset of AR database, more details are showed in Subsection 3.2). We choose the first sample of each subject for training and the rest 25 samples of each subject for testing. The recognition rates are recorded.

- We fix the parameters $\lambda_1 = 0.001$ and $\lambda_2 = 0.001$ in VFRC to check influence of γ on the recognition rate. As showed in Fig. 2(a), when $\gamma = 0.8$, VFRC method performs best.

- According to Fig. 2(a), we fix the parameters $\gamma = 0.8$ and $\lambda_2 = 0.001$ in VFRC to check the influence of λ_1. As showed in Fig. 2(b), when $\lambda_1 = 0.01$, VFRC performs better than others. Meanwhile, the variation of the recognition rate is limited, which means the variation of parameter λ_1 shares a little effect on the recognition rate.

- Then, we fix $\gamma = 0.8$ and $\lambda_1 = 0.01$ in VFRC. The recognition rates are showed in Fig. 2(c) with varying value of λ_2. When $\lambda_2 = 0.1$, the recognition rate is the highest.

- As the variations of γ and λ_2 bring big influence on the recognition results, we check again the effect of the parameter γ with fixed λ_1 and λ_2, the parameter λ_2 with fixed γ and λ_1, the parameter λ_1 with fixed γ and λ_2, one by one. As showed in Fig. 2(d), VFRC achieves the best recognition rate with $\gamma = 0.9$, when $\lambda_1 = 0.01$ and $\lambda_2 = 0.1$. Fig. 2(e) shows the recognition rate of VFRC with $\gamma = 0.9$, $\lambda_1 = 0.01$ and varying λ_2. When $\lambda_2 = 0.08$, VFRC performs better than others. After that, we set $\gamma = 0.9$ and $\lambda_2 = 0.008$, the recognition rate with varying λ_1 is showed in Fig. 2(f).

By means of those experiments above, we get the optimal value of parameters used in VFRC as: $\gamma = 0.9$, $\lambda_1 = 0.02$ and $\lambda_2 = 0.08$ on ARs database. The similar experiments have been done for other compared methods on ARs database. Depend on the results of these experiments, the parameters are set respectively as: $\lambda = 0.03$ in CRC; $\lambda = 0.01$ and $ps = 6$ in Block-LDA; $ps_{row} = 25$ and $ps_{column} = 20$ in DMMA; $\lambda = 0.02$ in ESRC. The parameter ps represents the patch size and λ is the penalty parameter in each model. On other three databases, the similar experiments are also carried out before the main experiments one by one.

3.2. Experiments with Complicated Varying Conditions

In this subsection, we aim at inspecting the effectiveness of VFRC for SSP-P problem with the complicated variational probe samples. Since AR face database contains 3,276 color images with difference illumination (all side lights on, left light on and right light on), varying facial expression (neutral, smile, anger and scream) and disguise (sunglasses and scarf) from 70 males and 56 females, we use it in this subsection. Each subject on AR database includes 26 images, as showed in Fig. 3. A subset (signed as ARs), which contains 50 male subjects and 50 female subjects with 26 images per subject, is selected for the experiments. On ARs, the first sample of each subject is ideal image, which is collected in gallery set. We also select 5 males and 5 females from the rest subjects with 26 samples to get the generic set. All images are cropped from original 768×576 to 50×40 gray scale.

Now we implement VFRC and other seven methods on ARs database. We choose the first sample of each subject for training; and choose randomly 5, 10, 15, 20, 24 samples, respectively, from the rest of each subject for testing. The experiments are repeated 10 times for each number of probe samples, and the average recognition rates and the corresponding standard deviation are listed in Table 1 and Table 2, respectively. As showed in Table 1 and Table 2, the numbers of the first row represent the number of probe samples chose from each samples (the same rule has been used in the following all tables).

Table 1: Average recognition rates (%) on ARs database

Method	5	10	15	20	24
SRC	65.94	67.08	65.29	71.20	69.95
CRC	69.22	69.96	68.43	73.74	72.52
Block–LDA	56.00	54.44	55.09	55.51	55.98
DMMA	66.52	66.49	65.27	65.01	64.15
AGL	70.12	69.50	67.80	68.32	68.88
ESRC	72.76	74.57	72.57	73.74	73.70
SVDL	72.94	76.15	76.06	74.82	74.81
VFRC	**76.52**	**76.94**	**79.49**	**79.87**	**78.65**

From Table 1, it is easy to see that the recognition rates of VFRC are higher than other related methods on ARs database. In addition, it is mentioned in Table 2 that VFRC is more stable than other compared methods. In a word, the proposed VFRC is effective for SSPP problem with complicated varying conditions, such as varying illumination, expression and occlusion.

3.3. Experiments with Obvious Variation on Illumination

In this subsection, the experiment focuses on whether VFRC is efficient for SSPP problem with obvious varying illumination. Extended Yale B face database [34, 35] is considered, since it includes 2,414 frontal-face images with 64 illumination conditions of size 640×480 from 38 subjects. For each subject, we reserve 59 images for experiments. All images (showed in Fig. 4) are cropped with dimension 96×84 and converted to gray scale. We choose the first 30 subjects with 59 gray images to establish the gallery set and the probe set. The gallery set is consisted of the first sample of each subject, and the rest samples are collected to the probe set. For generic set, the rest 8 subjects are considered. Each subject in the generic set contains 59 images.

For experiments of parameters setting, we choose the first sample per subject for training and the rest 58 samples for testing. According to the numerical

Table 2: The standard deviation on ARs database

Method	5	10	15	20	24
SRC	0.1041	0.0651	0.0381	0.0218	0.0078
CRC	0.0982	0.0507	0.0337	0.0191	0.0071
Block–LDA	**0.0569**	0.0399	0.0252	0.0214	0.0085
DMMA	0.1155	0.1315	0.0364	0.0254	0.0101
AGL	0.0656	0.0510	0.0361	0.0181	0.0096
ESRC	0.0761	0.0557	0.0263	0.0243	0.0091
SVDL	0.0701	0.0506	0.0292	0.0163	0.0087
VFRC	0.0834	**0.0395**	**0.0238**	**0.0133**	**0.0059**

results, the parameters are set respectively as: $\gamma = 0.993$, $\lambda_1 = 0.001$ and $\lambda_2 = 0.05$ in VFRC; $\lambda = 0.0083$ in CRC; $\lambda = 0.01$ and $ps = 12$ in Block-LDA; $ps_r = 48$ and $ps_c = 42$ in DMMA; $\lambda = 0.009$ in ESRC in the following experiments on Extended Yale B database.

Now, we implement VFRC and other methods on the above database. For every subject, we choose the first sample for training and the following first 10, 20, 30, 40, 50, 58 samples for testing, respectively. The recognition rates are listed in Table 3.

Table 3: Recognition rates (%) on Extended Yale B database - Fixed

Method	10	20	30	40	50	58
SRC	85.33	68.17	47.78	51.58	51.27	45.11
CRC	87.67	72.00	50.33	53.58	55.00	48.85
Block–LDA	76.33	61.33	43.56	47.17	44.40	39.71
DMMA	87.00	79.83	59.22	60.58	60.60	53.74
AGL	82.67	75.17	55.89	57.58	58.93	53.33
ESRC	88.33	85.67	67.67	67.00	67.87	62.36
SVDL	88.67	87.50	69.89	68.42	70.47	64.37
VFRC	**91.33**	**91.67**	**79.00**	**76.08**	**78.67**	**74.37**

As showed in Table 3, the more probe samples are added, the lower the recognition rate is. For different probe sample sizes, VFRC is 3% to 35% higher than other methods. According to the recognition results, the proposed VFRC keeps a high level of recognition rate and robustness among those methods. We can also conclude that VFRC is effective for SSPP problem with broad change in illumination.

Next, we carry out another experiments to check the performance of VFRC and other seven methods under the conditions with broad change of illumination on Extended Yale B database. For this, we take the first sample of each subject for training, and choose randomly 10, 20, 30, 40, 50, 57 samples from the rest samples for testing, respectively. Each experiment is carried out 10 times to guarantee stable experimental results, which are listed in Table 4.

Table 4: Recognition rates (%) on Extended Yale B database - Random

Method	10	20	30	40	50	57
SRC	45.83	46.60	43.17	45.19	45.41	44.95
CRC	52.50	50.33	49.24	49.11	48.87	49.16
Block–LDA	43.60	41.22	40.51	39.35	39.37	39.99
DMMA	46.40	56.02	53.81	53.28	53.51	53.27
AGL	52.53	51.32	53.46	53.68	53.85	53.40
ESRC	65.00	63.30	61.70	61.40	61.86	62.53
SVDL	68.20	65.83	63.89	64.16	64.56	64.42
VFRC	**74.20**	**72.43**	**75.16**	**75.87**	**74.25**	**74.37**

From Table 4, it is obvious that VFRC is very effective for SSPP problem with extreme variation on illumination.

3.4. Experiments with Change of Pose and Expression

The CMU-PIE database contains 41,368 source images of size 640×486 for 68 subjects with various conditions, i.e., 13 poses (from cameras No. 02, 05, 07, 09, 11, 14, 22, 25, 27, 29, 31, 34, and 37), 43 illuminations, 3 expressions and

3 talking situations for each subject. Depending on these factors, the images of each subject are classified into four different sets: expression, illum, lights and talking. The first expression set contains images with neutral expression, smiling, and blinking from all 13 cameras under the natural illumination. The second and third sets contain images with varying illumination conditions. The images from all 13 cameras are kept in the illum set, and from only 3 cameras (No. 05, 22, 27) are kept in lights set. The forth set contains 2 seconds (60 frames) of each subject talking when 3 cameras (No. 05, 22, 27) is available.

In this subsection, the aim of the experiment is to verify the performance of VFRC for SSPP problem with change of pose and expression. Thus, we choose a subset of CMU-PIE database, which contains 68 subjects with 37 images under two pose conditions from cameras No. 27, 05. For each subject, the first image is P27_07 (P27 represents the camera No. 27 and 07 represents the seventh light condition) from the lights set, which is regarded as the sample in gallery set. The following 36 images are consisted of 3 different expression images and 15 talking images from cameras No. 27 and 05. We choose the first 60 subjects with 37 images per subject for training and testing. The 296 images from rest 8 subjects are used for the generic set. In our experiment, all images are first eye-aligned and cropped, and resized into 112×100 pixels. The samples of the first subject we selected are showed in Fig. 5.

The experiments of parameters setting, which are analogous to those in Subsection 3.1, are done for the 60 subjects. According to the experiments results, the parameters are set as: $\gamma = 0.999$, $\lambda_1 = 0.02$ and $\lambda_2 = 0.03$ in VFRC; $\lambda = 0.003$ in CRC; $\lambda = 0.001$ and $ps = 50$ in Block-LDA; $ps_r = 50$ and $ps_c = 44$ in DMMA; $\lambda = 0.001$ in ESRC.

We carry out the experiments of the concerned eight methods for SSPP problem with changes of pose and expression on CMU-PIE database. The first sample of each subject is selected for training. From the rest samples, we choose randomly 6, 12, 18, 24, 35 samples for testing, respectively. Table 5 shows the average recognition rates of 10 times experiments on CMU-PIE database.

As the numerical results showed in Table 5, VFRC is effective for SSPP face

Table 5: Recognition rates (%) on CMU-PIE database

Method	6	12	18	24	30	35
SRC	24.83	24.19	23.99	24.01	23.65	23.89
CRC	30.86	28.43	30.25	29.28	29.66	29.39
Block–LDA	24.78	24.65	24.34	24.25	24.12	24.37
DMMA	39.69	31.00	29.13	29.90	29.63	29.08
AGL	17.67	16.65	17.59	17.31	17.32	17.43
ESRC	33.72	34.82	34.80	34.58	33.86	34.24
SVDL	40.31	41.00	40.31	40.86	41.13	40.79
VFRC	**42.64**	**41.39**	**41.92**	**41.24**	**41.33**	**41.25**

recognition with variation of pose and expression conditions.

3.5. Experiments with Obvious Occlusion

Comparing the experiment results showed in Table 5 with those showed in Table 1, Table 3 and Table 4, we can get the conclusion that the more complex (or obvious) the conditions of experiment databases is, the more observable the advantage of VFRC achieves. In order to adequately present the superiority of VFRC, we implement another two experiments with different occlusion conditions on ARs database and CMU-PIE databases.

In the occlusion experiments on ARs database, we choose the samples in the gallery set for training, and choose randomly 1 to 11 samples with occlusions (as sunglasses or scarf) for testing. The parameters, the gallery set and the generic set are the same as those in Subsection 3.2, respectively. Each experiment is repeated 10 times. The recognition rates are showed in Fig. 6. With different numbers of probe samples, VFRC always performs the best among the eight methods.

We carry out another four different random occlusion experiments on a subset of CMU-PIE database. In the following experiments on CMU-PIE database, the samples from camera No. 27 (showed in Fig. 5) are chosen for the gallery

set, the probe set and the generic set. The samples in the gallery set are for training, and random 3, 6, 9, 12, 15 probe samples with added occlusion are for testing, respectively. The occlusion rate is ranged from 5% to 20% by 5% a step. Every occlusion patch is randomly located on the samples in the probe set and the generic set (as showed in Fig. 7). The mean of the 10 trial data are showed in Fig. 8.

From Fig. 8, we can see that the performance of VFRC is obviously better than other seven methods, especially under the 15% and 20% random occlusion conditions. Comparing with Table 5, we can get the conclusion that the superiority of VFRC are presented prominently with severe variational information between the gallery set and the probe set.

3.6. Image-to-Set

In this subsection, we test the effectiveness of the proposed method for image-to-set experiments by using the LFW database [37], which is one of the most challenging databases in FR field. The database contains 13,233 face images of 5,749 different individuals, collected from the web, and 1,680 of the people pictured have two or more distinct images in this data set. The aligned vision LFWa[1] database is used in the following experiments. We select 151 subjects, that have more than ten images, to build the gallery and probe sets. The first ten images of each subject are selected for our experiments. For each subject, the first image is selected as the gallery sample, and the remaining images as the probe sample set. Furthermore, the first 250 subjects, that have more than two and less than ten images, are considered to build the generic set. In which, each subject have the first 3 images. All images are cropped into the size 64×64. Fig. 9 shows samples of first three subjects from the LFWa database for our experiments.

As the probe set per subject contains 9 probe samples, the final recognition

[1]The LFWa database is provided after alignment using commercial face alignment software, that can be downloaded at http://www.openu.ac.il/home/hassner/data/lfwa/.

results of SRC, CRC, Block-FLDA, DMMA, AGL, ESRC, SVDL and VFRC are obtained by the majority voting strategy, except for the image-to-set method LRP-SRSM. According to the parameter setting experiments, the parameters are set as: $\gamma = 0.9$, $\lambda_1 = 0.04$ and $\lambda_2 = 0.008$ in VFRC; $\lambda = 0.006$ in CRC; $\lambda = 0.006$ and $ps = 8$ in Block-LDA; $ps_r = 24$ and $ps_c = 24$ in DMMA; $\lambda = 0.009$ in ESRC; $k = 3$ in LRP-SRSM. The recognition rates and computational time are showed in Table 6.

Table 6: Image-to-set experiments on LFWa database

Method	SRC	CRC	Block−LDA	DMMA	AGL	ESRC	SVDL	LRP−SRSM	VFRC
Rate(%)	19.87	33.11	20.53	12.58	14.57	29.14	28.48	29.80	**43.05**
Time(s)	327.09	13.44	624.48	168.60	0.89	1730.71	12.30	2658.16	15.12

Because of the variations of varying lighting, expressions, poses, occlusions, and misalignment in LFWa database, the great challenges arise in our experiments. Fortunately, the results in Table 6 show the proposed VFRC is very effective for image-to-set SSPP problem with complicated variational information, no matter for the recognition accuracy or the computing cost.

3.7. Complexity and Computational Time

The experiments in the previous subsections demonstrate that the proposed VFRC method leads to higher accuracy and better robustness than other compared approaches. In this subsection, we analyze the computational complexity of VFRC method and display the comparisons of computational time about nine concerned approaches (including SRC, CRC, Block-LDA, DMMA, AGL, ESRC, SVDL, LRP-SRSM and the proposed VFRC).

In order to the convenient narrative, we let C be the number of sample in the gallery set, L be the column number of the variational matrix, and d be the pixel number of face sample. The proposed VFRC method can be divided into two stages: the training stage and the testing (recognition) stage. The former involves two optimal models with the close-form solutions (the computational complexity is $O((C + L)d^2)$), and is implemented as an off-line process. Therefore, the computational complexity of the training stage is usually ignored in

practical applications. On the contrary, the situation of the recognition stage has received the users' and researchers' continually attention. Our VFRC method has the computational complexity $O(d^2)$ in this stage.

The computational time of image-to-set experiments has been showed in Table 6. For the computational time of image-to-image experiments, as an example, we statistic the time spent on the recognition stage by the proposed VFRC method and the compared seven methods on Extended Yale B database with the same experiments setting of Table 4. The computational time showed in Table 7 is the average computational time of 10 times experiments for all subjects with each probe size conditions.

Table 7: computational time of experiments on Extended Yale B database (s)

Method	10	20	30	40	50	57
SRC	17.46	35.17	52.75	68.31	92.43	105.61
CRC	0.85	1.65	2.47	3.30	4.10	4.66
Block–LDA	27.33	54.07	81.67	108.45	130.61	149.66
DMMA	276.30	260.34	304.88	371.99	425.81	461.20
AGL	0.15	0.14	0.20	0.26	0.31	0.35
ESRC	102.26	219.70	335.41	439.67	582.00	592.68
SVDL	6.58	12.50	19.03	25.21	30.71	35.02
VFRC	2.78	4.27	5.84	7.19	8.74	9.68

From Tables 6 and 7, it can be seen that the computational time of CRC, AGL, SVDL and the proposed VFRC are obviously less than other methods for the identification process. Combined with the results of recognition rates, VFRC keeps a high accuracy with a favorable speed. This indicates the proposed method VFRC is of great availability for practical application of FR with SSPP problem.

4. Conclusions

In this paper, we proposed a Variational Feature Representation-based Classification (VFRC) method to improve the performance of the face recognition method for SSPP problem with non-ideal probe samples. In VFRC, the linear regression model was used to obtain the more precise variational feature of the probe sample. As the variational feature made the best of the coupled information with the gallery set and the generic set, the corresponding normal feature of the probe sample was got precisely. For the reason that the normal feature reserves few variational information, it is beneficial to enhance the identity information of the corresponding probe sample. Therefore, a combination of the normal feature and the non-ideal probe sample was used in our model.

Extensive numerical experiments were implemented on AR, Extended Yale B and CMU-PIE databases to verify the performance of VFRC for image-to-image SSPP problem under the complication conditions, broad changing in illumination conditions and change of pose and expression conditions, respectively. We also implemented VFRC on LFW database to check its performance for image-to-set SSPP problem with extremely complicated conditions. Eight related methods, SRC, CRC, Block-LDA, DMMA, AGL, ESRC, SVDL and LRP-SRSM, were chosen for the comparison with VFRC in recognition rates and computing time. The experiments results showed that VFRC is more efficient and robust than other eight methods for SSPP problems with non-ideal probe samples.

Acknowledgments

This work was partially supported by the National Natural Science Foundations of China (Grants 11171252 and 11431002).

References

[1] W. Zhao, R. Chellappa, P. J. Phillips, A. Rosenfeld, Face recognition: a literature survey, ACM Comput. Surv. 35 (4) (2003) 399–458.

[2] R. Brunelli, T. Poggio, Face recognition: features versus templates, IEEE Trans. Pattern Anal. Mach. Intell. 15 (10) (1993) 1042–1052.

[3] S. B. Chen, C. H. Ding, B. Luo, Extended linear regression for undersampled face recognition, J. Vis. Commun. Image R. 25 (7) (2014) 1800–1809.

[4] Z. H. Huang, W. J. Li, J. Wang, T. Zhang, Face recognition based on pixel-level and feature-level fusion of the top-levels wavelet sub-bands, Information Fusion 22 (2015) 95–104.

[5] B. Wang, W. Li, Z. Li, Q. Liao, Adaptive linear regression for single-sample face recognition, Neurocomputing 115 (2013) 186–191.

[6] H. Shahamat, A. A. Pouyan, Face recognition under large illumination variations using homomorphic filtering in spatial domain, J. Vis. Commun. Image R. 25 (2014) 970–977.

[7] Y. M. Chen, J. H. Chiang, Fusing multiple features for Fourier Mellin-based face recognition with single example image per person, Neurocomputing 73 (16) (2010) 3089–3096.

[8] A. Y. Yang, Z. Zhou, A. G. Balasubramanian, S. S. Sastry, Y. Ma, Fast l_1-minimization algorithms for robust face recognition, IEEE Trans. Image Process. 22 (8) (2013) 3234–3246.

[9] C. Y. Lu, H. Min, J. Gui, L. Zhu, Y. K. Lei, Face recognition via weighted sparse representation, J. Vis. Commun. Image R. 24 (2013) 111–116.

[10] M. Yang, L. Zhang, S. C. Shiu, D. Zhang, Gabor feature based robust representation and classification for face recognition with Gabor occlusion dictionary, Pattern Recognit. 46 (7) (2013) 1865–1878.

[11] X. Tan, S. Chen, Z. H. Zhou, F. Zhang, Face recognition from a single image per person: a survey, Pattern Recognit. 39 (9) (2006) 1725–1745.

[12] C. Lu, W. Liu, S. An, Face recognition with only one training sample, in: Chinese Control Conference (CCC), 2006, pp. 2215–2219.

[13] M. Turk, A. Pentland, Eigenfaces for recognition, J. Cognitive Neurosci. 3 (1) (1991) 71–86.

[14] P. N. Belhumeur, J. P. Hespanha, D. Kriegman, Eigenfaces vs. Fisherfaces: recognition using class specific linear projection, IEEE Trans. Pattern Anal. Mach. Intell. 19 (7) (1997) 711–720.

[15] T. Ahonen, A. Hadid, M. Pietikäinen, Face recognition with local binary patterns, in: European Conference on Computer Vision, 2004, pp. 469–481.

[16] D. Zhang, S. Chen, Z. H. Zhou, A new face recognition method based on SVD perturbation for single example image per person, Appl. Math. Comput. 163 (2) (2005) 895–907.

[17] Q. X. Gao, L. Zhang, D. Zhang, Face recognition using FLDA with single training image per person, Appl. Math. Comput. 205 (2) (2008) 726–734.

[18] G. McLachlan, Discriminant analysis and statistical pattern recognition, Vol. 544, John Wiley & Sons, 2004.

[19] R. Gottumukkal, V. K. Asari, An improved face recognition technique based on modular PCA approach, Pattern Recognit. Lett. 25 (4) (2004) 429–436.

[20] S. Chen, J. Liu, Z. H. Zhou, Making FLDA applicable to face recognition with one sample per person, Pattern recognit. 37 (7) (2004) 1553–1555.

[21] J. Lu, Y. P. Tan, G. Wang, Discriminative multimanifold analysis for face recognition from a single training sample per person, IEEE Trans. Pattern Anal. Mach. Intell. 35 (1) (2013) 39–51.

[22] H. Yan, J. Lu, X. Zhou, Y. Shang, Multi-feature multi-manifold learning for single-sample face recognition, Neurocomputing 143 (2014) 134–143.

[23] T. K. Kim, J. Kittler, Locally linear discriminant analysis for multimodally distributed classes for face recognition with a single model image, IEEE Trans. Pattern Anal. Mach. Intell. 27 (3) (2005) 318–327.

[24] X. Wang, X. Tang, Random sampling for subspace face recognition, Int. J. Comput. Vis. 70 (1) (2006) 91–104.

[25] Y. Su, S. Shan, X. Chen, W. Gao, Adaptive generic learning for face recognition from a single sample per person, in: IEEE Conference on Computer Vision and Pattern Recognition, 2010, pp. 2699–2706.

[26] J. Wright, A. Y. Yang, A. Ganesh, S. S. Sastry, Y. Ma, Robust face recognition via sparse representation, IEEE Trans. Pattern Anal. Mach. Intell. 31 (2) (2009) 210–227.

[27] H. Qiu, D. S. Pham, S. Venkatesh, W. Liu, J. Lai, A fast extension for sparse representation on robust face recognition, in: 20th International Conference on Pattern Recognition (ICPR), 2010, pp. 1023–1027.

[28] H. Qiu, D. S. Pham, S. Venkatesh, J. Lai, W. Liu, Innovative sparse representation algorithms for robust face recognition, Int. J. Innov. Comput. I 7 (10) (2011) 5645–5667.

[29] W. Deng, J. Hu, J. Guo, Extended SRC: undersampled face recognition via intraclass variant dictionary, IEEE Trans. Pattern Anal. Mach. Intell. 34 (9) (2012) 1864–1870.

[30] W. Deng, J. Hu, J. Guo, In defense of sparsity based face recognition, in: IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2013, pp. 399–406.

[31] M. Yang, L. Van, L. Zhang, Sparse variation dictionary learning for face recognition with a single training sample per person, in: IEEE International Conference on Computer Vision, 2013, pp. 689–696.

[32] D. Zhang, M. Yang, X. Feng, Sparse representation or collaborative representation: which helps face recognition?, in: IEEE International Conference on Computer Vision, 2011, pp. 471–478.

[33] A. M. Martinez, R. Benavente, The AR face database, Technical report, CVC Technical Report (1998).

[34] A. S. Georghiades, P. N. Belhumeur, D. Kriegman, From few to many: illumination cone models for face recognition under variable lighting and pose, IEEE Trans. Pattern Anal. Mach. Intell. 23 (6) (2001) 643–660.

[35] K. C. Lee, J. Ho, D. Kriegman, Acquiring linear subspaces for face recognition under variable lighting, IEEE Trans. Pattern Anal. Mach. Intell. 27 (5) (2005) 684–698.

[36] T. Sim, S. Baker, M. Bsat, The CMU pose, illumination, and expression database, IEEE Trans. Pattern Anal. Mach. Intell. 25 (12) (2003) 1615–1618.

[37] G. B. Huang, R. Manu, B. Tamara, L. M. Erik, Labeled faces in the wild: A database for studying face recognition in unconstrained environments, Tech. Rep. 07–49, University of Massachusetts, Amherst (October 2007).

[38] J. Lu, Y. P. Tan, G. Wang, G. Yang, Image-to-set face recognition using locality repulsion projections and sparse reconstruction-based similarity measure, IEEE Trans. Circuits Syst. Video Technol. 23 (6) (2013) 1070–1080.

[39] S. J. Kim, K. Koh, M. Lustig, S. Boyd, D. Gorinevsky, An interior-point method for large-scale l_1-regularized least squares, IEEE J. Sel. Top. Sign. Proces. 1 (4) (2007) 606–617.

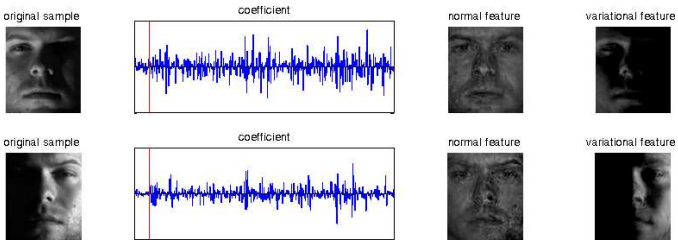

Fig. 1: The first column includes two probe samples of the first subject on Extended Yale B database. The second column shows the representation coefficients over the gallery set and the variation matrix. The third and fourth columns are the reconstructed normal faces and variational information of illumination, respectively.

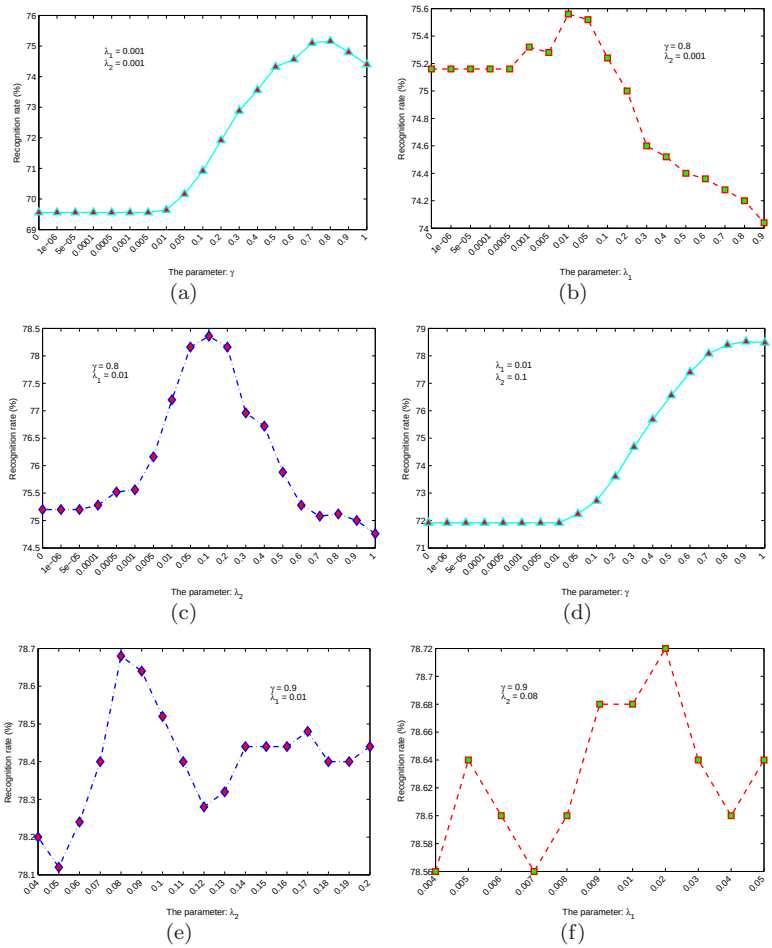

Fig. 2: Parameters setting

Fig. 3: All samples of the first subject on AR Database

Fig. 4: All samples of the first subject on Extended Yale B database

Fig. 5: All samples of the first subject on CMU-PIE database. The first column shows the sample of the first subject in the gallery set. The 18 samples of the first and second rows are the frontal (P27) probe samples of the subject with 3 expression images and 15 talking images. The rest 18 probe samples are from camera No. 05 with 3 expression images and 15 talking images.

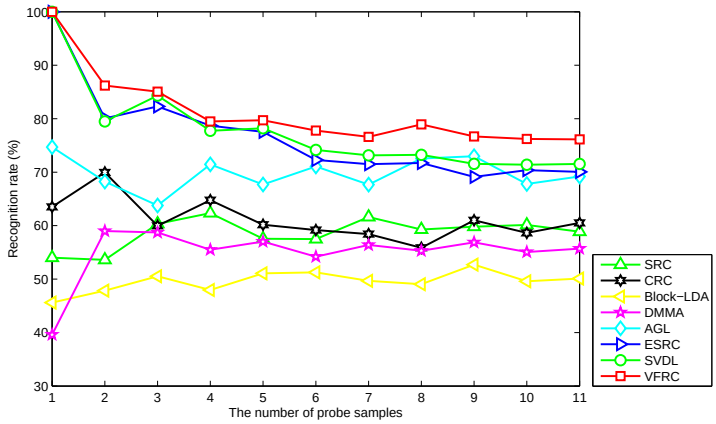

Fig. 6: The recognition results with different occlusion on ARs database.

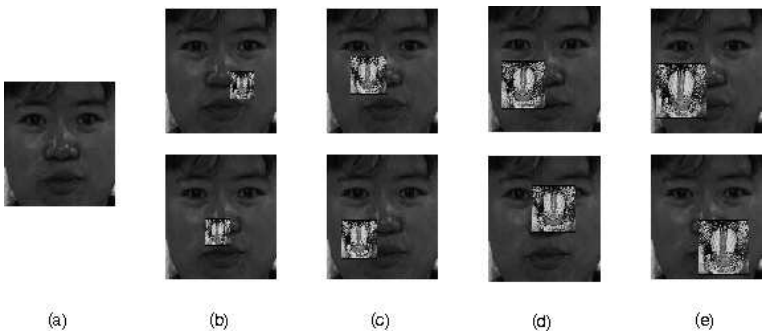

Fig. 7: The samples of the random occlusions in the probe set of the first subject on CMU-PIE database. (a) is the first original probe sample in this subject. (b) - (e) respectively show the sample (a) with 5%, 10%, 15% and 20% occlusion rates in two different random location.

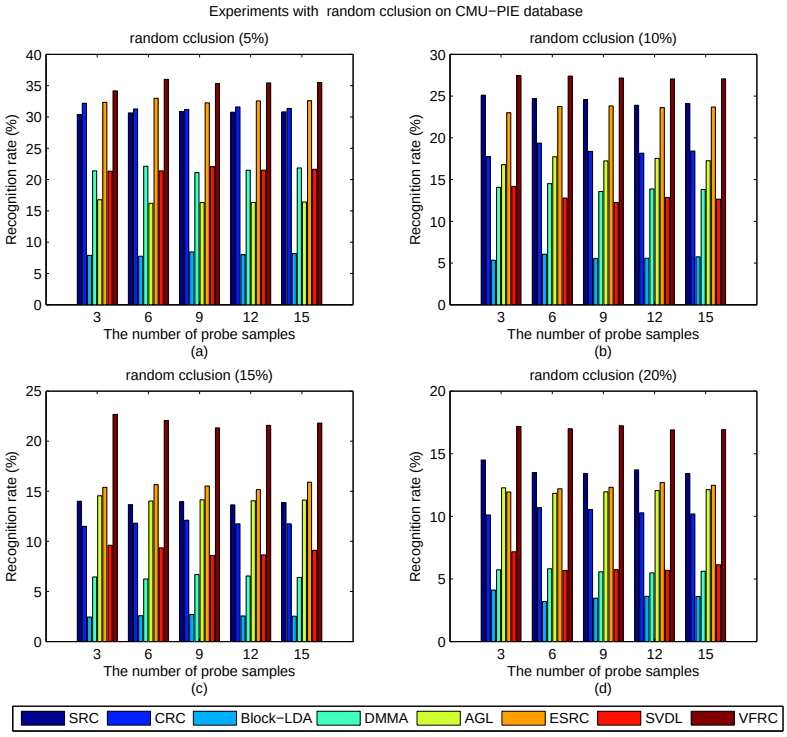

Fig. 8: The recognition results with different random occlusion on CMU-PIE database.

Fig. 9: Samples of the first three subjects from the LFWa database. The images in each row are from the same subject.