

# Occluded Face Recognition Based on Dictionary Learning and Sub-classifier Fusion

Nan Deng<sup>1</sup>, Zhengguang Xu<sup>1</sup>, and Wang Jun<sup>2</sup>

1. Institute of Control Science and Engineering, University of Science and Technology Beijing, 100083, China

2. School of Electronics and Information, Northwestern Polytechnical University, Xi'an 710072, China

e-mail:dengnan666666@163.com

**Abstract**—Facial recognition is a challenging area of research due to difficulties with robust face recognition (FR) under occlusion and sparse representation-based classification (SRC) only focusing on face global features. To solve these issues, we proposed an occluded FR method based on dictionary learning for sparse representation and sub-classifiers fusion (LSSRC), which efficiently combines local and overall characteristics of face images. First, we partitioned continuous but non-lapped blocks of the face by multi-resolution blocking. Then, for each block, SRC was used for feature extraction and face classification. We established a sub-block dictionary and conducted K-SVD dictionary learning, established sub-classifiers and determined weight. Finally, we conducted sub-classifier fusion recognition using voting rules with weight. Results using the AR and YaleA database showed that our algorithm achieved superior recognition performance to the existing sparse representation classification occluded FR method.

**Index Terms**—Face Recognition, K-SVD, Dictionary Learning, Sparse Representation, Sub-Classifier Fusion

## I. INTRODUCTION

Face recognition (FR) [1], as biological feature recognition, is extensively utilized nowadays in authentication, video monitoring, and information security. However, in actual face image processing, FR robustness is affected by illumination, posture, facial expression and occlusion such as glasses, masks, scarves and decorations. These occlusions not only cause the absence of local features in face image, but also have an influence on the accuracy of whole features. Determining how to quickly and accurately conduct occluded FR has become of increasing importance in recent years.

Sparse representation classification (SRC) [2-4] has become a representative occluded FR method in recent years. In the SRC method, the face image under occlusion can be regarded as the summation of the non-occluded face image and occlusion error. For a small percentage of occlusion, the non-occluded part is considered to be sparsely coded by the dictionary of non-occluded training samples only, and the occlusion error also has a sparse representation (SR) over the occlusion dictionary. As a result, classification can be performed by calculating the sparse representation coefficients over the expanded dictionary composed of the occlusion dictionary and

non-occluded training dictionary together with the minimum reconstruction error. The dictionary can be composed based on the original training sample or after conducting feature extraction of the original training sample, and the optimal dictionary can be obtained after K-SVD learning [5].

The SRC based occluded FR can be grouped into two categories. One category refers to SRC based on overall characteristics of the face image with limitations or prior knowledge. Wright et al. [2] proposed block sparse representation based on classification (BSRC), composed multiple dictionaries through combining SRC with face blocks and conducted recognition by voting algorithm. Zhai et al. [6, 7] combined the homotopy algorithm with color information fusion to recognize the occluded face image. Mao et al. [6] proposed a multi-resolution based sparse representation model for face recognition (MSRC). Yang et al. [8] established a Gabor occlusion dictionary to compress the number of atoms after filtering the original face using a Gabor filter (GSRC), and thus reduced computational cost in sparse coding and improved the occlusion recognition rate. The second category refers to sparse representation based on local characteristics of the face image, among which Zhen et al. [9] proposed the sparsely encoded local descriptor (SELD), fused the K-SVD learning dictionary with local information coding of an image, and determined categories using cosine similarity. The SELD was very different from the existing texture based local descriptor in that the sparse coding led to an image descriptor of summation of sparse coefficient vectors. Thus, SRC obtains good occluded recognition performance by means of combining total or partial characteristics of the occluded face.

Classifier fusion can be divided into abstraction, arrangement, and fraction fusions based on different classifier outputs [10]. Fraction fusion can be divided into product, sum, maximum, minimum, mid-value, and voting rules, which not only shield the diversity of the image features and complexity of the recognition process, but also store the similarity metrics of various features. Of them, the voting rule is characterized by simple algorithm, flexible use and wide application.

Based on the above, we proposed a partially occluded face recognition method based on dictionary learning and sub-classifiers fusion (LSSRC). We first partitioned

multi-resolution blocks (MAB) and feature extraction for the face, conducted sub-block dictionary learning using K-SVD, solved sparse vectors and established sub-classifiers, and finally conducted sub-classifier fusion recognition using voting rules with weight. This achieved a higher recognition rate compared to the existing SRC and BSRC methods.

The main contributions of this paper are as follows: First, the SRC method containing overall characteristic information was introduced to conduct local blocking of the face image, which excavated the local and overall characteristics of the face image for FR. Second, the K-SVD dictionary learning optimized the dictionary and SRC was introduced to establish sub-classifiers with various resolutions and space positions. Finally, the classification results of all the blocks were aggregated by statistical voting with weight. The advantage of this method is its ability to conduct fusion recognition based on different identification contributions of sub-classifiers.

## II. SPARSE REPRESENTATION CLASSIFICATION FOR FR

### A. Sparse Approximation Problem

Sparse representation [2] supposes that there have a dictionary  $D \in M^{N \times M}$  with properties as follows:

Firstly the dictionary is redundant:  $M \geq N$ ; Secondly the dictionary is full rank( $\text{rank}(D)=N$ ); Thirdly the dictionary is normalized:  $\|d_i\|_2=1, i=1, \dots, M$ . Then we can find a representation  $\alpha$  for a given signal  $f$ :

$$D\alpha = f \quad (1)$$

This is a under-determined linear equation—infinite number of solutions. The given signal  $f$  and dictionary of Sparse representation are shown in Figure 1.

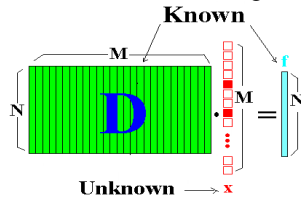


Figure 1. Sparse representation

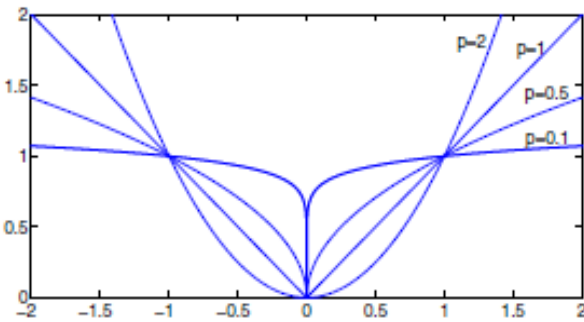


Figure 2.  $\|\alpha\|_p$ .

We think the “best” one is the sparsest, so we need to solve the following optimization problem:

$$\min_{\alpha} \|\alpha\|_0, \quad \text{s.t.} \quad D\alpha = f \quad (2)$$

where the  $l_0$ -norm means the number of nonzero entries in vector  $\alpha$ . (1) is called Sparsest representation problem.

$$\|\alpha\|_p := [\sum_{i=1}^M |\alpha(i)|^p]^{1/p} \quad (3)$$

is called  $l_p$ -norm. From the figure 2 we see that when  $0 < p < 1$ ,  $\|\alpha\|_p$  is concave; when  $p \geq 1$ ,  $\|\alpha\|_p$  becomes convex.

In the case of observed data  $f$  containing noise, the equality constraint in (1) should be relaxed to be inequality:

$$\min_{\alpha} \|\alpha\|_0, \quad \text{s.t.} \quad \|D\alpha - f\|_2 \leq \sigma \quad (4)$$

This is called Sparse Approximation problem. To overcome the concave difficulty and still promote sparsity we use  $l_{1-norm}$  as a closest approximation of  $l_0$ -norm, so that problem (1) is reformulated as:

$$\min_{\alpha} \|\alpha\|_1, \quad \text{s.t.} \quad D\alpha = f \quad (5)$$

which corresponds exact data and is called Basis Pursuit (BP) problem. To obtain the solution for problem (1), we need to exhaustively test all possible combinations of  $k$  ( $=1, 2, \dots$ ) atoms, until the constraint is satisfied. The number of combination for a fixed  $k$  is  $M/k(M-k)$ , it grows exponentially with  $k$ , so this is a NP—Hard problem, means computational intractable. Normally we solve it by Marching Pursuit algorithm (MP) [2, 8] or Orthogonal Matching Pursuit (OMP) method [2, 8].

### B. Marching Pursuit Algorithm and Orthogonal Matching Pursuit algorithm

The Marching Pursuit algorithm for solving problem (5) was proposed by Mallat and Zhang [2] in 1993. It is an iterative procedure.

Input:  $D \in M^{N \times M}$ ,  $f \in R^N$ , stopping criterion.

Initialization:  $r = f$  ( $r \in R^N$ ),  $\alpha = 0$  ( $\alpha \in R^M$ ),  $k=1$

While stopping criterion is not hold.

1. Seek one atom from  $D$ , which is most strongly correlated with  $r$ , that means

$$d_{i_k} = \max_{d \in D} \langle r, d \rangle \quad (6)$$

2. Set

$$\alpha(i_k) = \langle r, d_{i_k} \rangle \quad (7)$$

3. Update residual

$$\begin{aligned} r &\leftarrow r - \alpha(i_k) d_{i_k} \\ r &\leftarrow r - \end{aligned} \quad (8)$$

and

$$k \leftarrow k + 1 \quad (9)$$

End

Remark: For sparse coding, stop at  $k=K$ ; For sparse approximation, stop at  $\|r\| \leq \varepsilon$ ; The step 1 can be accomplished by compute  $\alpha = D^T r$  followed by sorting the coefficient  $\{\alpha\}$ .

The Orthogonal Matching Pursuit (OMP) method is an improved MP algorithm.

1. Initialize:  $\Omega = \emptyset$ ,  $r_0 = f$ ,  $k = 1$

2. Indentify: find an atom that is most strongly correlated with  $r_{k-1}$ :

$$d_k = \max_{j=1, \dots, M} |\langle r_{k-1}, d_j \rangle| \quad (10)$$

Then

$$\Omega \leftarrow \Omega \cup i_k \quad (11)$$

3. Estimation: Compute an current approximate representation  $\alpha_k$  by solving a LS problem:

$$\min_{\beta} \|f - D_{\Omega} \beta\|_2^2 \quad (12)$$

where  $D_{\Omega}$  is a  $N \times K$  matrix contenting those  $k$  atoms that belong to  $\Omega$ . So  $\beta \in R^k$ .

4. Update :

$$\alpha_k(i) = \begin{cases} \beta(p), & \text{if } i = i_p \in \Omega_k; \\ 0, & \text{otherwise} \end{cases};$$

$$r_k = f - D_{\Omega} \alpha_k; \quad err = \|r_k\|_2; \quad (13)$$

$$k \leftarrow k + 1$$

Repeat 2-4 until stepping criterion holds. The stopping criterion is the same as the Marching Pursuit algorithm.

The main difference between OMP and MP is in step 3 and 4. In MP,  $\alpha(i_p)$  ( $p=1, \dots, k-1$ ) are fixed, only  $\alpha(i_p)$  is computed; In OMP, all  $\alpha(i_p)$ ,  $p=1, \dots, k$  are updated, so OMP outperform MP.

### C. Parse Representation Classification

Sparse representation classification (SRC) [2] assumes that any one test image can be represented by the linear combination of the training images of the same face. The category is determined by calculating the sparse coefficient of the test image with respect to all training images. The formal description is as follows.

First, the set of training samples of the  $i_{th}$  class can be expressed as  $\Phi_i = [A_{i,1}, \dots, A_{i,j}, \dots, A_{i,n_i}] \in R^{m \times n_i}$ , where  $A_{i,j}$ ,  $j=1, \dots, n_i$  is a  $m$  dimensional column vector composed of the  $j_{th}$  sample image of the  $i_{th}$  training class. In the sparse representation model,  $A_{i,j}$  is also called atom, and the combination of the atoms composes a dictionary.

The testing sample  $x \in R^m$  from this class can be represented by  $x = \sum_{j=1}^{n_i} \theta_{i,j} A_{i,j} = \Phi_i \theta_i$ , where  $\theta_i = [\theta_{i,1}, \theta_{i,2}, \dots, \theta_{i,n_i}] \in R^{n_i}$  are the weights.

Denoting the over-complete dictionary matrix composed by atoms from the whole training images of all object classes as by, the testing image of the training subject can be represented as a linear combination of all training images:

$$x = \Phi \theta \in R^m \quad (14)$$

where

$$\theta = [\theta_1; \dots; \theta_i; \dots; \theta_k] = [0, \dots, 0, \theta_{i,1}, \dots, \theta_{i,n_i}, 0, \dots, 0]^T \in R^m.$$

All image samples are normalized and their dimensions are reduced as follows:

$$\tilde{x} = T \Phi \theta = \tilde{\Phi} \theta \in R^d \quad (15)$$

where  $T \in R^{d \times m}$ ,  $d \leq m$  is the feature transformation matrix. Similarly, the above testing image under occlusion can be rewritten as:

$$\tilde{x}_e = \tilde{x} + e = \tilde{\Phi} \theta + e = [\tilde{\Phi}, I] \begin{bmatrix} \theta \\ e \end{bmatrix} = \tilde{B} \omega \quad (16)$$

where  $\tilde{B} = [\tilde{\Phi}, I] \in R^{d \times (n+n_e)}$ .

Obviously the sparser  $\theta$  or  $\omega$  is, the more concise the image can be represented. The sparsest representation of image can be acquired by solving the  $l^0$ -minimization problem of  $\theta$  or  $\omega$ , i.e.

$$\hat{\theta}_0 = \arg \min \|\theta\|_0, \text{ s.t. } \|\tilde{x} - \tilde{\Phi} \theta\| \leq \varepsilon \quad \text{without occlusion or}$$

$$\hat{\omega}_0 = \arg \min \|\omega\|_0, \text{ s.t. } \|\tilde{x}_e - \tilde{B} \omega\| \leq \varepsilon \quad \text{with occlusion,}$$

where  $\|\cdot\|_0$  denotes  $l^0$ -norm [2]. In practice, this problem is efficiently solved by minimizing  $l^1$ -norm[11][12]:

$$\hat{\theta}_1 = \arg \min \|\theta\|_1, \text{ s.t. } \|\tilde{x} - \tilde{\Phi} \theta\| \leq \varepsilon \quad (17)$$

or

$$\hat{\omega}_1 = \arg \min \|\omega\|_1, \text{ s.t. } \|\tilde{x}_e - \tilde{B} \omega\| \leq \varepsilon \quad (18)$$

To identify the object, the reconstruction errors are calculated:

$$r_i(\tilde{x}) = \left\| \tilde{x} - \tilde{\Phi} \delta_i(\hat{\theta}_1) \right\|_2 \quad (i=1, \dots, k) \quad (19)$$

where  $\delta_i(\theta) \in R^n$  is the characteristic function which selects the coefficients of the  $i_{th}$  class. With occlusion, the above reconstruction errors are modified as:

$$r_i(\tilde{x}_e) = \left\| \tilde{x}_e - \hat{e}_1 - \tilde{\Phi} \delta_i(\hat{\theta}_1) \right\|_2 \quad (i=1, \dots, k) \quad (20)$$

where  $\hat{e}_1$  is the sparse error, and the classification is based on the minimization reconstruction errors

$$\begin{aligned} \text{identify}(\tilde{x}) &= \arg \min r_i(\tilde{x}) \text{ or } \text{identify}(\tilde{x}_e) \\ &= \arg \min r_i(\tilde{x}_e) \end{aligned} \quad (21)$$

The recognition of an occluded testing image by SRC is shown in Figure 1.

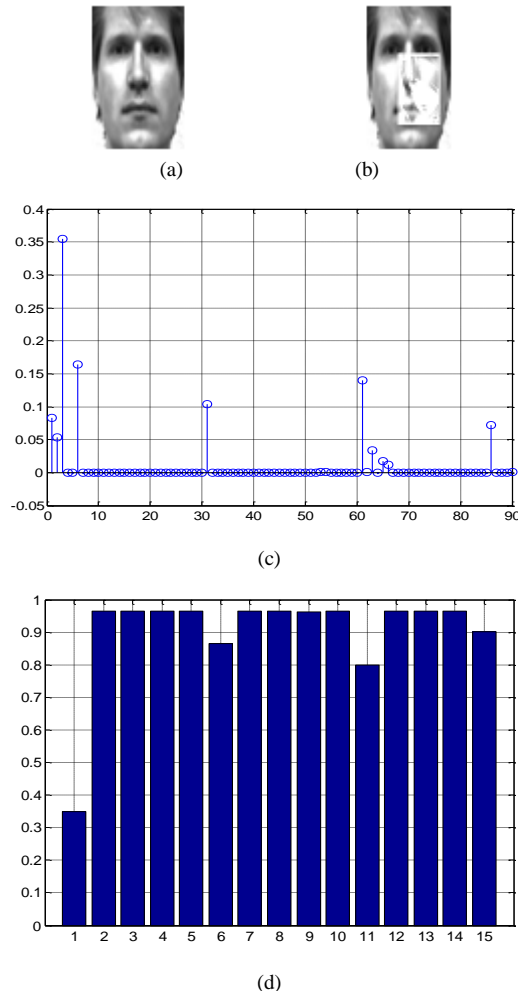


Figure 3. Recognition of an occluded testing image by SRC. (a) A test image. (b) 25% occluded test face image from YaleA database. (c) Values of sparse coefficients. (d) Reconstruction errors of test image.

The problem of finding the solution of the  $\ell_1$ -minimization [10] of an underdetermined system of linear equations is NP-hard [2], and the normally used methods include greedy matching pursuit method, relaxation optimization method, combinatorial optimization method, etc. The greedy matching pursuit (MP) method [11] utilizes local greedy method to find the solution. The modified MP method include the orthogonal matching pursuit (OMP) method [8, 11], the stage-wise orthogonal matching pursuit (StOMP) method [12], etc. Those modified methods have lower complexity, but the optimal result cannot be guaranteed. For the relaxation optimization method,  $\ell_1$ -norm or some other sparsity measuring functions are generally substituted for the original non-convex  $\ell_1$ -norm to transfer the original problem into a convex programming or nonlinear programming problem. The frequently used relaxation optimization methods include the basis pursuit method

[11], the focuss method [12], and the threshold method [12], etc. These methods have higher reconstruction precisions, but are more complex than the MP methods. The Heuristic search based combinatorial optimization methods, such as the simulated annealing method, the genetic algorithm, the ant algorithm [12], etc. can also be adopted in  $\ell_1$ -minimization. These kind of methods can lower the complexity of computation for the minimum value of  $\ell_1$ -norm, but tend to have a local optimal solution. The optimization method utilized in this paper is the relaxation optimization based primal-dual interior-point algorithm [11, 12].

### III. OCCLUDED FACE RECOGNITION BASED ON DICTIONARY LEARNING AND SUB-CLASSIFIER FUSION

Here we introduce the overall process of the partially occluded face recognition method based on dictionary learning and sub-classifiers fusion method (LSSRC). The key points of the method including dictionary learning for SRC and sub-classifier fusion are presented.

#### A. Total Flow of LSSRC Method for FR under Occlusion

In the training phase, we conducted multi-resolution blocks (MAB) for the face, and then completed feature extraction and dimensionality reduction for the sub-blocks of the training images at the corresponding positions by PCA [13], LDA [13], down-sampling [2], and random projection algorithm [2] generated by a Gaussian random matrix method (random) separately.

During the testing stage, the testing image was first processed by MAB to obtain sub-blocks as introduced above. The dimension of each sub-block was reduced by the characteristic transformation matrix obtained in the training stage. Then, each sub-block of the testing image was classified by SRC over the same sub-blocks of all training image. The sub-classifiers used K-SVD dictionary learning and greedy tracking algorithm, and determined weights based on the recognition rate of the sub-classifiers. Finally, the fusion recognition of the testing image was made by voting rules with weight [14-16].

The system diagram of the proposed method is shown in Figure 2.

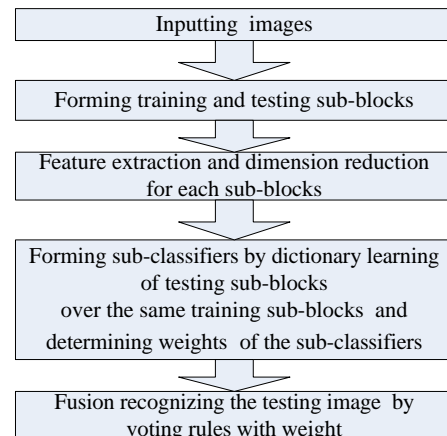


Figure 4. System diagram of the proposed LSSRC-based FR method.

### B. Image Preprocessing

The image preprocess in this paper includes geometry normalization and illumination normalization. The geometry normalization process was carried out, and all the face images were normalized to the same size based on the geometric center distance of eyes. Then, these images were processed by histogram equalization to reduce the effect of illumination. The image preprocessing results are shown in Figure 3.



Figure 5. Normalized face images

### C. Multi-resolution Block, Feature Extraction and Dimensionality Reduction of Face Image

Multi-resolution decomposition of image can be accomplished by pyramid decomposition in space domain or wavelet decomposition in frequency domain. The image pyramid is composed of a series of images in pyramid order with gradually decreasing resolutions and contains the image information both locally and globally. When face images with different resolutions are obtained by the pyramid decomposition, images with higher resolution can be divided into more blocks due to the fact that human eyes are able to observe more local information of object with higher resolution. Accordingly, images with lower resolution can be divided into fewer blocks. The adaptive blocking refers to the number of the sub-blocks depends on the resolution of the image in partitioning, which accords with human vision.

We first evenly divided the normalized image into sixteen blocks by image sub-sampling [4], and obtained the first layer of sub-blocks. We conducted down-sampling to evenly divide the sub-blocks into four blocks to obtain the second layer of sub-blocks; and continued down-sampling to obtain the third layer of sub-blocks. In total, twenty-four non-overlapping image sub-blocks were obtained as shown in Figure 4. Finally, After partition, the PCA, LDA, down-sampling, and Random methods are utilized in this paper to conduct feature extraction and dimension reduction on training and testing datasets.

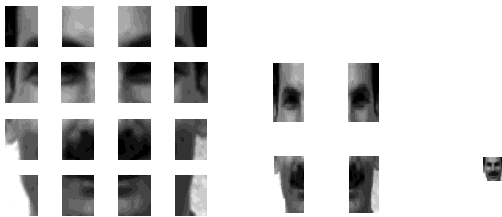


Figure 6. MAB.

### D. K-SVD-based Dictionary Learning and Sub-classifier

#### 1) K-SVD-based Dictionary Learning

K-SVD dictionary learning is used to find the optimal basis under sparse representation to meet its unique conditions. It can effectively reduce the number of atoms, improve atomic irrelevance, and obtain more sparse and accurate sparse solution.

Assuming that  $X = [x_1, x_2, \dots, x_M]$  is a training set,  $D$  is an over-complete dictionary, dictionary dimension is  $n$ ,  $k$  is the number of atoms in the dictionary,  $M$  is the number of samples, and  $A = [\alpha_1, \alpha_2, \dots, \alpha_M]$  is the corresponding coefficient of characterization, the sparse representation model is as follows:

$$\hat{\alpha} = \arg \min_{\alpha} \left\{ \|X - D\alpha\|_2^2 + \lambda \|\alpha\|_0 \right\} \quad (22)$$

and the iterative solution can be conducted in the following two stages.

The first stage is sparse coding. Assuming that dictionary  $D$  (first  $K$  samples can be used as an initialized dictionary) is a constant value, the sparse representation vector  $\alpha_i$  of each input training sample can be calculated by matching pursuit as follows:

$$\alpha_i = \min_{\alpha} \|\alpha\|_0, \quad s.t. \quad \|\mathbf{x}_i - D\alpha\| \leq \varepsilon; \quad i = 1, 2, \dots, M \quad (23)$$

The second stage is dictionary updating. Strategies are similar to that in the sparse coding stage, i.e., first fix the sparse representation vector matrix  $A$  obtained in sparse coding stage, assuming that  $\alpha^k$  is the  $k_{th}$  column of  $A$ , successively upgrade the first row of the dictionary (every atom  $d_k$ ), and total error is as follows:

Assuming that all  $d_i$  and  $\alpha_i$  are constant except  $d_k$  and  $\alpha^k$ , the error can be represented as:

$$E = \left\| \left( X - \sum_{j \neq k} \mathbf{d}_j \alpha^j \right) - \mathbf{d}_k \alpha^k \right\|_F^2 = \|E_k - \mathbf{d}_k \alpha^k\|_F^2 \quad (24)$$

In this formula,  $E_k$  represents the error matrix after removing  $d_k$  from dictionary  $D$ . Assuming that  $\omega_k = \{i | 1 < i < K, \alpha^k(i) \neq 0\}$  is a non-zero index entry of  $\alpha^k$ , defining the matrix  $\Omega_k$  with a size of  $M \times |\omega_k|$ , setting it to be 1 at  $(\omega_k(i), i)$  and 0 at other places, we can obtain dimensionality reduction matrix  $\alpha_R^k = \alpha^k \Omega_k$  and  $\alpha_R^k \in 1 \times |\omega_k|$  by multiplying  $\alpha^k$  and  $\Omega_k$ , i.e., remove non-zero item. In a similar way, we can obtain

$$x_R^k = x \Omega_k \quad (25)$$

$$x_R^k \in n \times |\omega_k| \quad (26)$$

and

$$E_R^k = E_k \Omega_k \quad (27)$$

The following formula can be gained after multiplying the above formula by  $\Omega_k$ :

$$\|E_k \Omega_k - d_k \alpha^k \Omega_k\|_F^2 = \|E_R^k - d_k \alpha_R^k\|_F^2 \quad (28)$$

Finally, we can conduct singular value decomposition (SVD) for  $E_R^k$  with the formula  $E_R^k = U \Delta V^T$ , upgrade  $d_k$  using the first column of matrix  $U$  after decomposition, upgrade  $\alpha_R^k$  using the first column of  $V \times \Delta(1,1)$ , and upgrade the first column of dictionary sequentially until the end.

The commonly used algorithms for  $l^0$  norm minimization include greedy tracking [8] and relaxation optimization [1]. This paper conducted sparse decomposition by the orthogonal matching pursuit method [8].

## 2) Establishment of Sub-classifier

We conducted multi-resolution blocks for the image, and established the sparse characterization model of sub-classifiers as follows:

$$\hat{\rho}_j = \arg \min_{\rho_j} \left\{ \|y_j - M_j \rho_j\|_2^2 + \lambda \|\rho_j\|_0 \right\} \quad (29)$$

In  $M_j = [M_{j1}, M_{j2}, \dots, M_{jc}]$ ,  $m_{jc}$  represents the sub-classifier dictionary of the  $j_{th}$  sub-block, and  $M_{jk}$  represents the sub-classifier dictionary of the  $j_{th}$  sub-block in the  $k_{th}$  category of Class C.

We solved each sub-classifier dictionary and sparse characterization vector by K-SVD and matching pursuit, and solved the sub-classifier recognition results  $z_j$  using minimum reconstruction residual as:

$$z_j = \text{identify}(y_j) = \arg \min_k (r_{jk}) \quad (30)$$

## E. Fusion Recognition of Sub-classifiers based on Different Weights

Average recognition rate of sub-classifiers reflects their different contributions to the final recognition result, so it can be used for calculating weight. We conducted fusion recognition for sub-classifiers by majority voting with weight [7].

Given a set of training samples including Class C, the  $1, \dots, j, \dots, J (J=21)_{th}$  sub-classifiers can be formed successively after partitioning. Assuming that  $\bar{n}_j$  represents average recognition rate of  $j_{th}$  sub-classifier, weight can be defined as  $w_j = \bar{n}_j / \sum_{j=1}^J \bar{n}_j$ , and face recognition as:

$$\text{identify}(y) = \arg \max_{1 \leq k \leq C} \sum_{j=1}^J w_j z_j \quad (31)$$

## IV. SIMULATION AND DISCUSSION

There are many standard databases available at present. These databases were selected in this experiment. The first was the AR [2] database, composed of 3288 frontal face images of 126 individuals under various conditions

such as expression, illumination, aging, and occlusion. In this experiment, 100 individuals were selected. For each individual, 26 images were taken in two separate sessions, among which 13 images from one session are shown in Figure 5.

The second database was the YaleA [2] database, which contained 165 images of 15 individuals with different expression and illumination as shown in Figure 6. And the third database was the ORL [2] database as shown in Figure 7.

All images were normalized to 100\*100 pixels in the experiment.

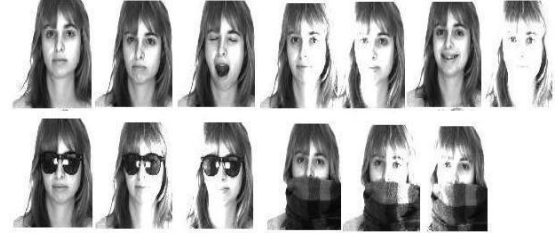


Figure 7. Face images on AR database.

## A. Various Feature Transformation and Classification Methods

The first experiment was carried out to evaluate the correct recognition rate of face images without occlusion on the AR and YaleA database utilizing various feature transformation to reduce dimension and classification methods.



Figure 8. Face images on YaleA database



Figure 9. Face images on ORL database

For each individual on the YaleA database, the first six images were selected as the training dataset, and the last five images were selected as the testing dataset. Seven images of 100 individuals in the first session without occlusion on the AR database were selected as the training datasets, and the other seven images in the second session without occlusion were selected as the testing datasets. The feature dimension was set at 90 in the experiment. The PCA, LDA, down-sampling and random method were respectively adopted to reduce the dimension. The SRC, BSRC, nearest neighbor (NN) [2, 17], and LSSRC method proposed in this paper were adopted for classification, where the sizes of the blocks were set at 25\*25 for the BSRC method. The results on



the AR and YaleA database are separately shown in Table 1 and Table 2.

TABLE I. CORRECT RECOGNITION RATES WITHOUT OCCLUSION ON YALEA DATABASE BY VARIOUS FEATURE TRANSFORMATIONS AND CLASSIFIERS

Feature extraction	SRC (%)	BSRC (%)	LSSRC (%)	NN (%)
LDA	91.6	92.0	92.6	85.0
PCA	91.4	91.8	92.7	83.2
Random	91.7	93.1	93.5	77.5
Down-sampling	88.4	91.7	92.0	69.0

TABLE II. CORRECT RECOGNITION RATES WITHOUT OCCLUSION ON AR DATABASE BY VARIOUS FEATURE TRANSFORMATIONS AND CLASSIFIERS

Feature extraction	SRC (%)	BSRC (%)	LSSRC (%)	NN (%)
LDA	85.7	86.1	87.2	76.4
PCA	84.1	87.0	87.4	78.1
Random	85.8	87.8	88.1	69.4
Down-sampling	82.6	86.3	87.5	71.3

When using the same dimensionality reduction method, the recognition rate of the LSSRC method was highest, followed by BSRC, with NN being the lowest, which can be attributed to the dictionary learning based SRC and sub-classifiers fusion. The random sampling matrix method had higher recognition rate compared to the other three feature transformation methods for SRC, BSRC and LSSRC method, which was probably because the transformed matrix of the random method had the lowest correlation coherence degree [2] with the dictionary matrix composed of training samples. Thus, the random method was used in the follow-up experiments for dimensionality reduction.

### B. Various Feature Dimensions and Classification Methods

We tested the impact of different dimensionality reduction modes and classification methods on the AR and YaleA database without occlusion. The training dataset and testing dataset used for the experiment were the same as those used for the first experiment. The random method was used for feature transformation to reduce the dimension to 60, 90, 120, 250, and 500 respectively. The SRC, BSRC, NN, and LSSRC methods were adopted respectively for classification, where the sizes of the blocks were set at 25\*25 for the BSRC method. The results on the YaleA and AR database are separately shown in Table 3 and Table 4.

TABLE III. CORRECT RECOGNITION RATES WITHOUT OCCLUSION ON YALEA FOR VARIOUS FEATURE DIMENSION AND CLASSIFIERS

Feature dimension	SRC (%)	BSRC (%)	LSSRC (%)	NN (%)
60	86.2	86.8	87.2	76.7
90	91.7	93.1	93.5	77.5
120	94.5	95.2	95.8	78.8
250	95.1	96.0	96.2	78.9
500	95.7	97.1	97.7	77.0

It can be seen from the results that the correct recognition rate increased with the increase in feature dimension for the four classification methods. The LSSRC achieved the highest recognition rate of 97.7% on

the YaleA database, which was due to the combination of SRC based ensemble modeling and voting with weight based sub-classifiers fusion. At the same time, the SRC, BSRC, and LSSRC methods achieved a recognition rate above 94.4% on the YaleA database and 87.4% on the AR database when feature dimension was larger than 120. Taking the recognition rate and the cost into consideration, the feature dimension was set at 120 for the subsequent experiment.

TABLE IV. CORRECT RECOGNITION RATES WITHOUT OCCLUSION ON AR FOR VARIOUS FEATURE DIMENSION AND CLASSIFIERS

Feature dimension	SRC (%)	BSRC (%)	LSSRC (%)	NN (%)
60	78.8	81.2	67.1	78.2
90	87.8	88.1	69.4	85.8
120	87.8	88.3	71.3	87.5
250	88.9	89.5	72.7	88.8
500	89.7	90.2	74.2	89.1

### C. Various Levels of Occlusion and Classification Methods

We tested the influence of different random occlusion sizes on face recognition. We selected the first six images from the face database to form a training set, and added different occlusion sizes to the last five images to form a testing set. The image dimension was reduced to 90, and occluded face recognition was conducted using SRC, BSRC (block sizes 25\*25, 25\*50 and 50\*50), NN, and the proposed method. The results on the YaleA, AR and ORL database are separately shown in Table 5, Table 6 and Table 7.

TABLE V. RECOGNITION RATE OF VARIOUS ALGORITHMS ON YALEA DATABASE UNDER DIFFERENT DEGREES OF OCCLUSION (%)

Recognition algorithms	Degree of occlusion			
	15%	25%	35%	50%
NN	76.8	57.6	48.9	39.0
SRC	83.4	77.8	65.9	41.2
BSRC(25*25)	93.9	89.8	78.2	62.6
BSRC(25*50)	92.0	88.5	73.4	57.7
BSRC(50*50)	85.5	81.3	68.4	45.0
LSSRC	97.9	95.8	87.8	75.6

TABLE VI. RECOGNITION RATE OF VARIOUS ALGORITHMS ON AR DATABASE UNDER DIFFERENT DEGREES OF OCCLUSION (%)

Recognition algorithms	Degree of occlusion			
	15%	25%	35%	50%
NN	64.8	51.6	42.9	32.4
SRC	79.4	75.8	64.9	42.2
BSRC(25*25)	90.9	85.4	81.9	65.7
BSRC(25*50)	88.2	83.4	79.8	54.7
BSRC(50*50)	79.0	75.8	68.4	44.3
LSSRC	95.7	93.4	85.2	73.4

TABLE VII. RECOGNITION RATE OF VARIOUS ALGORITHMS ON ORL DATABASE UNDER DIFFERENT DEGREES OF OCCLUSION (%)

Recognition algorithms	Degree of occlusion			
	15%	25%	35%	50%
NN	67.4	54.6	48.9	36.4
SRC	76.4	72.8	61.9	39.2
BSRC(25*25)	87.9	82.8	78.2	62.6
BSRC(25*50)	85.0	80.5	76.4	51.4
BSRC(50*50)	76.5	72.3	65.4	41.0
LSSRC	92.8	91.4	83.2	70.5

With the increase in occlusion, the recognition rates of the four algorithms decreased; of them, the recognition rate of the proposed method decreased at the lowest pace, and its recognition rate was highest under the same occlusion. This was because the proposed method was based on dictionary learning and sub-classifier fusion recognition, with the highest robustness of occlusion. In addition, under the same occlusion, with the increasing number of blocks, the BSRC recognition rate increased. The recognition rate of the proposed method with the largest number of blocks was highest, far more than the recognition rate of SRC and NN without blocks. This was because BSRC and the proposed method established SRC modes based on partial blocks. The more blocks there were, the greater was the detail. Thus, it was possible to obtain additional local information on images and improve the occlusion recognition rate [17].

## V. CONCLUSIONS

The proposed method conducted sub-block dictionary learning for face blocks and carried out weighted fusion recognition for multi-classifiers according to different recognition contributions of sub-classifiers. Results from the AR and YaleA database showed that the occlusion recognition rate of the proposed method was higher than that of SRC and BSRC. The classifier used by the LSSRC method was a voting algorithm with weight, and other classifiers fusion recognition using dictionary learning and image wavelet features will be used for face recognition under occlusion in our future research.

## ACKNOWLEDGEMENTS

This work was supported by the National Nature Science Foundation of China (Grant No. 60973064).

## REFERENCES

- [1] W. Zhao, R. Chellappa, P. J. Phillips, and A. Rosenfeld, "Face recognition: A literature survey", *ACM Computing Surveys*, vol. 35, no. 4, pp. 399-458, 2003.
- [2] Wright J, Yang A, Ganesh A, et al., "Robust face recognition via sparse representation", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 31, no. 2, pp. 210-227, 2009.
- [3] A. Wagner, J. Wright, A. Ganesh, Z. Zhou, and Y. Ma, "Toward a practical face recognition: Robust registration and illumination via sparse representation", *In Proceeding(s) of Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 597-604, 2009.
- [4] D. Donoho, "Compressed Sensing", *IEEE Transactions on Information Theory*, vol. 52, pp.1289-1306, 2006.
- [5] Q. Zhang, B. Li, "Discriminative K-SVD for dictionary learning in face recognition", *Proceeding(s) of Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 2691-2698, 2010.
- [6] Mao X. Nguyen, Quang M. Le, Vu Pham, Trung Tran, Bac H. Le, "Multi-scale Sparse Representation for Robust Face Recognition", *Proceeding(s) of Conference on Knowledge and Systems Engineering (KSE)*, pp. 195-199, 2011.
- [7] Yikui Zhai, Junying Gan, Jingwen Li, "Study of occluded robust face recognition approach based on homotopy algorithm and color information fusion", *Signal Processing*, vol. 21, no. 11, pp.1762-1768, 2011.
- [8] Yang M, Zhang Lei, "Gabor Feature based Sparse Representation for Face Recognition with Gabor Occlusion Dictionary", *Proceedings of the European Conference on Computer Vision, ECCV*, pp. 448-461, 2010.
- [9] Cui Zhen, Shan Shi-guang, Chen Xi-lin, et al., "Sparsely Encoded Local Descriptor for Face Recognition", *Proceedings of Automatic Face & Gesture Recognition and Workshops. FG*, pp. 149-154, 2011.
- [10] Liu Cheng-Lin, "Classifier combination based on confidence transformation", *Pattern Recognition*, vol. 38, no. 1, pp. 11-28, 2005.
- [11] Chen S, Donoho D, Saunders M, "Atomic decomposition by basis pursuit", *SIAM Review*, vol. 43, no. 1, pp. 129-159, 2001.
- [12] Beck A., Teboulle M., "A fast iterative shrinkage-thresholding algorithm for linear inverse problems", *SIAM Journal on Imaging Sciences*, vol. 2, no. 1, pp. 183-202, 2009.
- [13] Belhumeur, P.N., Hespanha, J.P., Kriegman, D.J., "Eigenfaces vs. fisherfaces: Recognition using class specific linear projection", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 19, no. 7, pp.711-720, 1997.
- [14] Wu Di, Cao Jie, Wang Jinhua, Li Wei, "Multi-feature fusion face recognition based on kernel discriminate local preserve projection algorithm under smart environment", *Journal of Computers (Finland)*, vol. 7, no. 10, pp. 2479-2487, 2012.
- [15] Soni, Mohit, Gupta, Phalguni, "A robust vein pattern-based recognition system", *Journal of Computers (Finland)*, vol. 7, no. 11, pp. 2711-2718, 2012.
- [16] Zhang Jing, Wu Gongqing, Hu Xuegang, Li Shiyong, Hao Shuilong, "A parallel clustering algorithm with MPI-Mkmeans", *Journal of Computers (Finland)*, vol. 7, no. 1, pp. 10-17, 2013.
- [17] Anil K. Jain, Stan Z. Li, "Handbook of Face Recognition", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2002.
- [18] Yang M, Zhang L, Zhang D. Efficient misalignment-robust representation for real-time face recognition. *Computer Vision-ECCV 2012. Springer Berlin Heidelberg*, 2012 pp. 850-863.



**Deng Nan** received the MSc degree (2008) in mathematics from NorthWest University, P. R. CHINA. She is now a PhD candidate at University of Science and Technology Beijing. Her current research interests include pattern recognition, image processing, non-Newtonian mechanical system control, artificial intelligence.

**Zheng-guang Xu** received the PhD degree (2000) in control science and engineering from University of Science and Technology Beijing, Beijing, P.R.CHINA. He is now a professor at University of Science and Technology Beijing. His current research interests include pattern recognition, image processing, non-Newtonian mechanical system control, Petri nets.