

Local Robust Sparse Representation for Face Recognition With Single Sample per Person

Jianquan Gu, Haifeng Hu, and Haoxi Li

Abstract—The purpose of this paper is to solve the problem of robust face recognition (FR) with single sample per person (SSPP). In the scenario of FR with SSPP, we present a novel model local robust sparse representation (LRSR) to tackle the problem of query images with various intra-class variations, e.g., expressions, illuminations, and occlusion. FR with SSPP is a very difficult challenge due to lacking of information to predict the possible intra-class variation of the query images. The key idea of the proposed method is to combine a local sparse representation model and a patch-based generic variation dictionary learning model to predict the possible facial intra-class variation of the query images. The experimental results on the AR database, Extended Yale B database, CMU-PIE database and LFW database show that the proposed method is robust to intra-class variations in FR with SSPP, and outperforms the state-of-art approaches.

Index Terms—Dictionary learning, face recognition (FR), illumination changes, single sample per person (SSPP), sparse representation.

I. INTRODUCTION

Face recognition (FR) has been one of the hot fields in computer vision and biometrics, since it is a challengeable work to identify a face image with varying expression, occlusion, disguise, and illumination. Many typical solutions are to collect plenty of training data for compensating the above intra-class variation problems [1]. However, collecting abundant data is hard to guarantee in many situations. In many real-world applications, e.g., driving license, visas, or student card, there may single training image of per subject is available, that leads to the trouble of single sample per person (SSPP) [2]. It is a bottleneck of FR since only limited face image information we obtain to compensate the possible intra-class variations in the query images. How to extract discriminative and robust

features to achieve excellent performance of FR with SSPP is a significant and difficult work.

For FR with SSPP, the generalization ability of the learned classifiers may be seriously reduced for the fundamental reason that the robustness of the extracted features is insufficient. Furthermore, the intra-class variation of the query images is hard to estimate and the traditional discriminative subspace learning approaches [3], [4] fail to work.

The existing solutions for FR with SSPP can be approximately categorized into two types: patch-based methods and generic learning methods.

Patch-based methods [5]–[8] divide face images into several patches, so that the discriminative feature information from all patches is extracted and the corresponding classification results of each patch are integrated to calculate the final recognition result. The algorithms that extracting discriminative feature information from all patches has considered local binary pattern (LBP) [9], manifold learning [6], or Gabor features [10]. In [5], Chen *et al.* learn a within-class matrix by considering each patch of each individual as the distributions of the current class. In [6], Lu *et al.* use the patches from each individual to construct a manifold and maximize the manifold margin to figure out a projection matrix. By obtaining the weak classifiers on all patches, some methods joint the weak classifiers to output the final classification result [7], [8]. In [7], Kumar *et al.* employ the nearest neighbor classifier (NNC) on all patches, and present a kernel plurality to combine the classification results of each patch. In [8], Zhu *et al.* employ the collaborative representation based classifier (CRC) [11] on all patches, and utilize the majority voting for classification. The patch-based algorithms such as [7], [8] enhance the FR performance to a certain degree, but they still fail to settle the matter of lacking the possible face variation information in the training samples. In [12], Zhu *et al.* proposes a local generic representation (LGR) based framework for face recognition by considering the facts that different parts of human faces have different importance. For the above patch-based methods, the classification performance is deteriorated when large intra-class variations exists between the probe images and the gallery images. To solve this problem, we introduce the generic variation data with weights for each patch to alleviate the loss of intra-class variations.

Taking into account that different face images share similarity in possible face variation, generic learning methods encourage employing external data to compensate the insufficient of gallery images in face recognition. The target of these algorithms is to model the intra-class variation [13], [14] and to learn the classifiers with promoted recognition abilities

Manuscript received April 12, 2016; accepted December 20, 2016. This work was supported in part by the National Natural Science Foundation of China (61673402, 61273270, 60802069), the Natural Science Foundation of Guangdong Province (2017A030311029, 2016B010109002, 2015B090912001, 2016B010123005, 2017B090909005), the Science and Technology Program of Guangzhou of China (201704020180, 201604020024), and the Fundamental Research Funds for the Central Universities of China. Recommended by Associate Editor Chengqing Zong. (Corresponding author: Haifeng Hu.)

Citation: J. Q. Gu, H. F. Hu, and H. X. Li, "Local robust sparse representation for face recognition with single sample per person," *IEEE/CAA J. of Autom. Sinica*, vol. 5, no. 2, pp. 547–554, Mar. 2018.

J. Q. Gu, H. F. Hu, and H. X. Li are with the School of Electronics and Information Technology, Sun Yat-sen University (SYSU), Guangzhou 510275, China, and SYSU-CMU Shunde International Joint Research Institute, Shunde 528300, China (e-mail: gujianquan123@163.com; huhaif@mail.sysu.edu.cn; lihaoxi@mail2.sysu.edu.cn).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/JAS.2017.7510658

[15], [16]. For instance, adaptive generic learning (AGL) [15] learned a within-class matrix by employing external data to classify all query images. In [16], Kan *et al.* presented a nonlinear framework to figure out the within-class matrix. The extended sparse representation coding (ESRC) [13] directly utilizes external data to consist of a generic dictionary to compensate the variation in the query images for recognition. In [14], Yang *et al.* link the external data by jointly learning a projection matrix. In [17], Yang *et al.* proposed a sparse representation classification based model by eliminating the occlusion components in the dictionary and using a weighted least absolute shrinkage and selection operator (LASSO) algorithm.

The holistic methods cannot guarantee desire performance on face recognition with the problem of varying illumination, expression, occlusion, and so on. Fortunately, patch-based discriminative feature learning and extraction is sensitive to those face image variations. Abundant intra-class variation feature information in the external data can be taken to represent the unknown variations of a query image for the reason that it brings useful information for discrimination.

Sparse representation based methods are broadly utilized in face recognition. In [18], Fang *et al.* propose to exploit local gabor features with multitask adaptive sparse representation for face recognition. In [19], Li *et al.* propose a customized sparse representation model for undersampled face recognition. He *et al.* propose two models based on sparse representation for face [20], [21]. In [20], a half-quadratic (HQ) framework is proposed by defining different kinds of half-quadratic functions, which aims to performing both error correction and error detection. In [21], a sparse correntropy framework is proposed for computing sparse representations of face recognition. The sparse correntropy framework aims to solve the problems of occlusion and corruption in face recognition which is based on the maximum correntropy criterion and a nonnegativity constraint. Different from He's methods, our method aims to solve the problem of single sample per person by combining a local sparse representation model and a patch-based generic variation dictionary learning model with the weights of each patch. Note that both the above two methods cannot solve the problem of face recognition with single sample per person.

In this paper, a novel model local robust sparse representation (LRSR) is proposed to solve the inter-class variation problems in FR with SSPP. The proposed model combines the patch-based method and dictionary learning method to extract robust feature information from external data. The local face recognition methods are beneficial to extract more image feature information for better tackling the problem of local variation. Learning local generic dictionaries for FR favor integrating the variation information of external data to solve the problems related to various inter-class variations in FR. For compensating the shortage of the face representation information in the training set with single sample per person, the proposed scheme collects a generic variation set from external data. Then we learn a generic variation dictionary from the external data. By employing the patch-based approach, the patches of each query sample are jointly represented by

the local gallery dictionary and the local generic variation dictionary. Using patch-based approach is beneficial to calculate the different weights of each patch of the face which take into account different regions of the human face with different structures. At last, the proposed model performs the classification of the query image according to the sums of the representation residuals of each patch over all subjects. The experimental results demonstrate that LRSR provides the highest recognition rates in FR with SSPP. Fig. 1 shows the flowchart of LRSR. In [22], He *et al.* proposes a two-stage sparse representation (TSR) framework, by decomposing the procedure of face recognition into outlier detection stage and recognition stage. TSR aims to solve the L_1 minimization problem in a low computation expensive. Our method is based on L_2 minimization by using a patch-based learning model, which is different from the above TSR method.

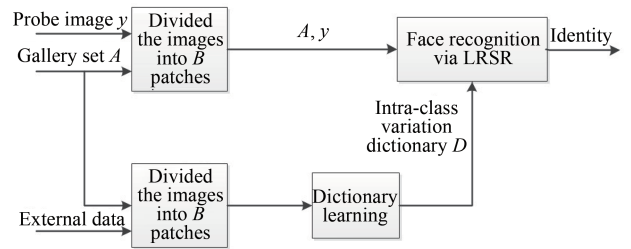


Fig. 1. Flowchart of the proposed framework for FR with SSPP.

By combining patch-based local sparse representation and the patch-based local generic variation dictionary learning from the external data, the proposed method brings three-fold benefits. First, the proposed method provides a novel framework for classifying face images with varying illumination, occlusion and expression variations by feat of the patch-based sparse coding and patch-based generic variation dictionary which extract robust variation features from external data to compensate the intra-class variation information of the interested subjects. Second, the proposed method calculates the class weights of each patch to balance the extremely local variations for restricting the outlier patches which pay an oversize contribution in the overall face classification. Third, the proposed method is able to overcome the difficulties of FR with SSPP, since just only single training image of per individual is required. By using a unified framework to integrate patch-based robust face recognition and local generic dictionary learning, LRSR enhances the performance of FR with SSPP.

The rest of this paper is organized as follows. In Section II, we present the proposed model local robust sparse representation (LRSR) for face recognition with single sample per person. In Section III, we present the experimental results on four benchmark face databases. In Section IV, we summarize the concluding remarks.

II. OUR APPROACH

In the scenario of FR with SSPP, we collect a gallery set $A = [a_1, \dots, a_2, \dots, a_j] \in \mathbb{R}^{d \times j}$, where $a_j \in \mathbb{R}^d$ is the single training sample of subject j , $j = 1, 2, \dots, J$. For a query

image $y \in \mathbb{R}^d$, we use the gallery dictionary A to represent the query image in such a SRC [23] based manner as:

$$y = Ax_a + e \quad (1)$$

where x_a denotes the coding vector of y on the gallery dictionary A , and e is the representation residual. Unfortunately, the majority of possible intra-class variation in the query image y is incapable of denoted by the single training sample from the correct subject. Therefore, the representation residual of y is large, and may lead to misclassification when utilize the representation manner of (1). Taking into account that the intra-class variation in face images share across different individuals, we adopt an external generic variation training dataset to learn a generic variation dictionary. It may lead to noisy for directly applying raw pixels from external data to consist of generic variation dictionary D , such as ESRC. In the proposed method, we support utilizing dictionary learning algorithm to extract representative feature from external data, and [24] has demonstrated that the methods employing dictionary learning algorithms from external data outperforms the algorithm employing predefined ones. The learned dictionaries are more suitable for image denoising and ensure the generalization ability for the interested individuals.

The sparse variation dictionary learning (SVDL) [14] aims to learn a generic variation dictionary by combining the gallery data and the external data. Considering that the SVDL has shown auspiciously performance in recent works of dictionary learning in FR, thus, we apply the SVDL to learn the external generic variation dictionary from external data. The SVDL exploit the inter-class variation feature of the external data by learning a compact dictionary with a projection matrix from the external data to the gallery data. The following is the learning process of the generic variation dictionary D by employing SVDL algorithm. Thus, we collect a generic variation training set $[D^r, D^v]$ from external data, where D^v is the variation subset and D^r is the reference subset. The reference subset $D^r \in \mathbb{R}^{d \times n}$ is constructed by neutral face images of each subject. For each external generic subject, we obtain the reference subject D_i^v and a variation subject, where $D_i^v = [D_1^v, \dots, D_k^v, \dots, D_K^v]$, D_k^v is the subset of the k th variation, $k = 1, 2, \dots, K$. Different gallery subjects may have same or different external generic variation dictionary representation since the generic dictionary is not related to any subjects of interest. For the i th gallery sample, denoted as A_i , $i = 1, 2, \dots, J$, the learning generic variation dictionary model of SVDL can be expressed as

$$\min_D \sum_{i=1}^c p(A_i, D_i^r, r_i) + q(D, D_i^v, r_i) \quad (2)$$

where r_i is the coding vector of A_i over D_i^r , $p(A_i, D_i^r, r_i)$ is the adaptive projection learning term, which exploits the relationship between the gallery subset and reference subset, and $q(D, D_i^v, r_i)$ denotes the variation dictionary learning, which is used for learning a sparse representation based dictionary by using the projection of the face variation from the external data over the projection matrix. D is the generic variation dictionary, which is learned by utilizing this SVDL

model. With the projection r_i , the variation matrix of gallery set can be obtained by projecting the variation matrix D_i^v onto the dictionary. Thus, r_i connect the external data with the gallery data, which guaranteeing the learned generic variation dictionary D is adaptive to the gallery data.

When obtaining the generic variation dictionary D , we represent the query sample y over the dictionary A and D as following:

$$y = Ax_a + Dx_d + e \quad (3)$$

where x_a and x_d is the coding vectors of y over A and D , and e denotes the residual.

Different face regions such as eye, mouth, and nose have different features and it is appropriate to learn region-specific features to obtain the essential information from the local patches in classifying the identification of a query face image. Considering this principle, we utilize a patch-based method on (3) by proposing a patch-based generic representation model to learn robust and discriminatively feature for different face patches individually.

For our proposed patch-based method, the query image y is divided into B patches which are represented as y_1, y_2, \dots, y_B . In the same way, the gallery dictionary A is partitioned as A_1, A_2, \dots, A_B . For each local patch i , we independently learn the local generic variation dictionary D_i by utilizing the SVDL algorithm at the corresponding patch of the external data. For each local patch y_i , $i = 1, 2, \dots, B$, it can be represented by the local gallery dictionary A_i and the local generic variation dictionary D_i , respectively. To explore the most of image information and enhance the representation and generalization ability of local gallery dictionaries of each patch, we extract the feature of the adjoining blocks at each patch of the training sample, and append them to A_i for constructing a dictionary with more gallery information. With A_i and D_i , we individually represent each local patch of query image y as

$$y_i = A_i x_{ia} + D_i x_{id} + e_i, \quad i = 1, 2, \dots, B \quad (4)$$

where x_{ia} and x_{id} are the coding vector of y_i over A_i and D_i , respectively, and e_i is the representation residual corresponding to local patch y_i , as shows in Fig. 2.



Fig. 2. Local sparse representation. The query image is divided into several blocks. The first image is the query image; the second image is recovered by the gallery set; the third is represented by the external data; the last is the representation residual.

It needs to define an appreciate loss function on the representation residual e_i and employ an appropriate regularization on the representation to calculate the optimal solutions of coding vectors x_{ia} and x_{id} . We take the following minimization problem into account

$$\min_x \sum_{i=1}^B (\|y_i - A_i x_{ia} - D_i x_{id}\|_2^2) + \lambda \|x\|_F^2 \quad (5)$$

where $x = x_1, x_2, \dots, x_B$ with $x_i = [x_{ia}; x_{id}]$. The solution of each $x_i = [x_{ia}; x_{id}]$ on (5) can be calculated by using a manner of least square regression, as following

$$[x_{ia}; x_{id}] = ([A_i, D_i]^T [A_i, D_i] + \lambda I)^{-1} [A_i, D_i]^T y_i. \quad (6)$$

When obtaining the optimal coding vectors $\{\tilde{x}_i\}_{i=1}^B$, we calculate the recognition result of query sample y . Let $A_i = [A_i^1, \dots, A_i^j, \dots, A_i^J]$ where A_i^j is the sub-gallery dictionary of i th patch associated with class j . The representation residual of each patch y_i on each class j can be calculated by utilizing the class-specific local gallery dictionary A_i^j and the local generic variation dictionary D_i . The query image can be classified by calculating the minimal sum of the representation residuals on each patch in the following manner

$$IDENTITY(y) = \arg \min_j \sum_{i=1}^B \left\| y_i - [A_i^j, D_i][x_{ia}^j; x_{id}] \right\|_2^2 \quad (7)$$

where x_{ia}^j is the j th coefficient of x_{ia} , which means the effect factor of j th subject in classification of i th patch. For classifying y , if subject j has a larger $\|x_{ia(j)}; x_{id}\|_2^2$, it means the i th patch of query image y maybe more close to the position-specific gallery patch of subject j . However, it also indicates that the classification result of the query image will be controlled by the patches with the extremely situation when $\|x_{ia(j)}; x_{id}\|_2^2$ is too large or too small. To solve this problem, we intend to figure out the class weights of each patch to adjust the classification effect of the representation residual in each patch. For the i th patch, the class weight is to weaken or to enhance the classification effect of the representation residual on each class as following

$$w_i^j = \frac{1}{\|x_{ia(j)}; x_{id}\|_2^2} \quad (8)$$

where x_{id} is a constant vector for each class in the same patch, and different patch with different x_{id} . $x_{ia(j)}$ is a value specific for the subject j in the i th patch. Our classification principle is to find the smallest sum of representation residual over each patch on each class weight w_i^j as the following

$$IDENTITY(y) = \arg \min_j \sum_{i=1}^B w_i^j \left\| y_i - [A_i^j, D_i][x_{ia}^j; x_{id}] \right\|_2^2. \quad (9)$$

When calculating the sums of the weighted representation residuals of each patch over all classes, the query sample y is classified to the class with the minimal sum.

III. EXPERIMENTS

We perform sufficient experiments to demonstrate the robustness of LRSR for FR with SSPP on four benchmark face databases, including AR database [25], Extended Yale B database [26], CMU-PIE database [27], and LFW database [28].

In all our experiments, to guarantee robustness and simplicity, the regularization parameter λ in (5) is fixed as $\lambda = 0.013$ throughout all the experiments. For all images throughout the experiments, we fix the size of each patch as 20×20 , and the

overlapped margin is 10 pixels. The three parameters $\lambda_1, \lambda_2, \lambda_3$ of the adopted local generic variation dictionary method SVDL are set as 0.0002, 0.0008, 0.00001, respectively. The face images on AR database, Extended Yale B database are resized to 80×80 .

We compare LRSR with the state-of-art approaches based on sparse representation including [11], SRC [23], ESRC [13], SVDL [14], robust sparse coding (RSC) [17] and LGR [12] to prove the performance of the proposed method for robust sparse representation in FR with SSPP. Note that method LRSR without using weight calculation (LRSRWW) means the proposed method LRSR without using weight calculation.

A. AR Database

The AR [25] face database includes 126 subjects which contains over 4000 images. For each individual, 13 face images were taken in each of two sessions, covering different expressions, illumination conditions, and occlusions.

In our SSPP experiments, the public available subset with face images consisting of 50 males and 50 females is chosen. In Session I, we choose the only 80 neutral images of the first 80 individuals to construct gallery data, and the remaining samples of these 80 subjects are used as probe data. The other 20 subjects from Session I are used as the external generic variation data. We also apply the same operation to Session II. Fig. 3 lists some examples of AR database.



Fig. 3. The first is the selected training image, and the remainder is four query images with variations of one individual in AR database.

To prove the robustness and generalization ability of LRSR with different external data, we add the experiments on the selected 80 subjects of AR database with the local generic variation dictionary learned from Multi-PIE database rather than the AR database. In these additional experiments, the external data is constructed by 20 subjects that are randomly selected from the Session I of Multi-PIE database, and all subjects have 40 face images from the POSE05_1, which include neutral face and a kind of expression variation and 20 kinds of illumination variations. Learning the local generic variation dictionary by using Multi-PIE database [29] makes the classification problem more challenging.

Table I lists the compared performance of FR with SSPP on the AR database with different external data over different algorithms. Fig. 4 demonstrates different performance on Session I of seven algorithms influenced by different number of dictionary dimensions on different external data, where the dimensionality is reduced by principal component analysis (PCA) [4]. It can be seen that LRSR is more stable under different dictionary dimensions compared with other algorithms. Fig. 5 demonstrates different performance of seven methods on Session II under different dictionary dimensions, where the dimensionality is reduced by PCA. From these

tables and figures, extensive experiments on AR database with different external data prove that LRSR outperforms the state-of-art methods. Although learning the local generic variation dictionary by employing Multi-PIE database, LRSR is also robust to the intra-class variations in the AR database. When using Multi-PIE database as external data, there is no occlusion information provide to predict the occlusion variations, and LRSR made desire performance by calculating the class weights of each patch on each class to classify the query image. It is notable that the performance of ESRC reduced singularly when the external data was chosen from Multi-PIE database rather than the AR database. This can be accounted for the reason that the generic variation dictionary in ESRC is directly constructed by external data without learning process. The Multi-PIE database does not contain any occlusion information which leading to the generic intra-class dictionary cannot compensate the occlusion information of query images. Note that the proposed method would not suffer from this problem.

TABLE I
PERFORMANCE COMPARISONS FOR FR WITH SSPP ON THE AR DATABASE, NOTE THAT THE GENERIC TRAINING DATA OF AR1 + MULTI-PIE AND AR2 + MULTI-PIE IS FROM MULTI PIE DATABASE (%)

Method	AR1	AR1 + Multi-PIE	AR2	AR2 + Multi-PIE
SRC	60.62	60.62	41.46	41.46
CRC	60.42	60.42	41.15	41.15
RSC	74.24	74.24	57.19	57.19
ESRC	70.10	64.90	60.35	45.31
SVDL	80.21	65.42	60.62	47.60
LGR	97.81	93.42	90.73	85.10
LRSRWW	85.63	76.15	75.21	66.15
LRSR	98.12	96.88	90.04	88.23

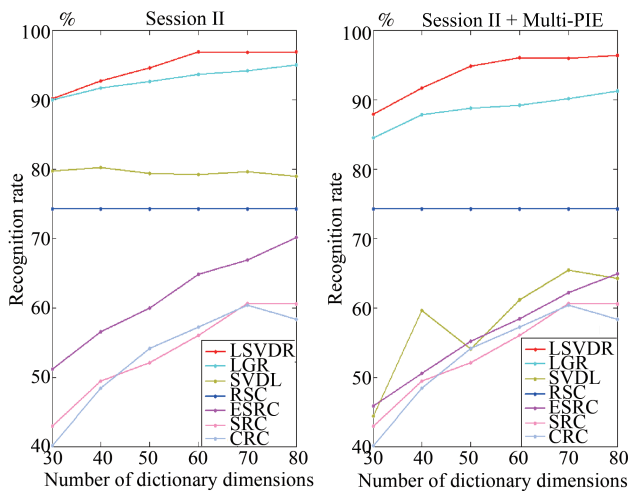


Fig. 4. Performance comparisons on the Session I of AR database with different dictionary dimensions. The left figure shows the local generic variation learned from AR database, and the right figure shows the performance with the local generic variation dictionary learned from Multi-PIE.

It is noteworthy that in Table I, the performance of ESRC is

worse than that reported in [13]. In our experimental settings, we use PCA to reduce the dimension of input features to 80 for further processing. Such dimension reduction process may lose some useful discriminant information for face recognition. This may be the reason of the different classification performance between our experiment and [13]. This further reflects the fact that the proposed method is more robust to the dimension reduction.

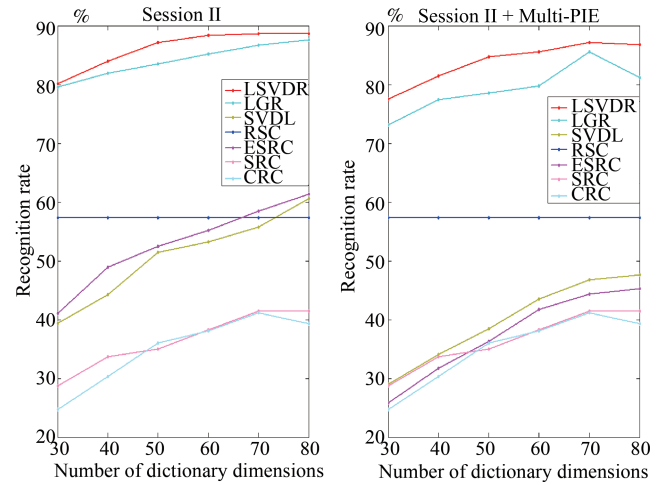


Fig. 5. Performance comparisons on the Session II of AR database with different dictionary dimensions. The left figure shows the local generic variation learned from AR database, and the right figure shows the performance with local generic variation dictionary learned from Multi-PIE database.

B. Extended Yale B Database

The Extended Yale B database [26] includes 38 individuals with 2434 frontal images. They were taken under 64 illumination conditions. Fig.6 lists some samples of the Extended Yale B database. The first 30 subjects from Extended Yale B database are used for training and testing, and the remaining 8 subjects are used external data. For each subject of interests, the single image under the illumination condition: P00A + 000E + 00 are used as the training sample, and the remainder 63 images are used for testing.

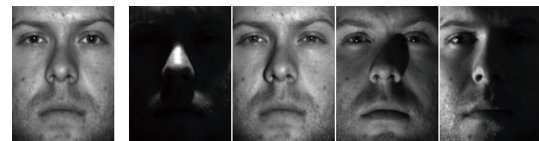


Fig. 6. The first is the selected training image, and the remainder is four query images with variations of one individual in Extended Yale B database.

Table II lists the superior performance of the compared methods on Extended Yale B database. Fig.7 shows the recognition results of various approaches in the condition with various training samples per subject, and the training samples of per subject is selected from the first several ones of each subject. With the augment of the training sample, the proposed method is still keeping a best recognition rate. The proposed method provides the highest performance for fully utilizing the robust feature of each patch and the possible

face variation information from external data. Note Extended Yale B database contains some samples taken under extremely illumination conditions. Therefore, this experiment shows the strong robustness of LRSR to illumination problem.

TABLE II
PERFORMANCE COMPARISONS FOR FR WITH SSPP ON
THE EXTENDED YALE B DATABASE (%)

Method	Extended Yale B
SRC	30.48
CRC	44.18
RSC	30.48
ESRC	56.93
SVDL	59.21
LGR	85.88
LRSRWW	92.37
LRSR	94.13

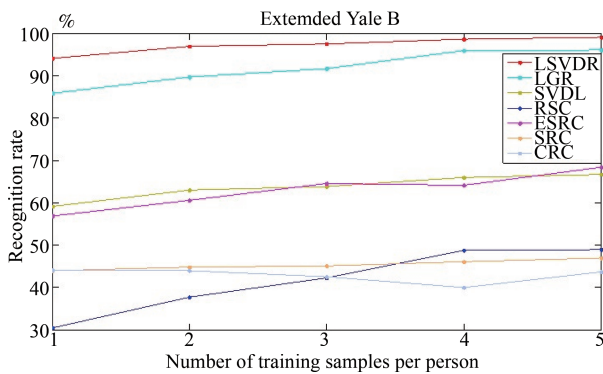


Fig. 7. Performance comparisons on Extended Yale B database with different number of training images per subject.

C. CMU-PIE Database

The CMU-PIE database [27] includes 68 individuals with 41 368 images, which contain intra-class variations including illumination, pose, and expression. Fig. 8 shows some samples of the CMU-PIE database. In the experiments, three subsets in CMU-PIE were selected, including: (C07 subset) where there is little pose variation, which contains 68 individuals and per individual with 24 samples; (C09 subset) where there is little pose variation, which contains 68 individuals and per individual with 24 samples; (C27 subset) where is the front faces, which contains 68 individuals and per individual with 49 samples. For these three subsets, the first 50 individuals are used for training and testing, and the remaining 18 individuals are used as external data to learn the local generic variation dictionary. For the C09 and C07 subsets, the 13th image of each individual are used as training sample and the remaining 23 images are used for testing. For each individual in the C27 subset, we take the 31th image as training sample and the rest 48 images as testing samples.

To prove the robustness of LRSR, we expand the experiments on CMU-PIE database by employing Extended Yale B database to learn the local generic intra-class dictionary.

20 individuals in Extended Yale B database were randomly picked out to construct the external data. Learning the local generic variation dictionary by using Extended Yale B database makes the classification problem more challenging.



Fig. 8. The first is the selected training image, and the remainder is four query images with variations of one individual in CMU-PIE database.

The comparison results of different algorithms on three subsets of CMU-PIE are listed in Table III, Table IV, and Table V respectively. These tables demonstrate that LRSR outperforms other algorithms and provides the highest performance. By using Extended Yale B database to learn local generic variation dictionary, LRSR still keeps the best performance and the recognition results do not declined much. The experiments on CMU-PIE database prove that LRSR is robust to variations with varying illumination, pose and expression.

TABLE III
PERFORMANCE COMPARISONS FOR FR WITH SSPP ON
C07 SUBSET OF CMU-PIE DATABASE, NOTE THAT
THE GENERIC TRAINING DATA OF THE C07 + EYB IS
FROM EXTENDED YALE B DATABASE (%)

Method	C07	C07 + EYB
SRC	68.09	68.09
CRC	70.79	70.79
RSC	68.48	68.48
ESRC	90.59	86.61
SVDL	83.96	77.50
LGR	92.24	89.45
LRSRWW	91.63	88.43
LRSR	93.46	91.01

TABLE IV
PERFORMANCE COMPARISONS FOR FR WITH SSPP ON
C09 SUBSET OF CMU-PIE DATABASE, NOTE THAT
THE GENERIC TRAINING DATA OF THE C09 + EYB IS
FROM EXTENDED YALE B DATABASE (%)

Method	C09	C09 + EYB
SRC	65.48	65.48
CRC	67.42	67.42
RSC	68.00	68.00
ESRC	88.61	88.40
SVDL	83.48	77.65
LGR	93.13	91.65
LRSRWW	91.39	88.26
LRSR	93.63	91.74

D. LFW Database

The LFW database [28] includes more than 13 000 face images with 5749 individuals which are captured from web

with unrestrict environment variations in pose, expression, illumination, and so on. A subset of 100 individuals with 14 samples per person from LFW database is selected. The selected images are cropped and resized to 80×80 , and aligned by [30]. For these subjects, we choose 80 subjects for training and testing, and the remaining 20 subjects are used to learn the local generic variation dictionary. We choose an image closed to neural face to construct gallery set in these 14 samples. Table VI shows the classification result of different methods in LFW database. It can be seen that the performance of the proposed method is better than the compared methods, but still can not achieve a promising performance. The reason can be explained by the fact that variation in LFW is yet very deteriorated compared with the images in the controlled environment although face alignment has been conducted.

TABLE V
PERFORMANCE COMPARISONS FOR FR WITH SSPP ON
C27 SUBSET OF CMU-PIE DATABASE, NOTE THAT
THE GENERIC TRAINING DATA OF THE C27 + EYB IS
FROM EXTENDED YALE B DATABASE (%)

Method	C27	C27+EYB
SRC	74.38	74.38
CRC	74.26	74.26
RSC	55.02	55.02
ESRC	86.69	85.27
SVDL	85.65	78.64
LGR	89.15	84.23
LRSRWW	89.07	87.28
LRSR	91.07	90.45

TABLE VI
PERFORMANCE COMPARISONS FOR FR WITH SSPP ON
THE LFW DATABASE (%)

Method	LFW database
SRC	37.12
CRC	25.19
RSC	29.81
ESRC	43.46
SVDL	30.58
LGR	49.52
LRSRWW	47.02
LRSR	50.10

IV. CONCLUSION

We propose LRSR framework that fully extracts the robust feature information of gallery set and the possible face variation information of external data. LRSR provides highest performance by combining the benefits of local sparse representation and local generic variation dictionary learning. By calculating the class weight of each patch on all classes, LRSR reduces the adverse effect of local extremely awful intra-class variations. The sufficient experiments performed on the AR database and Extended Yale B database demonstrate that

LRSR achieves the higher recognition performance compared with the state-of-the-art SSPP approaches.

REFERENCES

- [1] H. T. Zhao and P. C. Yuen, "Incremental linear discriminant analysis for face recognition," *IEEE Trans. Syst. Man Cybernet. B*, vol. 38, no. 1, pp. 210–221, Feb. 2008.
- [2] X. Tan, S. Chen, Z.-H. Zhou, and F. Zhang, "Face recognition from a single image per person: A survey," *Pattern Recognit.*, vol. 39, no. 9, pp. 1725–1745, Sep. 2006.
- [3] S. Z. Li and J. W. Lu, "Face recognition using the nearest feature line method," *IEEE Trans. Neural Netw.*, vol. 10, no. 2, pp. 439–443, Mar. 1999.
- [4] M. Turk and A. Pentland, "Eigenfaces for recognition," *J. Cognit. Neurosci.*, vol. 3, no. 1, pp. 71–86, 1991.
- [5] S. C. Chen, J. Liu, and Z. H. Zhou, "Making FLDA applicable to face recognition with one sample per person," *Pattern Recognit.*, vol. 37, no. 7, pp. 1553–1555, Jul. 2004.
- [6] J. W. Lu, Y. P. Tan, and G. Wang, "Discriminative multimanifold analysis for face recognition from a single training sample per person," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 1, pp. 39–51, Jan. 2013.
- [7] R. Kumar, A. Banerjee, B. C. Vemuri, and H. Pfister, "Maximizing all margins: Pushing face recognition with kernel plurality," in *Proc. 2011 IEEE Int. Conf. Computer Vision (ICCV)*, Barcelona, Spain, pp. 2375–2382, 2011.
- [8] P. F. Zhu, L. Zhang, Q. H. Hu, and S. C. K. Shiu, "Multi-scale patch based collaborative representation for face recognition with margin distribution optimization," in *Proc. 12th European Conf. Computer Vision-ECCV 2012*, Berlin Heidelberg, Germany, pp. 822–835, 2012.
- [9] T. Ahonen, A. Hadid, and M. Pietikainen, "Face description with local binary patterns: Application to face recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 28, no. 12, pp. 2037–2041, Dec. 2006.
- [10] J. Zou, Q. Ji, and G. Nagy, "A comparative study of local matching approach for face recognition," *IEEE Trans. Image Process.*, vol. 16, no. 10, pp. 2617–2628, Oct. 2007.
- [11] L. Zhang, M. Yang, and X. C. Feng, "Sparse representation or collaborative representation: Which helps face recognition," in *Proc. 2011 IEEE Int. Conf. Computer Vision (ICCV)*, Barcelona, Spain, pp. 471–478, 2011.
- [12] P. F. Zhu, M. Yang, L. Zhang, and I. Y. Lee, "Local generic representation for face recognition with single sample per person," in *Proc. Asian Conf. Computer Vision Computer on Vision-ACCV 2014*, Switzerland, pp. 34–50, 2014.
- [13] W. H. Deng, J. N. Hu, and J. Guo, "Extended SRC: Undersampled face recognition via intraclass variant dictionary," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 34, no. 9, pp. 1864–1870, Sep. 2012.
- [14] M. Yang, L. Van Gool, and L. Zhang, "Sparse variation dictionary learning for face recognition with a single training sample per person," in *Proc. 2013 IEEE Int. Conf. Computer Vision*, Sydney, Australia, pp. 689–696, 2013.
- [15] Y. Su, S. G. Shan, X. L. Chen, and W. Gao, "Adaptive generic learning for face recognition from a single sample per person," in *Proc. 2010 IEEE Conf. Computer Vision and Pattern Recognition (CVPR)*, San Francisco, USA, pp. 2699–2706, 2010.
- [16] M. N. Kan, S. G. Shan, Y. Su, D. Xu, and X. L. Chen, "Adaptive discriminant learning for face recognition," *Pattern Recognit.*, vol. 46, no. 9, pp. 2497–2509, Sep. 2013.
- [17] M. Yang, L. Zhang, J. Yang, and D. Zhang, "Robust sparse coding for face recognition," in *Proc. 2011 IEEE Conf. Computer Vision and Pattern Recognition (CVPR)*, Providence, RI, USA, pp. 625–632, 2011.
- [18] L. Y. Fang and S. T. Li, "Face recognition by exploiting local Gabor features with multitask adaptive sparse representation," *IEEE Trans. Instrum. Measur.*, vol. 64, no. 10, pp. 2605–2615, Oct. 2015.

- [19] Z. M. Li, Z. H. Huang, and K. Shang, "A customized sparse representation model with mixed norm for undersampled face recognition," *IEEE Trans. Inform. Forens. Secur.*, vol. 11, no. 10, pp. 2203–2214, Oct. 2016.
- [20] R. He, W. S. Zheng, T. N. Tan, and Z. N. Sun, "Half-quadratic-based iterative minimization for robust sparse representation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 36, no. 2, pp. 261–275, Feb. 2014.
- [21] R. He, W. S. Zheng, and B. G. Hu, "Maximum correntropy criterion for robust face recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 33, no. 8, pp. 1561–1576, Aug. 2011.
- [22] R. He, W. S. Zheng, B. G. Hu, and X. W. Kong, "Two-stage nonnegative sparse representation for large-scale face recognition," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 24, no. 1, pp. 35–46, Jan. 2013.
- [23] J. Wright, A. Y. Yang, A. Ganesh, S. S. Sastry, and Y. Ma, "Robust face recognition via sparse representation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 31, no. 2, pp. 210–227, Feb. 2009.
- [24] R. Rubinstein, A. M. Bruckstein, and M. Elad, "Dictionaries for sparse representation modeling," *Proc. IEEE*, vol. 98, no. 6, pp. 1045–1057, Jun. 2010.
- [25] A. Martínez and R. Benavente, "The AR face database," Centre de Visió per Comp., Univ. Autònoma de Barcelona, Spain, CVC Tech. Rep. #24, Jun. 1998.
- [26] A. S. Georghiadis, P. N. Belhumeur, and D. J. Kriegman, "From few to many: generative models for recognition under variable pose and illumination," in *Proc. 4th IEEE Int. Conf. Automatic Face and Gesture Recognition*, Grenoble, France, pp. 277–284, 2000.
- [27] T. Sim, S. Baker, and M. Bsat, "The CMU pose, illumination, and expression (PIE) database," in *Proc. 5th IEEE Int. Conf. Automatic Face and Gesture Recognition*, Washington, DC, USA, pp. 46–51, 2002.
- [28] G. B. Huang, M. Ramesh, T. Berg, and E. Learned-Miller, "Labeled faces in the wild: a database for studying face recognition in unconstrained environments," Univ. Massachusetts, USA, Tech. Rep. 07–49, 2007.
- [29] R. Gross, I. Matthews, J. Cohn, T. Kanade, and S. Baker, "Multi-pie," *Image Vision Comp.*, vol. 28, no. 5, pp. 807–813, May 2010.
- [30] T. Hassner, S. Harel, E. Paz, and R. Enbar, "Effective face frontalization

in unconstrained images," in *Proc. 2015 IEEE Conf. Computer Vision and Pattern Recognition*, Boston, MA, USA, pp. 4295–4304, 2015.



Jianquan Gu is currently a master student at the School of Electronics and Information Technology, Sun Yat-sen University. His research interests include computer vision, pattern recognition, deep learning, and digital image processing.



Haifeng Hu received the Ph.D. degree from Sun Yat-sen University (SYSU) in 2004. He is currently an Associate Professor at SYSU. He has published over 30 scientific papers in international journals and conferences on image processing and pattern recognition. His research interests include digital image processing, pattern recognition, and computer vision.



Haoxi Li is currently working toward the master degree at the School of Electronics and Information Technology, Sun Yat-sen University in Guangdong, China. His research interests include computer vision, metric analysis, and cross-age face recognition.