Integrating Conventional and Inverse Representation for Face Recognition

Yong Xu, Member, IEEE, Xuelong Li, Fellow, IEEE, Jian Yang, Member, IEEE, Zhizhi Lai, and David Zhang, Fellow, IEEE

Abstract—Representation-based classification methods are all constructed on the basis of the conventional representation, which first expresses the test sample as a linear combination of the training samples and then exploits the deviation between the test sample and the expression result of every class to perform classification. However, this deviation does not always well reflect the difference between the test sample and each class. With this paper, we propose a novel representation-based classification method for face recognition. This method integrates conventional and the inverse representation-based classification for better recognizing the face. It first produces conventional representation of the test sample, i.e., uses a linear combination of the training samples to represent the test sample. Then it obtains the inverse representation, i.e., provides an approximation representation of each training sample of a subject by exploiting the test sample and training samples of the other subjects. Finally, the proposed method exploits the conventional and inverse representation to generate two kinds of scores of the test sample with respect to each class and combines them to recognize the face. The paper shows the theoretical foundation and rationale of the proposed method. Moreover, this paper for the first time shows that a basic nature of the human face, i.e., the symmetry of the face can be exploited to generate new training and test samples. As these new samples really reflect some possible appearance of the face, the use of them will enable us to obtain higher accuracy. The experiments show that the proposed conventional and inverse representation-based linear regression classification (CIRLRC), an improvement to linear regression classification (LRC), can obtain very high accuracy and greatly outperforms the naive LRC and other state-of-the-art conventional representation based face recognition methods. The accuracy of CIRLRC can be 10% greater than that of LRC.

Index Terms—Face recognition, pattern recognition, representation-based classification.

I. INTRODUCTION

FACE recognition is an important branch of biometrics [1]–[6] and in this branch, representation-based classification methods (RBCM) have attracted much attention. In recent years, various RBMs have been proposed [7]–[11]. In RBCM, both the algorithm to obtain the representation of the test sample and the classification rule are very beneficial to evaluate the dissimilarity between the test sample and each class. Among RBMs, the sparse representation classification (SRC) was almost the earliest proposed one [7], [8]. SRC tries to obtain an optimal linear combination of the training samples, which has the minimum deviation from the test sample. We refer to this linear combination as optimal linear combination to classify the test sample. SRC requires that the $l_1$ norm of the coefficient vector of the linear combination be as small as possible. We refer to this kind of method as $l_1$ norm based representation method or naive sparse representation method (NSRM). Researchers also proposed $l_2$ norm based representation methods such as the collaborative representation classification (CRC) method [12] and the two-stage test sample sparse representation (TPTSSR) method [13]. These methods also obtain the optimal linear combination of the training samples to represent the test sample first, and then exploit this linear combination to classify the test sample. However, they differ from the NSRM as follows [14]–[21]. First, they require that the $l_2$ norm rather than $l_1$ norm of the coefficient vector of the linear combination be minimized. Second, they have analytic solutions whereas NSRM have no such solutions. The $l_2$ norm based representation methods show good performance in biometrics such as face recognition and palmprint recognition [14]–[21]. Linear regression classification (LRC) can be viewed as a $l_2$ norm based representation method [22]–[24]. However, as LRC and conventional $l_2$ norm based representation methods respectively use a linear combination of the training samples from each class and all the training samples to represent the test sample, they should respectively...
solve $C$ and one linear systems for classifying the test sample. $C$ is the number of the classes. The recently proposed kernel RBCM is a nonlinear extension of RBCM [19], [25]–[27]. The $l_1$ norm based representation method has also been extended to the complex space [21], and the corresponding method performs very well in bimodal biometrics. Moreover, the way to obtain the sparse representation has been also applied to improve other methods [28]–[30].

It should be pointed out that previous RBCMs are not perfect. For example, NSRM is usually computationally inefficient. Moreover, when implementing NSRM, we should properly set the values of the parameters. Since LRC depends on the deviation between the test sample and the optimal linear combination of the training samples of each class to classify the test sample, LRC might suffer from the problem shown in Section II-B. It has been shown that the $l_p$ norm with $p \leq 1$ might be a better measurement of sparsity in sparse representation and can perform very well in some cases [31]. However, sparse representation with this norm is a non-convex problem and its solution is only locally optimal.

All previous RBCMs first use a linear combination of the training samples to express the test sample, and then exploit the expression result to perform classification. We say that RBCM is based on conventional representation, which exploits the training samples to obtain a representation of the test sample. It seems that as the same face has varying images owing to changing pose, facial expression, and illumination, this conventional representation means is not very competent for evaluating the dissimilarity between the test sample and each class. Moreover, we identify that the inverse representation is also helpful for evaluating the dissimilarity between the test sample and every class. The inverse representation aims at obtaining a representation of the training sample by exploiting the test sample. Specifically, it expresses an individual training sample of a class by using the optimal linear combination of the test sample and training samples of the other classes.

In this paper, we propose to integrate conventional representation and the inverse representation to perform face recognition. The proposed method indeed uses two complementary ways to reflect and evaluate the dissimilarity between the test sample and each class. A reasonable score level fusion strategy is used to combine the scores generated from the two ways and classify the test sample. This enables us to obtain higher classification accuracy.

As the face image might have a severe variation with the various poses, facial expression and illumination and more training samples can better reflect these variations (i.e., provide more information of the face). The accuracy of face recognition is not only directly related with the classification method but also greatly influenced by the number of the training samples. If there are many training samples and they can sufficiently show possible poses, facial expression, and illuminations, we can usually obtain high accuracy. Nevertheless, if there are only a very small number of training samples, it will be very hard for the real-world applications to obtain a good face recognition performance.

In order to overcome the problem of insufficient training samples, previous literatures have proposed some ways to generate new face images from the true face images. For instance, Ryu et al. [32] exploited the distribution of the given training set to generate virtual training samples for face recognition. Sharma et al. [33] first obtained multiple virtual views of a person under different poses and illumination from a single face image, and then used the extended training samples to recognize the face. Beymer et al. [34] and Vetter et al. [35] also generated new samples with virtual views for face recognition. Based on prototype faces and the optic flow, Tang et al. [36] obtained virtual facial expression. Jung et al. [37] proposed to use the noise to produce new samples. Thin et al. [38] used simple geometric transformations to generate virtual samples. In addition, the synthesized virtual samples were also used to address the one training sample issue [34], [35], [39].

Previous studies have also shown that both the facial structure and expression are symmetrical [40]–[42]. Moreover, as a basic nature of the face, the symmetry of the face has been successfully applied to face detection [43]–[45]. However, we see that in real-world face recognition applications, most of face images are not symmetrical images. The main reason is that most of the faces do not have a strictly frontal and neutral pose when they are imaged. This motivates us to exploit the symmetry of the face to generate virtual samples, i.e., virtual face images. In this paper, we view the mirror image of the face image as new samples (i.e., virtual samples), and use the original and new samples to perform face recognition. We do this because the mirror image really reflects some possible change of the original face image caused by varying pose and illumination (for detail, please refer to Section IV). The experimental results on a number of face databases including the corrupted face images show that the proposed method can outperform the state-of-the-art face recognition methods.

This paper has the following main contributions. First, it proposes the inverse representation and integrates the inverse representation with LRC for face recognition. Second, it demonstrates the theoretical rationale to combine the inverse representation and conventional representation. Third, it takes into account the fact that the face itself has a symmetrical structure and the face image is usually not a symmetrical image. Based on this fact, the paper very reasonably uses the mirror image of the face image to reflect some possible pose and illumination change of the original face image. Fourth, the proposed method obtains very high accuracy and is quite robust to the noise.

The remainder of this paper is organized as follows. Section II presents LRC and its potential drawback. Section III describes the proposed method. Section IV presents the interpretation of the main steps of the proposed method. Section V shows the experimental results. Section VI offers the conclusion.

II. ANALYSIS OF LRC

A. Presentation of LRC [22]

In this subsection we present the algorithm of LRC as follows. We assume that there are $C$ classes and each class has $n$ training samples. Let $x_1, \ldots, x_n$ be all the $N$ training samples. Let $P_i$ be the probability that the $i^{th}$ training sample is from class $i$, where $i = 1, 2, \ldots, C$. The score of the test sample $s$ from class $i$ is given by:

$$s_i = \sum_{j=1}^{n} x_j \cdot y_j$$

where $x_j$ is the $j^{th}$ training sample, $y_j$ is the corresponding label, and $y_j$ is the inverse representation of the training sample.

The final score is given by:

$$s = \max_{i=1}^{C} s_i$$

The class with the highest score is the predicted class.

The remainder of this paper is organized as follows. Section II presents LRC and its potential drawback. Section III describes the proposed method. Section IV presents the interpretation of the main steps of the proposed method. Section V shows the experimental results. Section VI offers the conclusion.
samples \((N = Cn)\), \(x_{Cn+1:k}\) stands for the \(k\)-th training sample of the \(i\)-th subject. LRC should solve \(C\) equations each corresponding to one class. \(C\) is the number of all the classes.

The equation of the \(i\)-th class is

\[
y = X_i A_i
\]

where \(A_i = [a_i^1 \cdots a_i^d]\), \(X_i = [x_{i1} \cdots x_{in}]\). Column vector \(y\) denotes the test sample \(x_{i1}, \cdots, x_{in}\) are also column vectors. For image-based applications such as face recognition, the column vector is obtained by concatenating the rows of the image matrix. Equation (1) is solved using

\[
\hat{A}_i = (X_i^T X_i)^{-1} X_i^T y.
\]

The deviation between the \(i\)-th class and the test sample is defined as \(d_i = \|y - X_i \hat{A}_i\|\). If \(k = \arg \min d_i\), then the test sample is assigned to the \(k\)-th class.

### B. Potential Drawback of LRC

In this subsection we will show that under the condition that each class has a very large number of training samples, LRC might perform badly in classifying the test sample.

For simplicity of presentation, we refer to the following proposition as proposition 1: the training samples of every class form a linearly independent set. The following theorem as proposition 2: the training samples of every class is the same as the dimensionality of the sample, i.e., the number of training samples of every class is close to the dimensionality of the sample is very easy to satisfy.

**Theorem 1.** If proposition 1 and proposition 2 are simultaneously satisfied, then for an arbitrary class there exists a linear combination of the training samples of this class that can accurately express the test sample.

**Proof:** If proposition 1 and proposition 2 are simultaneously satisfied, then the training samples of an arbitrary class will form a maximum linear independence group. As a result, for an arbitrary class, there will exist a linear combination of the training samples that can accurately express the test sample. In other words, the deviation between the test sample and this linear combination of the training samples will have the same value, i.e., zero.

The formal proof is as follows. When proposition 1 and proposition 2 are simultaneously satisfied, the solution of (1) is \(\hat{A}_i = X_i^{-1} y\). As a result, \(d_i = \|y - X_i \hat{A}_i\| = 0\), \(i = 1, 2, \cdots, C\).

Theorem 1 indeed tells us that if in real-world applications proposition 1 and proposition 2 are simultaneously satisfied, the deviation of each class from the test sample will always be the same value, i.e., zero. Thus, it seems that LRC will be almost not able to identify the class to which the test sample truly belongs and it is hard to correctly classify the test sample. It will also lead to very low classification accuracy. Though some numerical computation techniques might be used to somewhat alleviate this problem, they still cannot obtain a great improvement in recognition accuracy. Theorem 1 also somewhat implies that LRC is not quite suitable for the pattern classification problem with very low-dimensional samples. The main reason is that in these kind of problems, the condition that the number of training samples of every class is close to the dimensionality of the sample is very easy to satisfy.

### III. DESCRIPTION OF PROPOSED METHOD

In this section we describe the main steps of the proposed method in detail. We assume that there are \(C\) classes and each class has \(n\) training samples. Let \(x_1, \cdots, x_N\) be all the \(N\) training samples \((N = Cn)\), \(x_{Cn+1:k}\) stands for the \(k\)-th training sample of the \(i\)-th class. Let \(y\) still stand for the test sample.

The main steps of the proposed method are as follows.

1. First the step uses the mirror image of the face image to obtain virtual training samples and test samples. The second step obtains the optimal linear combination of the original and virtual training samples from every class to represent the test sample, and calculates the score of each class. The third step uses the following procedure to implement the inverse representation.

2. The first step uses the mirror image of the face image to obtain virtual training samples and test samples. The second step takes the mean of all the distances corresponding to all the training samples from the \(j\)-th class as the distance between the test sample and the \(j\)-th class. The fourth step integrates the scores and distances, respectively produced from the second and third steps, for the ultimate classification. These steps are presented in detail below.

#### Step 1. Produce the virtual test sample for the original test sample.

For original test sample \(\tilde{y}\) in the form of face image, the virtual test sample is defined as \(y(p, q) = \tilde{y}(p, Q - q + 1)\), \(p = 1, \cdots, Q\) and \(Q\) stand for the numbers of the rows and columns of the face image matrix, respectively. The relationship between \(\tilde{y}\) and \(y\) is that column vector \(y\) is obtained by concatenating the rows of \(\tilde{y}\) in sequence. \(y(p, q)\) and \(y'(p, q)\) denote the pixels located in the \(p\)-th row and \(q\)-th column of \(\tilde{y}\) and \(y\), respectively. \(y\) is indeed the mirror image of \(\tilde{y}\). The virtual training sample corresponding to each original training sample is also generated in the same way. Let column vectors \(x_1^1, \cdots, x_N^1\) be the virtual training samples generated from original training samples \(x_1, \cdots, x_N\) in the form of face images, respectively. \(x_1^1, \cdots, x_N^1\) are obtained by concatenating the rows of \(x_1, \cdots, x_N\), respectively.

Now we convert \(y\) and \(x_1^1, \cdots, x_N^1\) into column vectors and hereafter we still use the same denotations to respectively stand for them. Hereafter we refer to \(y\) and \(x_1, \cdots, x_N\) as naive test sample and training samples, respectively.

#### Step 2. Let \(X\) be:

\[
X = [x_1 \cdots x_N \ x_1^1 \cdots x_N^1 \ x_{Cn+1} \cdots x_{C(n-1)+1} \ x_N \ x_N^1 \ x_{Cn+1} \cdots x_{C(n-1)+1} \ x_N],(N = Cn).
\]
The proposed method has two main contributions. The first contribution is that it for the first time directly uses the symmetry of the face to obtain virtual face images. The second contribution is that it not only exploits the conventional representation means for classification but also provides the inverse representation, i.e., the optimal linear combination of the naive training samples, naive and virtual test samples to represent a training sample. The above two representations are reasonable and beneficial to face recognition owing to the following factors. First, the symmetry of the face is a basic nature of the face itself. However, in real-world applications most of the faces do not have a strictly frontal and neutral pose when they are imaged. Thus, the face image is almost not a symmetrical image and the generated virtual face image obtained using our method will reflect possible variation of the pose. Moreover, the virtual face image completely seems to be a natural face image. This is illustrated by Figs. 1–3. The virtual face image also partially reflects possible variation of the illumination. For example, if in the naive sample the left face has a stronger illumination than the right face then in the virtual sample the right face will have a stronger illumination. Thus, both the naive and virtual training samples are helpful for recognizing the test sample, which usually contains pose and illumination different from those of the naive training sample of the same face. In other words, the combination of the naive and virtual training samples reflect more possible variation of the face image and enables the sample to be more correctly classified. Secondly, in the second step of the proposed method, the simultaneous use of the naive and virtual training samples will enable the representation result to be more useful for classifying the face. In particular, since in the second step every class has more training samples than previously proposed LRC, this step will allow the class result to be more useful for classifying the face. As a result, the second step will be more beneficial to classification than previous LRC. Third, since the representation scheme of the third step is very different from that of the second step, the result of the inverse representation obtained using the third step.
is complementary for the conventional representation obtained using the second step. Moreover, it seems that the second step mainly aims at independently determining the capability, of the naive and virtual training samples of each class, to represent the test sample. In other words, it establishes a linear system for every class and exploits it to evaluate this capability. However, the third step determines the capability, of the naive and virtual test samples, to represent the naive training samples of each class in a competitive way. In other words, for a naive training sample from a subject, it first constructs a linear combination of the naive and virtual test samples as well as the naive training samples from all the other subjects. The goal of this linear combination is to best approximate the naive training sample of the being processed subject. The third step then uses the deviation of the naive and virtual test samples from the naive training sample to evaluate the dissimilarity between the test sample and the subject from which the naive training sample truly originates.

The two representations in our paper, i.e., conventional representation and inverse representation can also be viewed as a multitask learning problem. Multitask learning has been widely studied and applied. For example, Yang et al. [46] proposed a multitask learning-based feature selection method and achieved very good performance in multimedia analysis.

The advantage to integrate the inverse representation and conventional representation can be formally presented as follows. As we know, in real-world applications, the error (or noise) exists in both the test sample and training sample. However, as conventional representation is based on the least-squares algorithm, it takes only the error in the test sample into account. Actually, for conventional representation, Eq. (1) can be rewritten as

\[ X_A = y + \Delta y \]  

(3)

where \( y_0 \) and \( \Delta y \) stand for the true test sample and error, respectively. According to the theoretical analysis, \( A_0 \), i.e., the solution of \( A \), obtained using conventional representation is indeed generated with the following objective function [47]-[49]:

\[
\begin{align*}
[A_0, \Delta y] &= \arg \min \| \Delta y \| \\
\text{s. t.} \quad X_A &= y_0 + \Delta y.
\end{align*}
\]  

(4)

Moreover, the inverse representation in our method takes only the error in the training sample into account. Specifically, in the inverse representation in our method, \( Z_{\beta k} = \lambda_k \) can be rewritten as

\[ Z_{\beta k} = \lambda_k + \Delta \lambda_k \]  

(5)

where \( \lambda_k \) and \( \Delta \lambda_k \) stand for the true value of \( \lambda_k \) and error, respectively. It is easy to know that \( \beta_k \) is obtained using the following objective function:

\[
\beta_k = \arg \min \| \Delta \lambda_k \| + \lambda \| | \beta_k | |
\]  

s. t. \( Z_{\beta k} = \lambda_k + \Delta \lambda_k \).  

(6)

From the above presentation, we see that the integration of the inverse representation and conventional representation allows the error in the test sample and training sample to be simultaneously taken into account and processed. This will be beneficial to achieve good face recognition performance.

We use the chart in Fig. 4 to summarize the effect of every step of the proposed method.

### B. More Quantitative Analysis

In this section, we perform the correlation analysis to show the difference between the score and distance respectively obtained using steps 2 and 3 of the proposed method. In score level fusion, if correlation coefficient between the two kinds of scores to fuse is low, the fusion result is usually good. In other words, a smaller correlation coefficient allows the fusion to better the accuracy. We use the chart in Fig. 4 to summarize the effect of every step of the proposed method.

**Table I**

<table>
<thead>
<tr>
<th>Training samples per class</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean of all the correlation coefficients</td>
<td>0.545</td>
<td>0.520</td>
<td>0.496</td>
</tr>
</tbody>
</table>

Fig. 4. Flow chart of the proposed method.

**Table I**

Mean of All the Correlation Coefficients of the Scores and Distances Respectively Obtained Using Step 2 and Step 3 of the Proposed Method. The First Three, Four and Five Face Images of Each Subject in the ORL Database are Respectively Used as Training Samples and the Others are Taken as Test Samples.
Fig. 5. Scores and distances, of the last test sample, obtained using steps 2 and 3 of the proposed method, respectively. The first four face images of each subject in the FERET database are used as training samples and the others are taken as test samples. The vertical axis shows the values of the score and distance. The horizontal axis shows the no. of the component of the score vector and distance vector. The scores and distances have been normalized.

Fig. 6. Scores and distances, of the last test sample, obtained using steps 2 and 3 of the proposed method, respectively. The first four face images of each subject in the subset of the FERET database are used as training samples and the others are taken as test samples. The vertical axis shows the values of the score and distance. The horizontal axis shows the no. of the component of the score vector and distance vector. The scores and distances have been normalized.

the others are taken as test samples. These two figures also visually tell us that the correlation between the scores and distances is low. This means that the distances generated from step 3 are very complementary to the scores generated from step 2 in recognizing the faces. The first row shows some test samples that were erroneously and correctly classified by LRC and the proposed method, and the corresponding training sample from the subject to which the test sample was erroneously assigned by LRC. The second row shows one training sample from the subject to which the test sample was erroneously assigned by LRC. The first face image of each subject in the ORL database were used as training samples and the others were taken as test samples. w₁ and w₂ were set to 0.6 and 0.4, respectively.

V. EXPERIMENT RESULTS
We used the FERET, ORL and Georgia Tech (GT) face databases to conduct experiments. Each sample was converted into a unit vector with length of 1 in advance. Besides our proposed method, several state-of-the-art face recognition methods such as SRC algorithm proposed in [50], LRC [22], relaxed collaborative representation (RCR) [18], and CRC [12] were also tested. In order to test the robustness of the methods, we also conduct experiments on corrupted face images and show the experimental results in Section V.D. In all the experiments, w₁ and w₂ satisfy the condition w₁ + w₂ = 1 and we show only the value of w₁. The code of our method will be available at.

### A. Experiments on FERET Face Database
We first used a subset of the FERET face database to test our method. This subset consists of 1400 images from 200 individuals, each providing seven images [51]. This subset was composed of images whose names are marked with two-character strings: “ba,” “bj,” “bk,” “be,” “bf,” “bd,” and “bg.” We resized each image to a 40 × 40 image using the down-sampling algorithm. We respectively took the first four face images of each subject as original training samples, and treated the remaining face images as test samples. λ in our method was set to 0.01. The experimental results show that λ has only little influence on the classification performance if it is set to a value near to 0.01. Table II shows the classification accuracies. This table shows that in most cases our method classifies much more accurately than all the other methods. For example, when the first four face images of each subject and the remaining face images were respectively used as original training samples and test samples, the classification accuracies (%) of the proposed method with w₁=0.8, SRC, LRC, CRC, and RCR are 78.63%, 60.00%, 59.62%, 44.37% and 46.12%, respectively. The fact that the accuracy of the well-known SRC is 18.63% lower than that of the proposed method illustrates

<table>
<thead>
<tr>
<th>Method</th>
<th>Training samples per class</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>LRC</td>
<td>44.92</td>
<td>64.20</td>
<td>59.62</td>
<td></td>
</tr>
<tr>
<td>CRC</td>
<td>44.33</td>
<td>58.40</td>
<td>44.37</td>
<td></td>
</tr>
<tr>
<td>RCR</td>
<td>41.67</td>
<td>55.10</td>
<td>46.12</td>
<td></td>
</tr>
<tr>
<td>SRC</td>
<td>50.25</td>
<td>64.80</td>
<td>60.00</td>
<td></td>
</tr>
<tr>
<td>The proposed method (w₁ = 0.8)</td>
<td>55.17</td>
<td>78.10</td>
<td>78.63</td>
<td></td>
</tr>
<tr>
<td>The proposed method (w₁ = 0.7)</td>
<td>53.83</td>
<td>75.40</td>
<td>75.00</td>
<td></td>
</tr>
<tr>
<td>The proposed method (w₁ = 0.6)</td>
<td>51.08</td>
<td>75.40</td>
<td>69.50</td>
<td></td>
</tr>
</tbody>
</table>
that the proposed method can perform very well in recognizing the face.

B. Experiments on ORL Face Database

The ORL database [52] includes 400 face images taken from 40 subjects each providing ten face images. For some subjects, the images were taken at different times, with varying lighting, facial expressions (open/closed eyes, smiling/not smiling), and facial details (glasses/no glasses). Each image was also resized to an image with one half of the original size by using the down-sampling algorithm. We respectively took the first two, three, and four face images of each subject as original training samples and treated the remaining face images as test samples. The experimental results were shown in Table III. It shows again that our method is able to perform better than all the other methods.

C. Experiments on GT Face Database

In this subsection we use the Georgia Tech face database to test our method. The Georgia Tech (GT) face database was built at Georgia Institute of Technology [53]. This database includes face images of 50 people taken in two or three sessions. All people in the database were represented by 15 color JPEG images with cluttered background taken at the resolution of 640×480 pixels. The pictures show frontal and/or tilted faces with different facial expressions, lighting conditions and scale. Each image was manually labeled to determine the position of the face in the image. We use the face image with the background removed. Each used image has the size of 30 by 40. The first three, four and five face images of each subject are used as training samples, respectively. The remaining images are taken as test samples. From Table IV, we see that the proposed method can outperform all the other methods.

D. Experiments on Corrupted Face Images

In order to test the robustness of the methods, we also conduct experiments on corrupted face images. The sets of training samples and test samples are the same as those in Sections V-A–V-C, respectively. For obtaining corrupted face images, we use MATLAB function imnoise to add Gaussian white noise of zero mean and variance of 0.01 to the tested face images, whereas, the training samples are not dealt with so. Fig. 8 shows some of the corrupted face images. Tables V, VI and VII show the classification accuracies of all the methods on the corrupted FERET, ORL, and GT face images, respectively. We see again that the proposed method outperforms all the other methods.

### TABLE III

<table>
<thead>
<tr>
<th>Training samples per class</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>LRC</td>
<td>79.06</td>
<td>81.79</td>
<td>85.00</td>
</tr>
<tr>
<td>CRC</td>
<td>83.44</td>
<td>86.07</td>
<td>89.17</td>
</tr>
<tr>
<td>RCR</td>
<td>77.19</td>
<td>81.07</td>
<td>82.08</td>
</tr>
<tr>
<td>SRC</td>
<td>85.00</td>
<td>85.71</td>
<td>90.00</td>
</tr>
<tr>
<td>The proposed method ($w_1 = 0.8$)</td>
<td>85.62</td>
<td>89.64</td>
<td>89.17</td>
</tr>
<tr>
<td>The proposed method ($w_1 = 0.7$)</td>
<td>86.88</td>
<td>90.36</td>
<td>90.83</td>
</tr>
</tbody>
</table>

### TABLE IV

<table>
<thead>
<tr>
<th>Training samples per class</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>LRC</td>
<td>51.00</td>
<td>55.27</td>
<td>59.40</td>
</tr>
<tr>
<td>CRC</td>
<td>45.33</td>
<td>47.09</td>
<td>48.80</td>
</tr>
<tr>
<td>RCR</td>
<td>36.67</td>
<td>38.36</td>
<td>40.80</td>
</tr>
<tr>
<td>SRC</td>
<td>52.00</td>
<td>56.73</td>
<td>59.80</td>
</tr>
<tr>
<td>The proposed method ($w_1 = 0.9$)</td>
<td>57.00</td>
<td>60.36</td>
<td>62.20</td>
</tr>
<tr>
<td>The proposed method ($w_1 = 0.8$)</td>
<td>55.33</td>
<td>57.82</td>
<td>62.20</td>
</tr>
<tr>
<td>The proposed method ($w_1 = 0.7$)</td>
<td>55.33</td>
<td>57.82</td>
<td>59.60</td>
</tr>
</tbody>
</table>

### TABLE V

<table>
<thead>
<tr>
<th>Training samples per class</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>LRC</td>
<td>19.75</td>
<td>39.00</td>
<td>38.00</td>
</tr>
<tr>
<td>CRC</td>
<td>25.67</td>
<td>27.90</td>
<td>18.03</td>
</tr>
<tr>
<td>RCR</td>
<td>18.58</td>
<td>16.80</td>
<td>12.50</td>
</tr>
<tr>
<td>SRC</td>
<td>34.50</td>
<td>40.90</td>
<td>43.00</td>
</tr>
<tr>
<td>The proposed method ($w_1 = 0.8$)</td>
<td>36.92</td>
<td>55.50</td>
<td>54.13</td>
</tr>
<tr>
<td>The proposed method ($w_1 = 0.7$)</td>
<td>36.50</td>
<td>50.80</td>
<td>49.12</td>
</tr>
<tr>
<td>The proposed method ($w_1 = 0.6$)</td>
<td>35.58</td>
<td>45.70</td>
<td>44.75</td>
</tr>
</tbody>
</table>
representation based face recognition methods. The devised inverse representation samples will enable the test sample to be better represented which indeed reflects possible variation of the face pose.

Thus, the simultaneous use of the original and virtual training which almost always not an axis-symmetrical image owing to the non-absolute frontal view pose makes it reasonable to exploit the mirror image of the face to generate virtual face image, nonnegative matrix factorization with application to facial expression classification and face recognition experiments," in Proc. IEEE Int. Conf. Audio, Speech, Signal Process., 2010, pp. 1–6.

The fact that the symmetry of the face is a basic nature of the face and in real-world applications the face image is almost always not an axis-symmetrical image owing to the non-absolute frontal view pose makes it reasonable to exploit the mirror image of the face to generate virtual face image, which indeed reflects possible variation of the face pose. Thus, the simultaneous use of the original and virtual training samples will enable the test sample to be better represented by a linear combination of the training samples truly from the same class as the test sample and will enable higher accuracy to be obtained. The devised inverse representation is also helpful and is very complementary to the conventional representation in evaluating the dissimilarity between the test sample and each class. The experimental results show that the proposed method can outperform a number of state-of-the-art representation based face recognition methods.

VI. CONCLUSION

The fact that the symmetry of the face is a basic nature of the face and in real-world applications the face image is almost always not an axis-symmetrical image owing to the non-absolute frontal view pose makes it reasonable to exploit the mirror image of the face to generate virtual face image, which indeed reflects possible variation of the face pose. Thus, the simultaneous use of the original and virtual training samples will enable the test sample to be better represented by a linear combination of the training samples truly from the same class as the test sample and will enable higher accuracy to be obtained. The devised inverse representation is also helpful and is very complementary to the conventional representation in evaluating the dissimilarity between the test sample and each class. The experimental results show that the proposed method can outperform a number of state-of-the-art representation based face recognition methods.

REFERENCES


malization for robust face recognition using discrete cosine transform in


feature extraction for face representation and recognition,” IEEE Trans.

representation for computer vision and pattern recognition,” Proc. IEEE,


pp. 1033–1040.


representation: Which helps face recognition,” in Proc. IEEE Int. Conf.


[16] X. Gao, N. Wang, X. Li, and D. Tao, “Face sketch-photo synthesis and

scores with conventional matching scores for multimodal biometrics and


[23] I. Naseem, R. Togneri, and M. Bennamoun, “Robust regression for face


representation for pattern classification,” in Proc. Int. Conf. Pattern
Recognit., 2010, pp. 115–120.


Vis., 2010, pp. 1–14.


[29] B. Zhu, M. Floyd, Q. Ruan, and W. B. Kleijn, “Graph-preserving sparse
nonnegative matrix factorization with application to facial expression

the eigenstructure for discrimination,” IEEE Trans. Neural Netw.


[42] X. Xu, X. Li, J. Yang, and D. Zhang, “Integrate the original face image and its mirror image for face recognition,” Neurocomput., to be published.


Yong Xu (M’06) received the B.S. and M.S. degrees from the Air Force College of Meteorology, Nanjing, China, in 1994 and 1997, respectively, and the Ph.D. degree in pattern recognition and intelligence system from Nanjing University of Science and Technology, Nanjing, China, in 2001.

From 2005 to 2007, he was with Shenzhen Graduate School, Harbin Institute of Technology, Harbin, China, as a Post-Doctoral Research Fellow. He is currently an Associate Professor at Shenzhen Graduate School, Harbin Institute of Technology. His current interests include pattern recognition, biometrics, image processing, and video analysis.

Xuelong Li (M’02-SM’07-F’12) is currently a Full Professor with the Center for Optical Imaging and Intelligent System (COIIIS), State Key Laboratory of Transient Optics and Photonics, Chinese Academy of Sciences, Shandong, China.

Jian Yang (M’08) received the B.S. degree in mathematics from Xinhua Normal University, Xiamen, China, in 1995, the M.S. degree in applied mathematics from the Changsha Railway University, Changsha, China, in 1998, and the Ph.D. degree in computer science from the Nanjing University of Science and Technology (NJUST), Nanjing, China, in 2002. He was a Post-Doctoral Researcher at the University of Zaragoza, Zaragoza, Spain, in 2003. From 2004 to 2006, he was a Post-Doctoral Fellow with Hong Kong Polytechnic University, Hong Kong. From 2006 to 2007, he was a Post-Doctoral Fellow with the Department of Computer Science, New Jersey Institute of Technology, NJ, USA. Since 2007, he has been a Professor with the School of Computer Science and Technology, NUST. He has authored more than 50 scientific papers in pattern recognition, computer vision, and the related areas. His journal papers have been cited more than 1100 times in the ISI Web of Science, and 2400 times on Scholar Google.

Dr. Yang is an Associate Editor of Pattern Recognition Letters and IEEE TRANSACTIONS ON NEURAL NETWORKS.

Zhihui Lai received the B.S. degree in mathematics from South China Normal University, Guangdong, China, the M.S. degree from Jinan University, Guangdong, and the Ph.D. degree in pattern recognition and intelligence system from Nanjing University of Science and Technology, Nanjing, China, in 2002, 2007, and 2011, respectively.

He was a Research Associate and a Post-Doctoral Fellow at the Hong Kong Polytechnic University, Hong Kong, from 2010 to 2013. He is currently a Research Associate and a Post-Doctoral Fellow at the Bio-Computing Research Center, Shenzhen Graduate School, Harbin Institute of Technology, Harbin, China. He has authored more than 30 scientific papers in pattern recognition and computer vision. His current research interests include face recognition, image processing and content-based image retrieval, pattern recognition, compressive sensor, and human vision modeling and applications in the fields of intelligent robot research.

David Zhang (F’08) received the bachelor’s degree in computer science from Peking University, Beijing, China, the M.Sc. and Ph.D. degrees from the Harbin Institute of Technology, Harbin, China, in 1982 and 1985, respectively, and the second Ph.D. degree in electrical and computer engineering from the University of Waterloo, ON, Canada, in 1994. From 1986 to 1988, he was a Post-Doctoral Fellow at Tsinghua University, Beijing, and then an Associate Professor at the Academia Sinica, Beijing. He is currently a Chair Professor at the Hong Kong Polytechnic University, Hong Kong, where he is also the Founding Director of the Biometrics Technology Centre (UGC/CRC), supported by the Hong Kong SAR Government in 1998. He also serves as the Visiting Chair Professor with Tsinghua University, and an Adjunct Professor with Shanghai Jiao Tong University, Shanghai, Peking University, Harbin Institute of Technology, and the University of Waterloo. He has authored more than ten books and 200 journal papers.

De Zhang is the Founder and Editor-in-Chief of the International Journal of Image and Graphics, the Book Editor for Springer International Series on Biometrics, and an Associate Editor of more than ten international journals including the IEEE Transactions and Pattern Recognition. He was also the organizer of the 1st International Conference on Biometrics Authentication, and the Technical Committee Chair of IEEE CIS. He is a Croucher Senior Research Fellow, Distinguished Speaker of the IEEE Computer Society, and a fellow of IAPR.