Pattern Recognition 54 (2016) 68–82

Contents lists available at ScienceDirect

Pattern Recognition

journal homepage: www.elsevier.com/locate/pr

Approximately symmetrical face images for image preprocessing in face recognition and sparse representation based classification

Yong Xu [a,b], Zheng Zhang [a,b,*], Guangming Lu [a], Jian Yang [c]

[a] Bio-Computing Research Center, Shenzhen Graduate School, Harbin Institute of Technology, Shenzhen 518055, Guangdong, P.R. China
[b] Key Laboratory of Network Oriented Intelligent Computation, Shenzhen 518055, Guangdong, P.R. China
[c] College of Computer Science and Technology, Nanjing University of Science and Technology, Nanjing 210094, Jiangsu, P.R. China

ARTICLE INFO

Article history:
Received 17 July 2015
Received in revised form
28 October 2015
Accepted 28 December 2015
Available online 25 January 2016

Keywords:
Approximately symmetrical face
Virtual sample
Face image preprocessing
Sparse representation
Face recognition

ABSTRACT

Though most of the faces are axis-symmetrical objects, few real-world face images are axis-symmetrical images. In the past years, there are many studies on face recognition, but only little attention is paid to this issue and few studies to explore and exploit the axis-symmetrical property of faces for face recognition are conducted. In this paper, we take the axis-symmetrical nature of faces into consideration and design a framework to produce approximately axis-symmetrical virtual dictionary for enhancing the accuracy of face recognition. It is noteworthy that the novel algorithm to produce axis-symmetrically virtual face images is mathematically very tractable and quite easy to implement. Extensive experimental results demonstrate the superiority in face recognition of the virtual face images obtained using our method to the original face images. Moreover, experimental results on different databases also show that the proposed method can achieve satisfactory classification accuracy in comparison with state-of-the-art image preprocessing algorithms. The MATLAB code of the proposed method can be available at http://www.yongxu.org/lunwen.html.

1. Introduction

In the face recognition community, various algorithms have been proposed to remedy difficulties caused by the small sample size (SSS) problem and variations of poses, illuminations and facial expressions [1,2]. It should be pointed out that the problem of varying poses, illuminations and facial expressions usually has the following characteristics. Each face usually has severe variations of poses, illuminations and facial expressions but a face has only a limited number of available face images [3]. This of course causes a great intra-subject difference and the difference may even be greater than the inter-subject difference. This may incur the result that different faces have similar patterns [4,5]. As a consequence, the performance of face recognition algorithms will be seriously affected.

Many efforts have been made to address these difficulties from the viewpoint of methodology. For example, when linear discriminant analysis (LDA) [6,8] is applied to face recognition, it usually suffers from the SSS problem. In other words, because the number of face samples for training is almost always smaller than the dimension of the face sample, LDA cannot be directly implemented. A number of variants of the original LDA algorithm have been proposed to overcome this problem. Moreover, various methods have been proposed to perform face recognition across poses, illuminations and facial expressions and to reduce disadvantageous effect on recognition of faces of variations of poses, illuminations and facial expressions [7].

Besides special algorithms are designed to overcome above difficulties, to obtain more virtual face samples is another important way to improve the performance of face recognition [9,10]. In recent years researchers have made notable advances in this way. For example, symmetrical face images proposed in [11] are very beneficial to overcome the problem of varying poses and illuminations. The mirror face image based face recognition method proposed in [12] has the following merits. The used mirror face images not only can provide possible poses of faces that are not shown in the original face images but also offer us natural face images. Li et al. [13] proposed to enlarge the number of the training samples by exploiting the inter-class relationship and combined the original samples and obtained virtual samples to perform face recognition. The simultaneous use of these synthesized images and original face images can lead to very good performance. Besides the above symmetrical face images and mirror face images, noised or corrupted face images are also available virtual face samples [14,15].

We see that a few methods take the facial geometry into account when generating more available face samples. However, when

* Corresponding author at: Bio-Computing Research Center, Shenzhen Graduate School, Harbin Institute of Technology, Shenzhen 518055, Guangdong, P.R. China
E-mail address: darrenzz219@gmail.com (Z. Zhang).

http://dx.doi.org/10.1016/j.patcog.2015.12.017
0031-3203/

humans recognize faces, they often exploit the facial geometry information. A common property of most faces is that they have a symmetrically geometric structure. In other words, the left-half face is almost always the "mirror image" of the right-half face. This motivates us to explore the way to exploit the symmetrical structure of faces when designing automatic face recognition algorithms. It has been proved that the performance of some tasks such as face detection can be promoted by using the symmetrical structure of faces [16–18]. The facial symmetry can also be applied for shadow compensation and analysis [19,20]. Moreover, the symmetry of faces has also been shown to be partially beneficial to 3D face recognition [21–23]. However, it seems that how to effectively exploit the facial symmetry to perform 2D face recognition is not studied in-depth. Especially, in the 2D face recognition community, there are only few studies on how to design an algorithm of automatically generating symmetrical face images. For 2D face recognition, the literature showed that if the 2D face image is completely an axis-symmetrical image, then we can use only one-half of the face image for the classification of faces [24]. However, most of the real 2D face images are not axis-symmetrical. Thus, if only one-half of the face image is used for matching and classification of faces, there will be much information loss and the classification performance will be degraded. Previous research also demonstrates that enhancing the symmetry of a face image is a meaningful and effective strategy [25,26].

In the past, there were some studies on preprocessing of face images [29,30]. For example, some methods to reduce effects on face recognition of illuminations have been proposed [27,28]. With the development of the time–frequency theory, frequency domain features such as discrete cosine transform (DCT) feature [31] were extensively used for face recognition. Because the small-scale feature is more robust to illumination variations in recognition tasks, the logarithmic total variation (LTV) model [32] and the logarithmic wavelet transform (LWT) method [33] based on the multi-scale facial structure representation were proposed as image preprocessing methods for face recognition. In the light of the fact that each face image can be constructed by the integration of the large- and small-scale features, Xie et al. [34] proposed to integrate the large- and small-scale features (LSSF) of face images to perform normalization of face illumination. Zhang et al. [35] proposed to use gradientfaces (GRF) and Šrtuc and Pavešić [36] presented a non-local means based normalization method (NLM) as image preprocessing techniques for robust face recognition. Cyganek et al. [37] proposed a very impressive method, which used Support Vector Machines with specific tensor kernel and achieved high accuracy for processing multidimensional images (among which are face datasets) without any need for sophisticated pre-processing or posing restrictions on the face pose. In addition, the Lambertian reflectance model also has received much attention in remedying the problems caused by unbalanced illuminations [39]. However, a real face image is usually not completely subject to the assumption of this model. The illuminating light field method also seems to be applicable for face recognition across illuminations and poses [40]. However, this method is still established on the basis of the Lambertain reflectance model.

In this paper, we propose a novel face image preprocessing method and also use it to generate a virtual dictionary for image classification. The proposed method aims at producing virtual face images with an approximately symmetrical structure, which will be utilized in both image preprocessing and image classification. Our method has the following notable merits. (1) Our method can well reduce the negative effect on face images of heterogeneous illuminations and it can be used as a face image preprocessing method. (2) The proposed method can automatically produce an approximately axis-symmetrical virtual face image from an original face image. Because the virtual face image is approximately axis-symmetrical, it can effectively reduce negative effects on face recognition of varying face appearance. As a result, the use of the virtual face images enables the accuracy of a face recognition algorithm to be improved. Thus, this not only provides a good solution to the face image preprocessing problem, but also can generate geometrically symmetrical and more attractive face images. To evaluate the performance of our method in two different applications, i.e. face image preprocessing and generation of virtual dictionary, we design experiments composed of two parts. We first applied our method as an image preprocessing method to all images including training samples and test samples from different datasets, and then exploited the nearest neighbor classifier for classification. In addition, we also applied our method on different face databases to generate virtual training dictionaries and then employed different sparse presentation based classification algorithms for classification. It should be pointed out that in the second part of experiments, our method was applied only to the original training sample and the test samples were not used to generate virtual samples. This makes use of our method partially similar to the dictionary learning method in face recognition. In summary, the paper has the following important contributions:

- It proposes an originally creative idea and algorithm to automatically produce approximately axis-symmetrical virtual face images.
- The proposed method can be treated as an image preprocessing method for face recognition, and extensive experiments on different face databases show the effectiveness of our method.
- The images synthesized by our method can be used as a virtual image dictionary, and the strong recognition capability of our method is verified in comparison with state-of-the-art dictionary learning algorithms in image classification.
- As an unsupervised image classification and image preprocessing method, our method can work much better than the benchmark that uses the original image datasets for classification.

The rest of our paper is organized as follows. Section 2 presents the description of the proposed method in detail. The analysis and advantages of the proposed method and the obtained virtual images are shown in Section 3. In Section 4, we introduce the experimental evaluation of our method as the image preprocessing method on some widely used databases. In Section 5, the experimental evaluation of generating the virtual dictionary is reported. The analysis and discussion are explicitly presented in Section 6. Finally, we give our conclusions in Section 7.

2. The description of the proposed method

In this section we describe the main steps of the proposed method in detail. Let $\boldsymbol{z}_1$ and $\boldsymbol{z}_2$ denote the left vector and right vector of an original face image, respectively. In particular, both $\boldsymbol{z}_1$ and $\boldsymbol{z}_2$ are supposed to be column vectors. The procedure to obtain $\boldsymbol{z}_1$ and $\boldsymbol{z}_2$ is as follows. Let $\boldsymbol{F}$ stand for an original face image. Let $\boldsymbol{F}_j$ $(j = 1, 2, \ldots, J)$ denote that j-th column of 2D image matrix $\boldsymbol{F}$. The pixel at the i-th row and j-th column of $\boldsymbol{F}$ is denoted by $\boldsymbol{F}_{ij}$. Suppose that J is an even number. $\boldsymbol{z}_1$ and $\boldsymbol{z}_2$ are defined as

$$\boldsymbol{z}_1 = \begin{bmatrix} \boldsymbol{F}_1 \\ \boldsymbol{F}_2 \\ \cdot \\ \cdot \\ \cdot \\ \boldsymbol{F}_{J/2} \end{bmatrix}, \quad \boldsymbol{z}_2 = \begin{bmatrix} \boldsymbol{F}_J \\ \boldsymbol{F}_{J-1} \\ \cdot \\ \cdot \\ \cdot \\ \boldsymbol{F}_{J/2+1} \end{bmatrix} \tag{1}$$

In other words, $\boldsymbol{z}_1$ is obtained by concatenating the first column to the $\frac{J}{2}$-th column of $\boldsymbol{F}$ one by one and $\boldsymbol{z}_2$ is obtained by concatenating the last column to the $\frac{J}{2}+1$-th column of $\boldsymbol{F}$ one by one.

Fig. 1 presents one simple example of implementing the above procedures and Fig. 2 intuitively shows how to obtain the vectors of the left-half and right-half faces of a real face image from the Extended YaleB face database [41].

It should be pointed out that the definitions in (1) are just initial values of $\boldsymbol{z}_1$ and $\boldsymbol{z}_2$, respectively, denoted by $\boldsymbol{z}_1^0$ and $\boldsymbol{z}_2^0$. Our method has the goal to finally obtain approximately equal $\boldsymbol{z}_1$ and $\boldsymbol{z}_2$. It is a proper way to use the gradient descent algorithm to iteratively update the left and right vectors. As we know, the gradient descent procedure of function $f(x)$ is

$$x^{t+1} = x^t - \eta \nabla f(x^t) \tag{2}$$

where η is the learning ratio, x^{t+1} and x^t are the values of x at time $t+1$ and t, respectively. The gradient descent algorithm usually can fast get the minimum value of $f(x)$ and the corresponding optimal value of x. For our problem, a reasonable function can be defined as $L(\boldsymbol{z}_1, \boldsymbol{z}_2) = \|\boldsymbol{z}_1 - \boldsymbol{z}_2\|_2^2$. It is clear that approximately equal $\boldsymbol{z}_1$ and $\boldsymbol{z}_2$ means that $L(\boldsymbol{z}_1, \boldsymbol{z}_2)$ should reach its minimum value. It is easy to get $\nabla_{\boldsymbol{z}_1} L(\boldsymbol{z}_1, \boldsymbol{z}_2) = 2\boldsymbol{z}_1 - 2\boldsymbol{z}_2$ and $\nabla_{\boldsymbol{z}_2} L(\boldsymbol{z}_1, \boldsymbol{z}_2) = 2\boldsymbol{z}_2 - 2\boldsymbol{z}_1$. As a result, the formulae to iteratively update $\boldsymbol{z}_1$ and $\boldsymbol{z}_2$ are

$$\boldsymbol{z}_1^{t+1} = \boldsymbol{z}_1^t - \eta(\boldsymbol{z}_1^t - \boldsymbol{z}_2^t) \tag{3}$$

$$\boldsymbol{z}_2^{t+1} = \boldsymbol{z}_2^t - \eta(\boldsymbol{z}_2^t - \boldsymbol{z}_1^t) \tag{4}$$

In order to make the solution numerically stable, we set η to $\eta = \frac{\alpha}{t}$. α is a positive constant and t is the number of iterations. Because η decreases with the increase of t, the solution is easy to converge. After the optimal vectors $\boldsymbol{z}_1^t$ and $\boldsymbol{z}_2^t$ are obtained, the new synthesized approximately axis-symmetrical face image is achieved by concatenating them. We summarize the main steps of the proposed method as follows.

Step 1: We obtain the initial values of the $\boldsymbol{z}_1$ and $\boldsymbol{z}_2$, i.e. $\boldsymbol{z}_1^0$ and $\boldsymbol{z}_2^0$, from an original face image using Eq. (1). Specifically, we separate the original image into the left-half and right-half face images, and reverse the right-half face, as shown in the upper-half of Fig. 2. All the samples are normalized to unit vectors with l_2-norm.

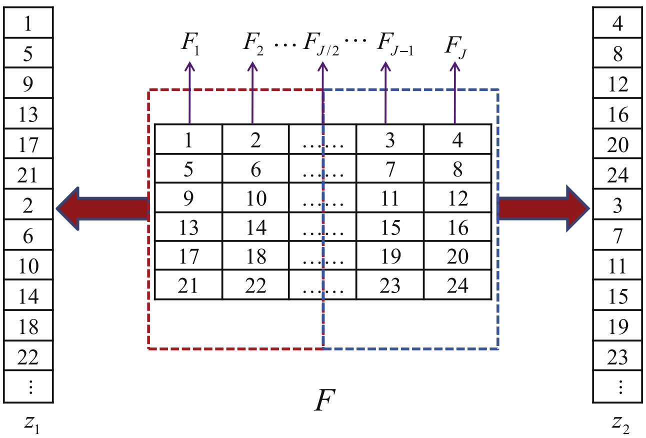

Fig. 1. A simple example for converting an image to the left and right vectors.

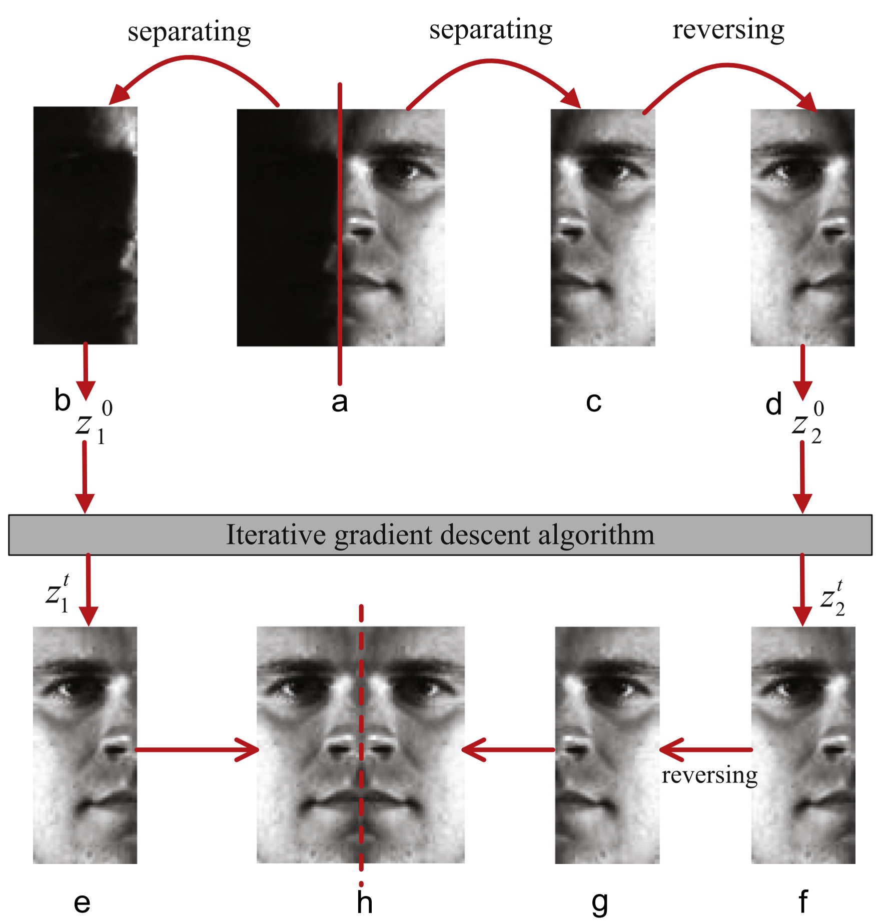

Fig. 2. The general framework of synthesizing an approximately axis-symmetrical face image. (a) is one original image from the Extended YaleB database [41]. (b) is the separated left-half face, which is converted to vector $\boldsymbol{z}_1^0$. (c) and (d) are respectively the separated right-half face image and its mirror image, which is converted to vector $\boldsymbol{z}_2^0$. (e) and (f), i.e. optimized $\boldsymbol{z}_1^t$ and $\boldsymbol{z}_2^t$, are the final optimal left-half and right-half face images corresponding to (b) and (d), respectively. (g) is the mirror image of image (f). (h) is the finally synthesized approximately axis-symmetrical face image.

Step 2: We update $\boldsymbol{z}_1$ and $\boldsymbol{z}_2$ by exploiting the iterative gradient descent algorithm presented in Eqs. (3) and (4). The iterative updating is not terminated until either of the following two conditions is satisfied. (i) The number of iterations is greater than the predefined maximum value. (ii) $\|\boldsymbol{z}_1^{t+1} - \boldsymbol{z}_1^t\| < \varepsilon$. $\boldsymbol{z}_1^t$ denotes the value of $\boldsymbol{z}_1$ at time t and ε stands for a small positive constant.

Step 3: After the optimal $\boldsymbol{z}_1^t$ and $\boldsymbol{z}_2^t$ are obtained, they are connected to form a new face image with the same size as the original image. Specifically, as shown in the lower-half of Fig. 2 (e)–(f) presents the process to connect the optimal $\boldsymbol{z}_1^t$ and $\boldsymbol{z}_2^t$ to obtain a final approximately axis-symmetrical face image, i.e. virtual face image (h).

Step 4: Apply a face recognition algorithm to the obtained approximately axis-symmetrical face images to determine the label of the test sample to be dealt with.

Generally, the proposed method can be applied to two main fields: image preprocessing and image classification. We first applied Steps 1–3 of our method as an image preprocessing method to all images including the training samples and test samples from different datasets and the processed images were evaluated by exploiting Step 4 of our method. Moreover, we also use our method as a dictionary learning method by first applying it to generate virtual training dictionaries, and then perform image classification. In image classification application, only the original training samples were used to generate virtual samples. This makes the use of our method partially similar to the dictionary learning method in face recognition.

In this paper, the nearest neighbor classifier is utilized to verify the effectiveness of our method in the image preprocessing problem. On the other hand, several sparse representation based classification (SRC) methods are used to perform image classification based on the obtained virtual images. For image classification application, we treat n samples from the virtual image dictionary, i.e. the set of virtual images, as training samples. Let $X = [x_1, x_2, \ldots, x_n]$ be all the virtual images from c classes, X_i denotes the virtual images of the i-th class and the testing sample is y. The general scheme of sparse representation based classification methods [43] is summarized in the following SRC Algorithm.

SRC Algorithm. The scheme of SRC methods.

Step 1: Normalize all the samples in the virtual image dictionary with the unit l_2-norm.

Step 2: Employ all the training samples to represent the test sample with the l_1-norm regularization

$\boldsymbol{\alpha}^* = \arg\min \|\boldsymbol{\alpha}\|_1 \quad \text{s.t.} \ \|\boldsymbol{y} - X\boldsymbol{\alpha}\|_2^2 \le \xi$.

Step 3: Compute the representation residual for each class

$\boldsymbol{r}_i = \|\boldsymbol{y} - X_i \boldsymbol{\alpha}_i^*\|_2^2$

where $\boldsymbol{\alpha}_i^*$ denotes the representation coefficients vector associated with the i-th class.

Step 4: The test sample to be dealt with is assigned to the i-th class by judging $label(\boldsymbol{y}) = \arg\min_i(\boldsymbol{r}_i)$.

3. Analysis and rationale of the proposed method

3.1. Analysis and advantages of the proposed method

The proposed method has the following advantages. First, it can easily generate symmetrical virtual face images from original face images. This not only can reduce the negative effect on face recognition of facial pose and illumination variations but also may make the geometry of face images look more consistent with that of true faces. Second, the proposed method is simple and easy to implement. It produces approximately symmetrical face images at a low computational cost. In addition, we say that the proposed method also provides alternative representation of faces.

Another merit of the proposed method is that this method can work well for the recovery of face images, which is an important issue in applications on face images. For example, the proposed method can be used as a preprocessing stage of some applications that have non-ideal original face images but need good-quality neutral face images. Especially, because symmetrical faces are more attractive, the proposed method is beneficial to obtain more beautiful face images. Moreover, the virtual image obtained using our method not only can increase the similarities between images from the same class but also can enhance the differences between different classes. The proposed method also seems to be applicable for some other pattern classification issues.

3.2. Illustrations of virtual face images

In this subsection, we provide figures to intuitively show results and virtual face images of our method. In order to more clearly describe the performance in enhancing the symmetry of the face image of our method, we implemented our method on the original images without normalizing the image vectors to unit vectors. Here we exploited some images from the Extended YALEB face dataset. From Fig. 3, we see that our method can well alleviate the negative effect of unbalanced illuminations and obtain good-quality virtual face images. The distance between the left-half and right-half faces is defined as the l_2-norm of the left-half image vector minus right-half image vector, i.e. $d_{lr} = \|\boldsymbol{z}_1 - \boldsymbol{z}_2\|_2$. Table 1 shows the distance differences between the left-half and right-half faces of the original face images and the corresponding virtual face images shown in Fig. 3. The virtual face images are obtained under the condition that our image preprocessing method is terminated if the maximum iteration number reaches 500 or the distance between the left-half and right-half faces is less than 150. It is clear that the distance between the left-half and right-half faces of the virtual face image is much less than the distance between the left-half and right-half faces of the corresponding original face images. This means that the virtual face image is approximately symmetrical whereas the original face image is not.

Fig. 4 shows distance variations between the left-half and right-half faces of the four virtual face images shown in Fig. 3 with the number of iterations. Here the original face image is not normalized to a unit vector. The fact that the distance decreases with the increase of the number of iterations demonstrating that our method is convergent. As we know, the convergence is a basic nature that an iterative method should possess.

3.3. Comparison with other face image preprocessing methods

It is demonstrated that variations of illuminations, facial expressions and poses can significantly influence the accuracies of face recognition methods and the removal of such noise is beneficial to improve the accuracy of image classification methods. Extensive image preprocessing methods have been proposed and the original purpose of image preprocessing in face recognition is mainly to improve the image quality, which is favorable to visual perception and classification of images. Thus, we compare the visual results of different image preprocessing methods to evaluate the effectiveness of these methods and use Fig. 5 to illustrate the image preprocessing results. We can see that all methods can somewhat eliminate the illumination variations and alleviate the distortion. However, all of the compared methods cannot perfectly address this problem. For example, DoG features [27] shown in Fig. 5(b) cannot effectively remove the influences of illumination changes, and images obtained

Fig. 3. Several original face images and the corresponding virtual face images of a subject in the YALEB face dataset. The first and second rows show the original face images and corresponding virtual face images, respectively.

Table 1
The distance between the left-half and right-half faces of the original face images and corresponding virtual face images shown in Fig. 3.

Distance between the left and right faces of the original face image	7948.0	6431.1	8631.8	7084.5
Distance between the left and right faces of the corresponding virtual face image	458.5	449.8	483.6	449.3

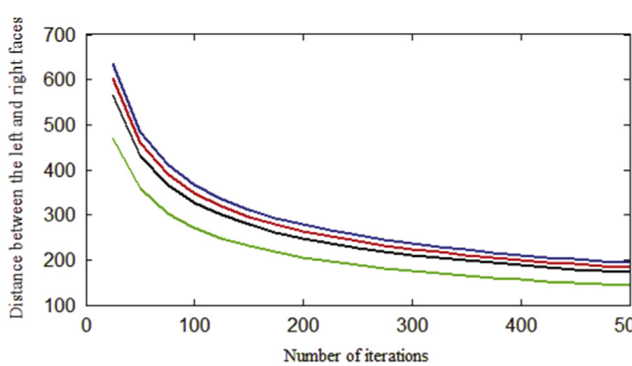

Fig. 4. Distances between the left-half and right-half faces of the four virtual face images shown in Fig. 3. The lines with four different colors denote the distances corresponding to the four virtual face images, respectively. (For interpretation of the references to color in this figure caption, the reader is referred to the web version of this paper.)

using TT in [27] just preserve a small quantity of image information and cannot substantially restore all facial details. Fig. 5(g) shows that LSSF [34] can slightly remove some shadows, but the obtained images seem to be abnormal and some regions around eyes appear to be strange. In contrast, our method not only can effectively reconstruct the missing information caused by illumination variations but also can recover a natural face image from the degraded images by preserving much more facial detail information. Furthermore, by applying the proposed method, we can always obtain approximately axis-symmetrical images. However, the other methods cannot do so and are not able to well exploit the intrinsic geometry of the face.

4. Experimental evaluation of our method as the image preprocessing method

In our experiments, the proposed method was evaluated by using the FERET [44], ORL [45], Georgia Tech (GT) [46] and Labeled Faces in the Wild (LFW) databases [47]. In order to evaluate the capability of our method in image preprocessing, we compared our method with state-of-the-art image preprocessing methods, i.e. the logarithmic wavelet transform (LWT) method [33], logarithmic total variation (LTV) [32], local binary pattern histogram Fourier (LBPHF) features [48], large- and small-scale features normalization technique (LSSF) in [34], non-local means based normalization technique (NLM) [36], gradientfaces (GRF) in [35], DCT normalization technique in [31], Tan and Triggs normalization technique (TT) in [27], DoG filtering-based normalization technique (DoG) [27], and multiscale Weberfaces normalization technique (MSW) in [28]. All these methods are summarized in Table 2. All the images are first transformed into gray scale images and the nearest neighborhood (NN) classifier is used for classification of faces.

4.1. Data sets and experimental setup

FERET: The first dataset to be tested is a subset of the well-known FERET face dataset, which is publicly available for evaluating face recognition algorithms and contains 14,126 images from 1199 people. The FERET database is viewed as one of the largest facial images databases. The used subset contains 1400 face images. The images were collected in a semi-controlled environment. These 1400 face images were from 200 subjects and each subject had seven different face images. This subset was composed of images in the original FERET face dataset whose names are marked with two-character strings: ba, bj, bk, be, bf, bd, and bg. Every face image was resized to a 40 by 40 image. Fig. 6(a) shows some cropped face images from this subset. We respectively took the first 1, 2, 3, 4 and 5 face images of each subject as training samples and regarded the remaining face images as test samples.

GT: The Georgia Tech (GT) face database contains images of 50 subjects and each subject has 15 face images. All pictures in the database are represented by 15 color JPEG images. The pictures show frontal and/or tilted faces with different facial expressions, lighting conditions and scales. Fig. 6(b) presents some original face images from the GT face database. In our experiments, all images in the database were manually cropped and resized to 30 by 40 images. Most of the complex background was excluded for cropped images. They are further converted to gray level images before the methods are tested. In our experiments on this database, the first 1, 2, 3, 4 and 5 images of each subject were treated as training samples and the remaining face images were used as test samples.

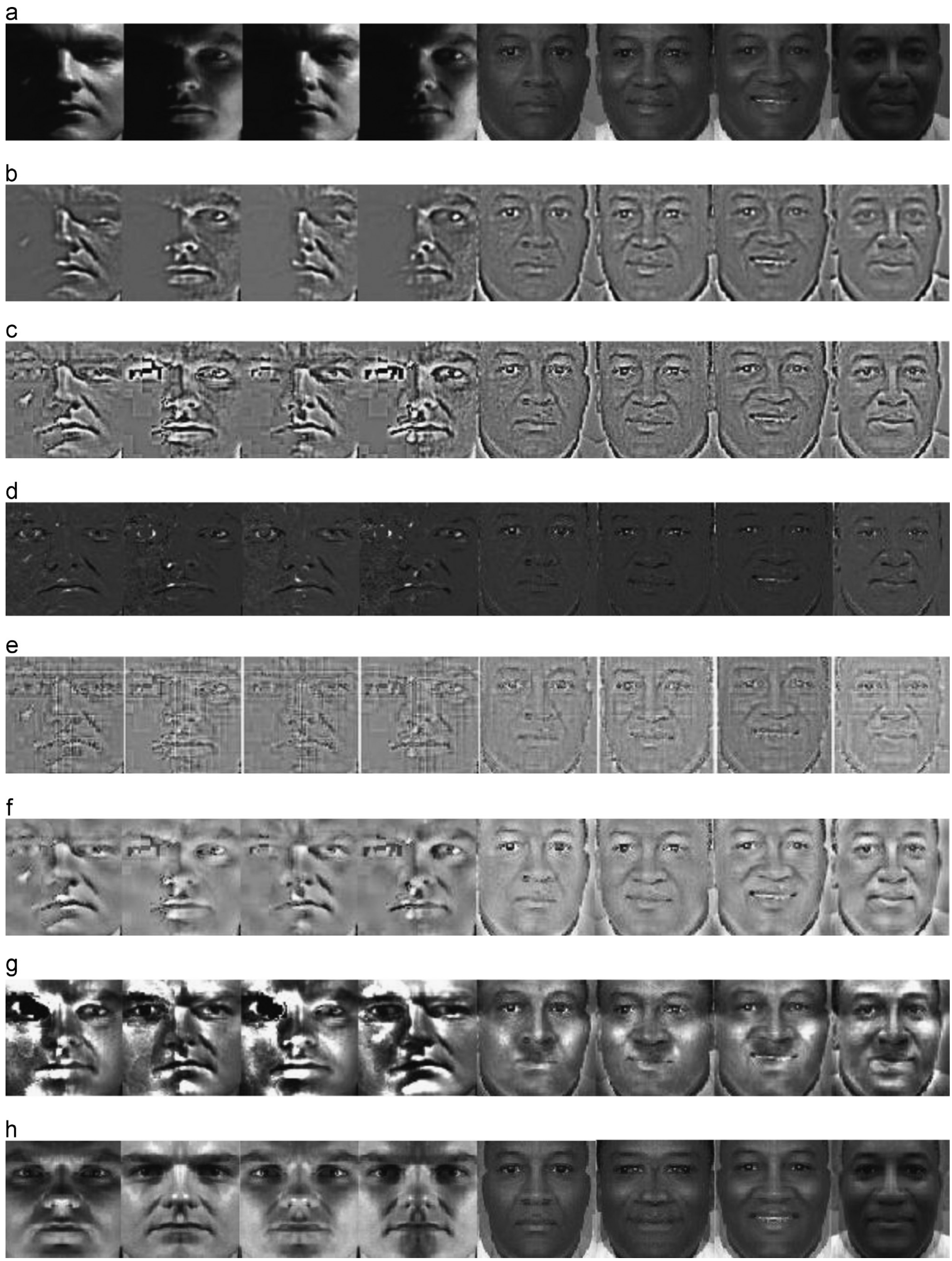

Fig. 5. Examples of image preprocessing results of images from the Extended YaleB and FERET databases. The first four images and the rest four images of each row are some samples from the Extended YaleB and FERET face databases, respectively. (a) original images. (b)–(g) are resultant images of the DoG [27], TT [27], LTV [32], LWT [33], DCT [31] and LSSF [34] methods, respectively. (h) are resultant images of our method.

Table 2
List of important acronyms.

Acronym	Description
LSSF	Large- and small-scale features [34]
GRF	Gradientfaces normalization technique [35]
LWT	Logarithmic wavelet transform technique [33]
LTV	Logarithmic total variation technique [32]
NLM	Non-local means based normalization technique [36]
DCT	Discrete cosine transform normalization technique [31]
LBPHF	Local binary pattern histogram Fourier features [48]
TT	Tan and Triggs normalization technique [27]
DoG	DoG filtering-based normalization technique [27]
MSW	Multi-scale Weberfaces normalization technique [28]
SRC	Sparse representation based classification method [38]
l1_ls	Large-scale l1-regularized least squares [54]
FISTA	Fast iterative shrinkage-thresholding algorithm [55]
Homotopy	Homotopy algorithm based classification [56]
LCKSVD	Label consistent K-SVD [57]
DKSVD	Discriminative K-SVD [58]
SFTS	Symmetrical face training samples [11]
MFTS	Mirror face training samples [3]
ICR	Inter-class relationship [13]
MFL	Metaface learning [59]
RNVS	Random noise based virtual sample [14]

ORL: The ORL database includes 400 face images taken from 40 subjects each providing 10 face images. For some subjects, the images were taken at different time, with varying lighting, facial expressions (open/closed eyes, smiling/not smiling), and facial details (glasses/no glasses). All the images were taken against a dark homogeneous background with the subjects in an upright, frontal position. Each image was resized to a 56 by 46 image matrix by using the down-sampling algorithm. Fig. 6(c) shows some cropped face images from the ORL face database. We also selected the first 1, 2, 3, 4 and 5 images of each subject as training samples and used the remaining face images as test samples.

LFW: The Labeled Faces in the Wild (LFW) face database is designed for the study of unconstrained identity verification and face recognition. It contains more than 13,000 images of faces collected from the web under the unconstrained conditions. Each face image has been labeled with the name of the people pictured. 1680 of the people pictured have two or more distinct photos in the database. In our experiments, we chose 1251 images from 86 people and each subject has 10–20 images [49]. Each image was manually cropped and resized to 32×32 pixels. Fig. 6(d) shows some cropped samples from the LFW database. The first 4, 5, 6, 7 and 8 images of each subject were regarded as training samples and the remaining face images were exploited as test samples.

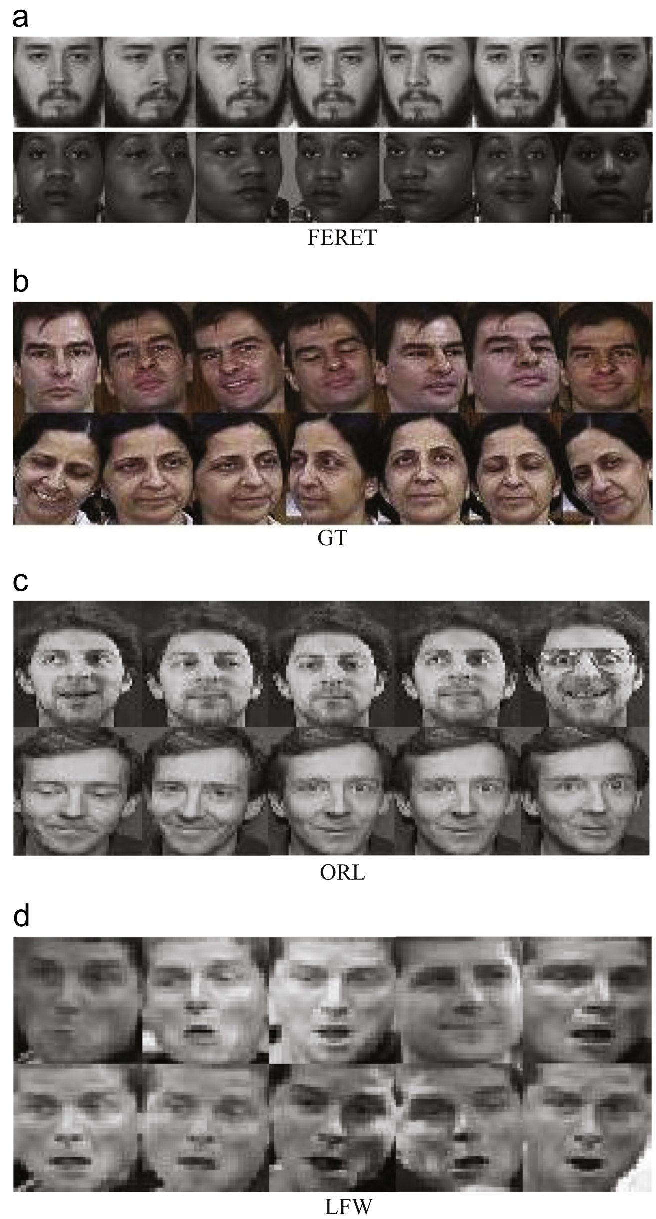

Fig. 6. Some images from different face databases. (a), (b), (c) and (d) show images from the FERET, GT, ORL and LFW databases, respectively. (For interpretation of the references to color in this figure caption, the reader is referred to the web version of this paper.)

4.2. Experimental configuration

In this subsection, we provide the implementation details of the proposed method and parameter setting of the compared methods. We first directly applied different image processing methods to all the original face images from different databases, and then employed a nearest neighbor classifier for classification. After image pre-processing, any dimension reduction method is not used. In other words, all the experiments were performed on the original images.

In this study, we apply this classifier for its simplicity and we replace the Euclidean metric in this classifier by the l_1-norm based distance metric, which provides better results. The definition of the classification accuracy is as follows: $acc = \frac{n_a}{n}$ where acc is the classification accuracy, n is the number of all the test samples, and n_a denotes the number of all the samples, which are correctly classified.

As shown in Fig. 1, we first acquire the left-half face vector and right-half face vector of an original face image, and then use the gradient descent algorithm to obtain the approximately symmetrical face image and the number of the maximum iteration is set to 30.

The implementation details of the compared methods are introduced. For the gradientfaces based method, i.e. GRF, the smoothing step smoothed the input image by convolving it with the Gaussian kernel function, and the standard deviation parameter of the convolution operator σ was set to 0.75. For LWT, the value of the decomposition level of the wavelet transform was set to three as suggested by the original literature [33]. That is, three-level wavelet decomposition was implemented for each image. For NLM, as suggested in [36], the decay parameter h of the exponential function was set to 80 and the size of the patches to be used in the non-local means algorithm, i.e. the size of the local neighborhood, was set to 2×2. For the DCT normalization method, the number of DCT coefficients to be set to zero was defined as 20. For the DoG features based method, the standard deviations of the inner and outer Gaussian functions used in the DoG filter were respectively set to 1 and 2, which typically produced the best results [27]. For LBPHF, the magnitude of the neighboring points and a radius of the neighborhood was respectively set to 24 and 3, which were recommended for a good result [48]. All the compared algorithms were implemented with suggested parameter settings.

4.3. Experiments and analysis

Tables 3, 4, 5, 6 summarize the classification accuracies of different preprocessing methods on the FERET, GT, ORL and LFW databases, respectively. The first column of each table denotes the algorithms for comparison and the first row of each table is the number of training samples per class. As can be seen, in these tables, our method is superior to the other approaches and can achieve the highest classification accuracy. For example, for the experimental results on the FERET face dataset shown in Table 3, when the first 2 face images of each subject were treated as training samples and the remaining face images were used as test samples, the proposed method can achieve the best classification accuracy of 72.50%, whereas the second best method, i.e. GRF, only gets a classification accuracy of 62.10%. Thus, the proposed preprocessing method has apparent advantage and achieves much more accurate classification than the other methods on this dataset.

As shown in Table 4, the proposed method can obtain the best performance on the GT database in comparison with the other methods. For example, when the number of the training samples per class is set to 5, the classification accuracy of our method is 63.60% and the recently proposed method TT only obtains an accuracy of 46.00%. The classification accuracy of the proposed method again verifies that the proposed method can perform very well in preprocessing face images.

Moreover, the experimental results on the ORL database presented in Table 5 again show that the proposed method always can obtain the best classification performance and our method consistently outperforms all other preprocessing methods. Especially, when the number of the training samples per class is 5, our method can achieve an accuracy of 91.5%, which is the highest classification accuracy among all methods. The classification results on the LFW database presented in Table 6 also demonstrate that the proposed method outperforms all other methods. The performance of the well-known LTV method is relatively bad on the LFW dataset, and the possible reason is that the LTV model more severely suffers from the complex background of images and variations of poses and facial expressions. In contrast, our method is not seriously influenced by these data uncertainties.

Generally, there is no doubt that our method is very efficient. For each iteration, the major computational burden of our method is plus and minus of vectors. So the computational complexity of our method is O(n) where n is the number of image pixels. Moreover, our method always can converge around 30 times for different image datasets with varying image sizes. So, our method can be efficiently applied to perform image preprocessing for larger images. Because our method has computational efficiency, it is absolutely suitable for real-time online classification. To demonstrate that the proposed method is more computationally efficient in comparison with other methods, we conducted some experiments on the ORL face database. We respectively applied different image preprocessing methods to the whole image dataset and the average computation time is summarized in Table 7. We can see that the average processing speed of the proposed method is lower than all the compared methods. Moreover, from Table 5, we see that our method can achieve the highest classification results on the ORL dataset. So the proposed method has great superiorities in face image preprocessing. Thus, it is demonstrated that our method not only is computationally efficient but can achieve better preprocessing results in comparison with the state-of-the-art face image preprocessing algorithms.

Table 3
Classification accuracies (%) of different methods on the FERET face database.

Training samples per class	1	2	3	4	5
Our method	**56.42**	**72.50**	**74.63**	**78.00**	**80.00**
LSSF	37.58	49.30	46.63	65.67	74.75
GRF	50.50	62.10	55.38	69.83	78.75
LWT	46.08	58.40	55.00	74.17	79.25
LTV	10.33	15.30	18.13	19.83	25.00
NLM	46.00	57.30	56.00	70.67	80.00
DCT	30.50	39.90	40.38	52.17	65.50
LBPHF	34.90	41.80	40.60	49.33	50.50
TT	45.33	59.10	55.63	73.67	78.75
DoG	41.33	56.50	52.38	65.17	78.25
MSW	39.25	53.70	50.00	66.67	78.25

Table 4
Classification accuracies (%) of different methods on the GT face database.

Training samples per class	1	2	3	4	5
Our method	**44.71**	**56.00**	**58.50**	**62.73**	**63.60**
LSSF	18.86	25.54	28.50	32.91	37.60
GRF	20.86	29.54	33.67	36.55	40.60
LWT	22.86	29.08	32.67	35.82	38.40
LTV	32.14	43.85	51.33	55.09	59.00
NLM	29.43	38.92	40.17	43.82	47.40
DCT	13.43	20.46	23.17	24.55	26.80
LBPHF	32.71	37.53	37.83	42.54	43.00
TT	24.57	35.08	37.83	41.64	46.00
DoG	21.00	30.62	33.50	38.91	44.00
MSW	23.57	32.62	36.50	40.91	44.00

Table 5
Classification accuracies (%) of different methods on the ORL face database.

Training samples per class	1	2	3	4	5
Our method	**72.22**	**86.25**	**89.64**	**89.17**	**91.50**
LSSF	49.72	63.75	65.71	69.17	70.00
GRF	61.94	73.13	76.79	82.08	80.50
LWT	60.28	74.36	75.71	80.83	81.00
LTV	70.00	81.88	83.21	86.67	88.50
NLM	63.61	78.44	82.86	86.25	85.00
DCT	63.06	76.88	80.36	84.58	84.00
LBPHF	62.50	70.93	72.14	80.00	86.50
TT	67.50	79.69	82.86	87.08	88.50
DoG	53.89	70.63	72.50	77.08	76.00
MSW	52.78	66.25	66.79	75.42	74.00

Table 6
Classification accuracies (%) of different methods on the LFW face database.

Training samples per class	4	5	6	7	8
Our method	**23.93**	**27.53**	**28.57**	**30.35**	**31.08**
LSSF	15.99	18.88	19.18	20.80	22.38
GRF	19.40	21.32	21.50	24.50	25.58
LWT	20.73	22.90	24.22	26.50	27.89
LTV	9.60	9.87	11.39	12.45	13.61
NLM	18.96	19.98	20.54	23.42	22.73
DCT	11.58	12.67	13.20	15.25	14.03
LBPHF	7.17	7.19	8.30	9.09	8.88
TT	17.75	19.12	19.86	22.03	23.98
DoG	21.28	22.90	23.95	26.50	27.89
MSW	18.63	21.32	22.45	25.58	26.29

5. Experimental evaluation on the generated virtual dictionary

The goal of the experiments in this section is to test the effectiveness of the virtual dictionary generated using our method.

Table 7
Average image preprocessing speed on the ORL database.

Method	Time costs (MS)
Our method	**2.0**
LSSF	29.6
GRF	3.9
LWT	13.2
LTV	15.5
NLM	79.5
DCT	6.4
LBPHF	2.7
TT	4.4
DoG	4.6
MSW	26.0

Our method is first applied to the original training samples to obtain the virtual training samples, and then it constructs the virtual dictionary by stacking these virtual sample vectors to form a matrix. The original test sample is not processed such that our method is similar to the dictionary learning method in face recognition. Because the dictionary learning based face recognition methods also exploit only the original training samples to obtain a dictionary and use this dictionary to perform classification, they are used as comparing methods. We applied our method to all the training samples from the CMU PIE [50], FERET [44], Georgia Tech (GT) [46], ORL [45], Yale [53] and extended YaleB face databases [41], to acquire respective virtual training dictionary, and then used the sparse representation based classification (SRC) algorithms such as l1_ls [54], FISTA [55] and Homotopy [56] to obtain the final classification results of the test samples by exploiting only the virtual training dictionary. We compared our method with several state-of-the-art dictionary learning methods such as DKSVD [58], LCKSVD [57] and metaface learning (MFL) [59], and some methods were proposed to generate virtual samples. For example, symmetrical face training samples (SFTS) [11], inter-class relationship (ICR) [13], random noise based virtual sample (RNVS) [14] and mirror face training samples (MFTS) [12] were all proposed for face recognition. Especially, in order to demonstrate that the synthesized virtual samples are better than the original samples, we implement different SRC algorithms based on the original dataset as a comparison task, which is denoted as “original dataset”.

5.1. Implementation details

In this subsection, we provide the implementation details of the compared methods. The parameter setting of our method is the same as that in Section 4.2.

It is demonstrated that synthesizing a virtual dictionary is an effective way to reduce the uncertainty of the face [3]. In order to measure the performance of different methods, we perform classification by applying different SRC algorithms, i.e. *l*1_*ls*, homotopy and FISTA, to the synthesized virtual dictionaries and the dictionaries obtained using dictionary learning methods.

As one of the most significant applications of the sparse representation, dictionary learning has attracted considerable attention. Here we employed some representative dictionary learning methods as the compared dictionary learning methods. As for dictionary learning methods DKSVD and LCKSVD, the sizes of the dictionaries are set to the same as the size of the training sample set. In order to evaluate the performance of the learned dictionaries, we first implemented the original dictionary learning methods to obtain the learned dictionaries, and then exploited different SRC algorithms and the learned dictionaries to obtain classification results of the test samples. For example, as for the dictionary learning method LCKSVD, we first obtained the learned dictionary, and then replaced the OMP algorithm in the classification procedure by other different sparse representation algorithms with the l_1-norm regularization such as *l*1_*ls*, FISTA and homotopy, to obtain the representation coefficients and do not change the original classification scheme. LCKSVD can achieve better classification results by replacing OMP with the sparse representation algorithms with the l_1-norm regularization, which is also suggested as an alternative strategy in the original paper of LCKSVD [57]. The scalar parameter in DKSVD is all set to 0.01, which is the best value obtained using the cross validation on all training data. The sparsity factor in LCKSVD is set to 0.05 times of the size of the training sample set and the parameters α and β are all set to 0.01 and 0.001 by n-fold cross validation on all the training data, respectively. As for the MFL method for dictionary learning, the number of the metafaces of each class is set to 4 as suggested in [59].

The regularization parameter λ of each sparse representation algorithm is determined by adjusting it to get the best performance on each database. For $l1_{ls}$ algorithm, the optimal value of the regularization parameter λ is 0.01 for all databases except for the ORL and FERET databases, where the optimal value of λ is 0.05. Without loss of generality, the regularization parameter λ for all other SRC algorithms is set to 0.01.

5.2. Experimental results on the CMU PIE face database

The CMU PIE face database contains 41,368 face images from 68 subjects, and the images of each person were captured under 13 different poses, 43 different illumination conditions, and 4 different expressions. As described in the literature [51,52], preprocessing to locate the faces was applied. Original images were normalized (in scale and orientation) such that the two eyes were aligned at the same position and the facial areas were cropped into the final images for matching. We chose two subsets of the CMU PIE database to perform experiments for evaluating the performance of different methods.

In the first experiment, we selected an illumination variation subset, where both of the pose and expression were fixed, to evaluate the performances of different methods under varying illumination conditions. This subset contains 1428 frontal face images of 68 people and each people has 21 face images under different lighting conditions [51], denoted by *CMU PIE-*1 *subset*. Each image is cropped to 32 by 32 pixels. Fig. 7 shows some cropped samples from the CMU PIE face database. We took the first 1, 2, 3, 4 and 5 face images of each subject as training samples and used the rest of the samples as test samples. Fig. 8 respectively shows the classification accuracies of all methods with three different SRC algorithms. In summary, the original training sample set and the virtual dictionaries obtained by seven different compared methods and our method were evaluated by using different SRC algorithms. From the classification results shown in Fig. 8, one can see that our method in most cases can achieve higher classification accuracies than LCKSVD, DKSVD, SFTS, MFTS, ICR, MFL and RNVS.

For the second subset of the CMU PIE database, the main purpose of the experiments is to evaluate the performances of different algorithms under the conditions of lighting and expression variations with fixed pose. Thus, the images are captured under different lighting and expression conditions and only the pose is fixed. This subset has 2856 face images of 68 subjects and each subject has 42 face images at Pose 27 under different illumination conditions and different expressions [52]. This subset is denoted as *CMU PIE-*2 *subset*. All the images have been aligned based on the eye position and cropped to 32 by 32 pixels. We also treated the first 1, 2, 3 and 4 images of each subject as original training samples and regarded the remaining images as test samples. The classification result curves are depicted in Fig. 9. Thus, one can see that our method in most cases can achieve better results than DKSVD, LCKSVD, MFL and other methods. We can draw a

Fig. 7. Some images from the CMU PIE database.

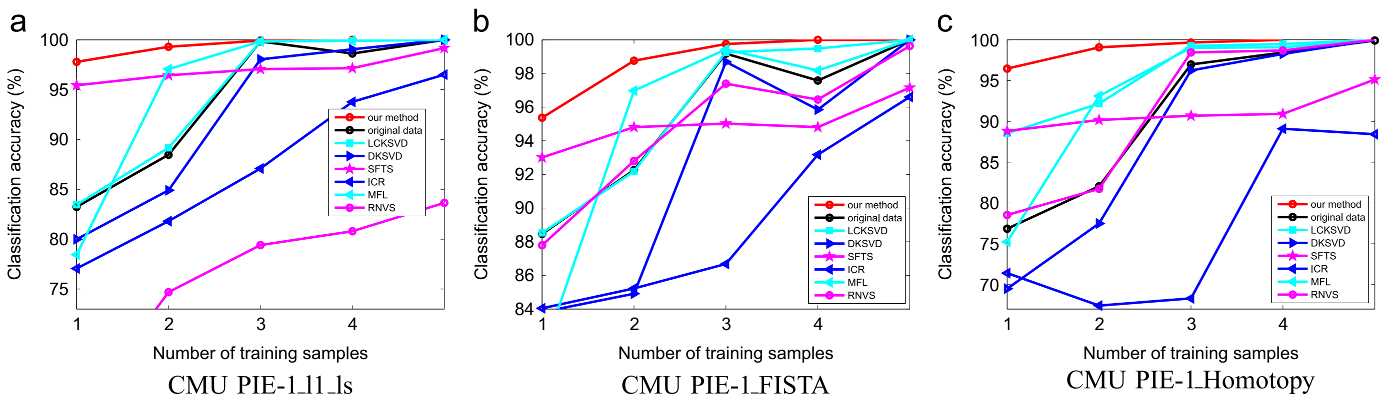

Fig. 8. Classification accuracies (%) of different methods on the CMU PIE-1 subset using different SRC algorithms. (a)–(c) are the experimental results evaluated by l1_ls, FISTA and homotopy, respectively.

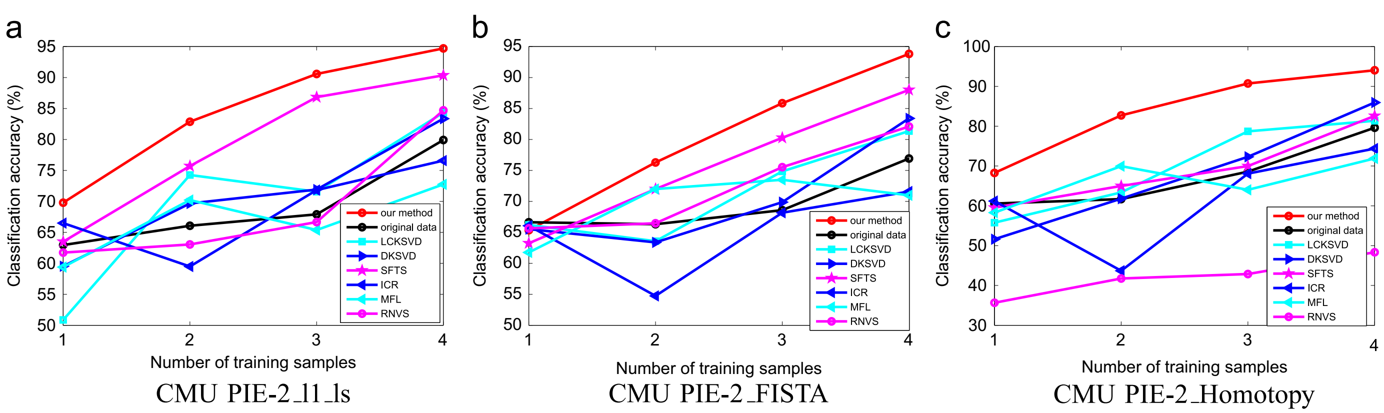

Fig. 9. Classification accuracies (%) of different methods on the CMU PIE-2 subset using different SRC algorithms. (a)–(c) are the experimental results evaluated by l1_ls, FISTA and homotopy, respectively.

conclusion that the proposed method can achieve the maximum classification accuracy in comparison with other methods.

5.3. Experimental results on the FERET face database

The detailed description of the FERET face database can be found in Section 4.1. We also used the subset presented in Section 4.1 to perform experiments. We respectively selected the first 1, 2, 3 and 4 face images of each subject as original training samples and used the remaining face images as test samples. Fig. 10 shows the classification accuracy curves of different SRC algorithms on the dictionaries obtained using different methods. The experimental results show that in most cases our method leads to much more accurate classification than all other methods.

5.4. Experimental results on the GT face database

The detailed description of the GT face database can be found in Section 4.1. We respectively selected the first 1, 2, 3 and 4 face images of each subject as original training samples and used the remaining face images as test samples. Fig. 11 shows again that the virtual dictionary obtained using the proposed method is able to produce better classification performance in comparison with the other methods. For instance, when we take the first two images of each subject as training samples and employ the *l1_ls* algorithm to classify the test samples, which are the rest of the images, the classification accuracy of the proposed method is 56.15%, whereas the classification result obtained using the original data is only 49.85%. Thus, the synthesized images contain more discriminative information in comparison with the original data. Moreover, compared to the dictionary learning algorithms, our method is really efficient because our method is based on very simple numerical computations. Experimental results presented in Fig. 11 demonstrate that our method can be applied to improve different SRC algorithms. It is clear that the accuracy improvement of our method is significant and our method outperforms the other methods.

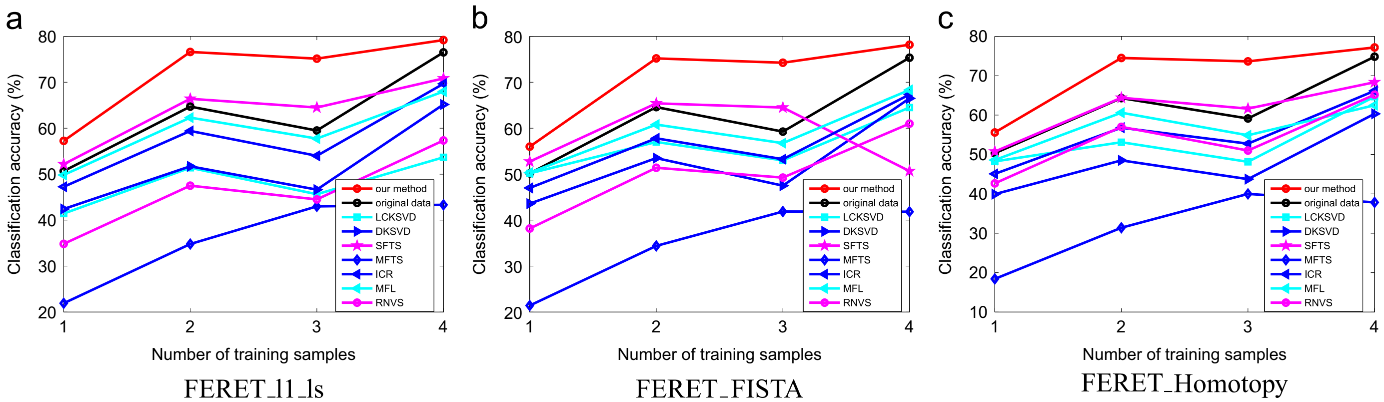

Fig. 10. Comparison of different methods on the FERET database using different SRC algorithms. (a)–(c) are the experimental results evaluated by l1_ls, FISTA and homotopy, respectively.

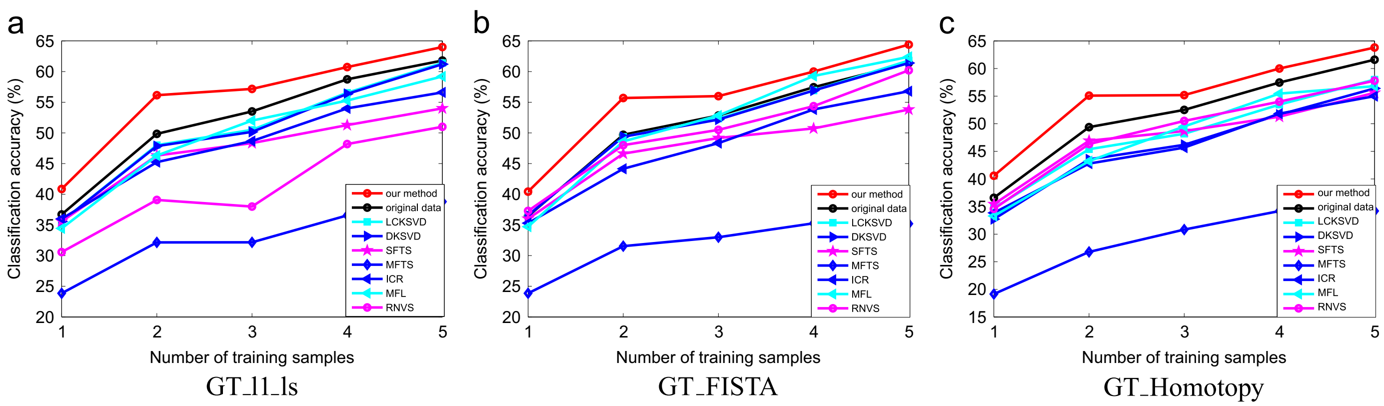

Fig. 11. Comparison of different methods on the GT database using different SRC algorithms. (a)–(c) are the experimental results evaluated by l1_ls, FISTA and homotopy, respectively.

5.5. Experimental results on the ORL face databasee

The description of the ORL database can be found in Section 4.1. We respectively selected the first 1, 2, 3 and 4 face images of each subject as training samples and treated the remaining face images as test samples. The experimental results have been summarized in Fig. 12. It verifies that the proposed method is able to obtain satisfactory classification results and also can perform better than the compared methods. The fact that the classification accuracy of the well-known dictionary learning method, i.e. LCKSVD, is lower than that of our method also demonstrates that our method has a better advantage in face recognition.

5.6. Experimental results on the Yale and Extended YaleB face database

The final experiment was performed on the Yale and extended YaleB databases to test the performance of the proposed method. The Yale face database contains 165 grayscale images of 15 subjects. There are 11 images per subject, and each person includes variations in facial expressions and facial details. All images were simply cropped and resized to 80×100 pixels. In our experiments, we took the first 1, 2, 3 and 4 images of each subject as training samples and regarded the remaining images as test samples. The classification accuracies of different SRC algorithms are presented in Fig. 13. We can see that the proposed method almost performs better than these dictionary learning methods.

Compared with the Yale database, the extended YaleB database has more illumination changes. The extended YaleB face database contains 2432 frontal-face images of 38 individuals. We used the cropped and unnormalized face images with the size of 84×96 pixels which were collected under various illumination conditions for our experiments. Some samples from the extended YaleB dataset are presented in Fig. 14. In our experiments, we treated the first 1–8 images of each subject as training samples and took the remaining images as test samples. The classification accuracy curves obtained using different SRC algorithms on different virtual or learned dictionaries and original dataset are presented in Fig. 15. It is obvious that the proposed method achieves the best classification accuracy in comparison with the other methods.

6. Limitations and discussion

In this paper, we proposed an efficient and effective method to perform face image preprocessing and to synthesize a virtual dictionary for image classification. Our method directly utilizes the geometrical midline of the face image to separate it into two-half images with the same size and exploits both of the left-half and right-half images of the face image to obtain a symmetrical face image for image preprocessing and classification. The general framework of generating an approximately axis-symmetrical face image has been presented in Fig. 2.

In our paper, we regard the geometrical midline of the face image as the symmetry line of the face to iteratively generate the approximately axis-symmetrical face images. If the face image is strictly symmetric, the true symmetry line of the face is exactly the geometrical midline of the image. On the other hand, for face recognition applications especially applications in video surveillance, we should consider the problem that face images captured are usually not

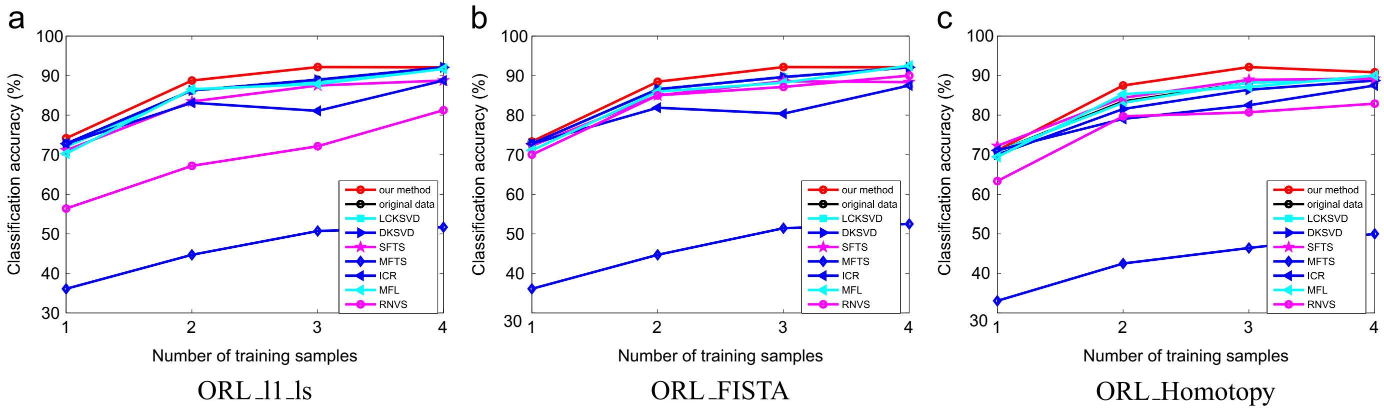

Fig. 12. Comparison of different methods on the ORL database using different SRC algorithms. (a)–(c) are the experimental results evaluated by l1_ls, FISTA and homotopy, respectively.

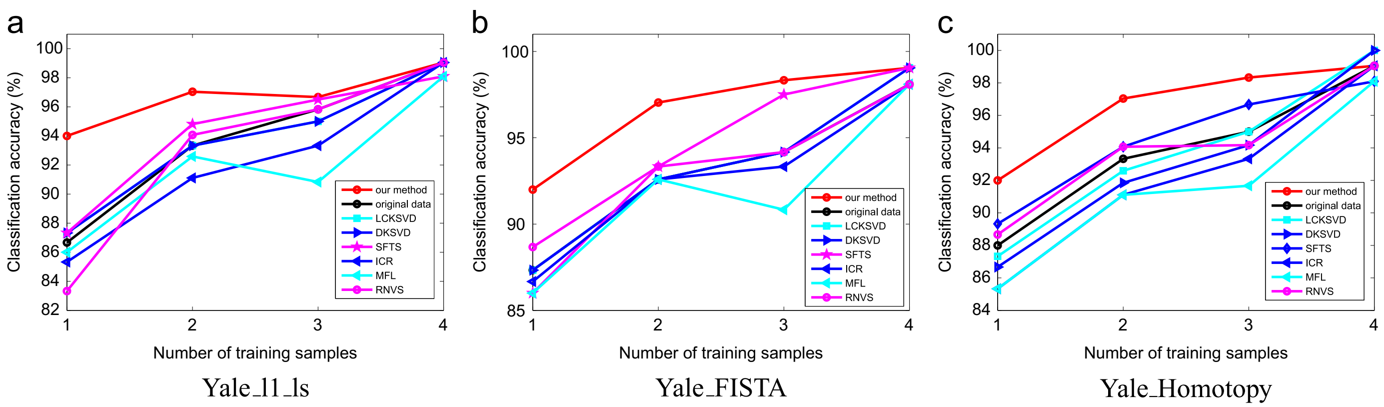

Fig. 13. Comparison of different methods on the Yale database using different SRC algorithms. (a)–(c) are the experimental results evaluated by l1_ls, FISTA and homotopy, respectively.

symmetric, and then the geometrical midline of the face image is not the exact symmetry line of the face. In this case, the virtual images synthesized using our method may look unnatural or little ugly when the original images have pose variations with large degrees as shown on the right-side four images of Fig. 5h. So the main limitation of our method is that our method cannot generate very satisfactory images in this case, but our method can effectively process the frontal face images under the varying lighting and expression situations, and even for face images with slight pose variation.

One may think that the accurate symmetry line is very important to generate an approximately axis-symmetrical face image. If the face image is separated into left-half and right-half face images based on the imprecise symmetry line of the face, the classification accuracy may be influenced. However, all these concerns are not definitely necessary for numerical operations on electronic images. Although the obtained virtual face image may look a little unnatural, blurred or ugly, our method really can improve the classification accuracy. This is because that all the face images can be transformed to approximately axis-symmetrical face images such that the influences of pose variations can be somewhat alleviated. It does not matter that the used geometrical midline of the face image is not strictly the same as the true symmetry line of the face, because the whole image information is still preserved and the proposed method just converts the face images into approximately axis-symmetrical face images based on numerical computations. The virtual image obtained using our method not only can increase the similarities between images from the same class but also can enhance the differences between different classes. So the proposed method is meaningful and also can achieve better performance even for not strictly aligned face images. Furthermore, though our method would not be effective for face images that have high position variance, i.e., are far from being frontal images and this would significantly reduce the usefulness of our method, the proposed method has been verified by extensive well-known face databases, which also demonstrates that our method is flexible and feasible in different conditions and even in large-scale applications.

It is worth noting that our method can be used with various other classifiers. To evaluate the performance of our method in image preprocessing, here we only use the nearest neighbor (NN) classifier with the l_1-norm based distance metric to make image classification, but any other image classifier such as SVM and NN classifier with Euclidean metric can be employed. On the other hand, for the experimental evaluation on the generated virtual dictionary, we use three well-known sparse representation based classification algorithms to demonstrate the effectiveness of our method. Thus, extensive experiments have definitely shown the flexibility of our method.

More importantly, to our knowledge, no similar algorithm has been proposed in the past. Besides, our method can be used to produce symmetrical face images to obtain good image preprocessing results and to lead to high face recognition accuracy, it can also be applied to other similar problems such as classification of other axis-symmetrical objects. Another remarkable advantage of our method is that it is computationally efficient. Furthermore, our studies also imply that in the vision-based recognition of humans, the difference of the two halves of the face image may be ignored, and for automatic face recognition, the reduction of this difference is also beneficial to more accurate recognition of faces.

Fig. 14. Some images from the Extended YaleB database.

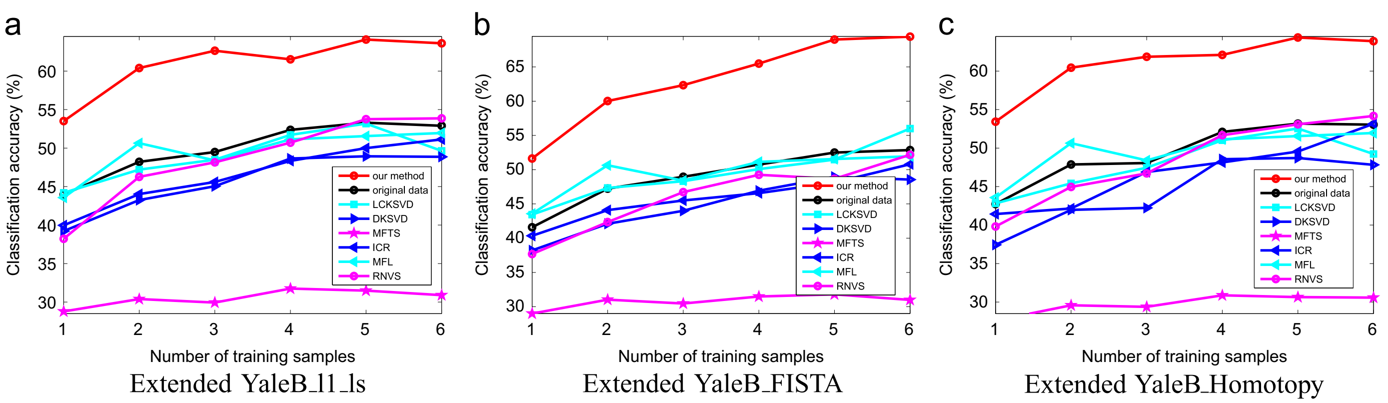

Fig. 15. Comparison of the original database on the Extended YaleB database using different SRC algorithms. (a)–(c) are the experimental results evaluated by l1_ls, FISTA and homotopy, respectively.

7. Conclusions

Our work brings a novel viewpoint and way to automatically produce approximately axis-symmetrical virtual face images and to explore them for better recognition of faces. It seems that no similar works are presented in the past. The designed method is not only novel and simple but also is very competent. Moreover, after approximately axis-symmetrical virtual face images are obtained, an arbitrary classification method or classifier can be applied for recognition of faces. This means that what we propose in this work is a framework with flexibility. Numerous experiments on multiple datasets and different sparse representation algorithms comprehensively demonstrate the good performance in face recognition of our framework.

The proposed face image preprocessing method doubtlessly shows that producing approximately axis-symmetrical face images from the original face images is beneficial to better recognition of faces. The main reason is that the produced approximately axis-symmetrical face images are partially immune to varying face appearances, which usually have severe negative effects on recognition of faces. It is also notable that the designed preprocessing method can automatically produce approximately axis-symmetrical virtual face images and can also well address the problem of heterogeneous illuminations of frontal-face images. Actually, this preprocessing method can generate good neutral and normal face images from naive face images with heterogeneous illuminations. Images obtained using the proposed method can achieve very good visual effects and different classification methods implemented on the obtained images always result in higher classification accuracy. On the other hand, the proposed method can efficiently generate more effective virtual dictionary in comparison with state-of-the-art dictionary learning methods. The experimental results on several different face databases also demonstrate that the virtual dictionary produced by the proposed method is very effective.

Our method is validated by the results of two main applications, i.e. face image preprocessing and generating of the virtual dictionary. We demonstrate that producing facial geometry based axis-symmetrical face images is greatly meaningful. One of our future studies is to extend our idea and approach to solve the face alignment problem. In particular, we will apply the property of facial geometrical symmetry to design a flexible approach to perform face pose estimation.

Conflict of interest

There is no conflict of interests.

Acknowledgments

This paper is partially supported by the National Natural Science Foundation of China (Nos. 61370163, 61233011, 61332011, 61271344), Shenzhen Municipal Science and Technology Innovation Council (Nos. JCYJ20130329151843309 and JCYJ20140904154630436), Shenzhen Fundamental Research Fund (JCYJ20140508160910917 and JCYJ20150403161923528).

References

[1] W. Deng, J. Hu, J. Lu, J. Guo, Transform-invariant PCA: a unified approach to fully automatic facealignment, representation, and recognition, IEEE Trans. Pattern Anal. Mach. Intell. 36 (2014) 1275–1284.

[2] H.T. Ho, R. Chellappa, Pose-invariant face recognition using Markov random fields, IEEE Trans. Image Process. 22 (2013) 1573–1584.

[3] Y. Xu, X. Fang, J. Yang, J. You, H. Liu, S. Teng, Data uncertainty in face recognition, IEEE Trans. Cybern. 44 (2014) 1950–1961.

[4] Y. Wang, J. Liu, X. Tang, Robust 3D face recognition by local shape difference boosting, IEEE Trans. Pattern Anal. Mach. Intell. 32 (2010) 1858–1870.

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

[6] X. Chen, J. Yang, D. Zhang, Complete large margin linear discriminant analysis using mathematical programming approach, Pattern Recognit. 46 (2013) 1579–1594.

[7] J. Gui, T. Liu, D. Tao, Z. Sun, T. Tan, Representative vector machines: a unified framework for classical classifiers, IEEE Trans. Cybern. (2016). http://dx.doi.org/10.1109/TCYB.2015.2457234.

[8] Y. Tao, J. Yang, Enhanced iterative projection for subclass discriminant analysis under EM alike framework, Pattern Recognit. 47 (2014) 1113–1125.

[9] W. Gao, S. Shan, X. Chai, X. Fu, Virtual face image generation for illumination and pose insensitive face recognition, in: Proceedings of the ICME, 2003.

[10] Z. Zhang, L. Wang, Q. Zhu, Z. Liu, Y. Chen, Noise modeling and representation based classification methods for face recognition, Neurocomputing 148 (2015) 420–429.

[11] Y. Xu, X. Zhu, Z. Li, G. Liu, Y. Lu, H. Liu, Using the original and 'symmetrical face' training samples to perform representation based two-step face recognition, Pattern Recognit. 46 (2013) 1151–1158.

[12] Y. Xu, X. Li, J. Yang, Z. Lai, D. Zhang, Integrating conventional and inverse representation for face recognition, IEEE Trans. Cybern. 44 (2014) 1738–1746.

[13] Q. Li, J.H. Wang, J. You, et al., Enlarge the training set based on inter-class relationship for face recognition from one image per person, Plos One 8 (2013) e68539.

[14] D. Tang, N. Zhu, F. Yu, et al., A novel sparse representation method based on virtual samples for face recognition, Neural Comput. Appl. 24 (2014) 513–519.

[15] C. Sanderson, K.K. Paliwal, Noise compensation in a person verification system using face and multiple speech feature, Pattern Recogn. 36 (2003) 293–302.

[16] M.C. Su, C.H. Chou, Application of associative memory in human face detection, in: Proceedings of the International Conference on Neural Networks, vol. 5, 1999, pp. 3194–3197.

[17] S. Saha, S. Bandyopadhyay, A symmetry based face detection technique, in: Proceedings of the IEEE WIE National Symposium Emergency Techniques, 2007.

[18] E. Saber, A.M. Tekalp, Frontal-view face detection and facial feature extraction using color, shape and symmetry based cost functions, Pattern Recognit. Lett. 19 (1998) 669–680.

[19] W.Y. Zhao, R. Chellappa, Illumination-insensitive face recognition using symmetric shape-from-shading, in: Proceedings of the CVPR, 2000.

[20] Y.J. Song, Y.G. Kim, U.D. Chang, H.B. Kwon, Face recognition robust to left/right shadows; facial symmetry, Pattern Recognit. 39 (2006) 1542–1545.

[21] G. Passalis, P. Perakis, T. Theoharis, I.A. Kakadiaris, Using facial symmetry to handle pose variations in real-world 3D face recognition, IEEE Trans. Pattern Anal. Mach. Intell. 33 (2011) 1938–1951.

[22] J. Harguess, J.K. Aggarwal, Is there a connection between face symmetry and face recognition? in: Proceedings of the CVPRW, 2011.

[23] L. Zhang, A. Razdan, G.E. Farin, J. Femiani, M. Bae, C. Lockwood, 3D face authentication and recognition based on bilateral symmetry analysis, Vis. Comput. 22 (2006) 43–55.

[24] A.K. Singh, G.C. Nandi, Face recognition using facial symmetry, in: Proceedings of the Conference on Computer Science Engineering Inference Technology, 2012.

[25] A.C. Little, Domain specificity in human symmetry preferences: symmetry is most pleasant when looking at human faces, Symmetry 6 (2014) 222–233.

[26] Q. Liao, X. Jin, W. Zeng, Enhancing the symmetry and proportion of 3D face geometry, IEEE Trans. Vis. Comput. Graph. 18 (2012) 1704–1716.

[27] X. Tan, B. Triggs, Enhanced local texture feature sets for face recognition under difficult lighting conditions, IEEE Trans. Image Process. 19 (2010) 1635–1650.

[28] B. Wang, W. Li, W. Yang, et al., Illumination normalization based on weber's law with application to face recognition, IEEE Signal Process. Lett. 18 (2011) 462–465.

[29] M.J. Tarr, A.S. Georghiades, C.D. Jackson, Identifying faces across variations in lighting: psychophysics and computation, ACM Trans. Appl. Percept. 5 (2008) 10.

[30] H. Han, S. Shan, X. Chen, et al., A comparative study on illumination preprocessing in face recognition, Pattern Recognit. 46 (2013) 1691–1699.

[31] W. Chen, M.J. Er, S. Wu, Illumination compensation and normalization for robust face recognition using discrete cosine transform in logarithm domain, IEEE Trans. Syst. Man Cybern. 36 (2006) 458–466.

[32] T. Chen, W. Yin, X.S. Zhou, et al., Total variation models for variable lighting face recognition, IEEE Trans. Pattern Anal. Mach. Intell. 28 (2006) 1519–1524.

[33] T. Zhang, B. Fang, Y. Yuan, Y.Y. Tang, Z. Shang, D. Li, F. Lang, Multiscale facial structure representation for face recognition under varying illumination, Pattern Recognit. 42 (2009) 252–258.

[34] X. Xie, W.S. Zheng, J. Lai, et al., Normalization of face illumination based on large-and small-scale features, IEEE Trans. Image Process. 20 (2011) 1807–1821.

[35] T. Zhang, Y.Y. Tang, B. Fang, Z. Shang, X. Liu, Face recognition under varying illumination usng gradientfaces, IEEE Trans. Image Process. 18 (2009) 2599–2606.

[36] V. Šrtuc, N. Pavešić, Illumination invariant face recognition by non-local smoothing, in: Proceedings of the Biometric ID Management and Multimodal Communication, 2009.

[37] B. Cyganek, B. Krawczyk, M. Wozniak, Multidimensional data classification with chordal distance based kernel and support vector machines, Eng. Appl. Artif. Intell. 46 (2015) 10–22.

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

[39] R. Basri, D. Jacobs, Lambertian reflectance and linear subspaces, IEEE Trans. Pattern Anal. Mach. Intell. 25 (2003) 218–233.

[40] S.K. Zhou, R. Chellappa, Illuminating light field: image-based face recognition across illuminations and poses, in: Proceedings of the Conference on AFGR, 2004.

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

[43] Z. Zhang, Y. Xu, J. Yang, X. Li, D. Zhang, A survey of sparse representation: algorithms and applications, IEEE Access 3 (2015) 490–530.

[44] P.J. Phillips, H.M. Syed, A. Rizvi, P.J. Rauss, The FERET evaluation methodology for face-recognition algorithms, IEEE Trans. Pattern Anal. Mach. Intell 22 (2000) 1090–1104.

[45] F.S. Samaria, A.C. Harter, Parameterisation of a stochastic model for human face identification, in: Proceedings of the IEEE Workshop on Applications of Computer Vision, 1994.

[46] Georgia Tech Face Database, available on: ⟨http://www.anefian.com/face_reco.htm⟩, 2007.

[47] G.B. Huang, M. Ramesh, T. Berg, E. Learned-Miller, Labeled Faces in the Wild: a Database for Studying Face Recognition in Unconstrained Environments, University of Massachusetts, Amherst, Technical Report 07-49, October 2007.

[48] G. Zhao, T. Ahonen, J. Matas, M. Pietikinen, Rotation-invariant image and video description with local binary pattern features, IEEE Trans. Image Process. 21 (2012) 1465–1467.

[49] S.J. Wang, J. Yang, M.F. Sun, et al., Sparse tensor discriminant color space for face verification, IEEE Trans. Neural Netw. Learn. Syst. 23 (2012) 876–888.

[50] T. Sim, S. Baker, M. Bsat, The CMU pose, illumination, and expression (PIE) database, in: Proceedings of the Conference on AFGR, 2002.

[51] X. He, D. Cai, P. Niyogi, Laplacian score for feature selection, in: Proceedings of the Advances in Neural Information Processing System, 2005.

[52] D. Cai, X. He, J. Han, T.S. Huang, Graph regularized nonnegative matrix factorization for data representation, IEEE Trans. Pattern Anal. Mach. Intell. 33 (2011) 1548–1560.

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

[54] S. Kim, K. Koh, M. Lustig, S. Boyd, D. Gorinevsky, An interior-point method for large-scale l1-regularized least squares, IEEE J. Sel. Top. Sig. Process. 1 (2007) 606–617.

[55] A. Beck, M. Teboulle, A fast iterative shrinkage-thresholding algorithm for linear inverse problems, SIAM J. Imaging Sci. 2 (2009) 183–202.

[56] M.S. Asif, J. Romberg, Dynamic updating for l1 minimization, arXiv preprint arXiv:0903.1443, 2009.

[57] Z. Jiang, Z. Lin, L.S. Davis, Label consistent K-SVD: learning a discriminative dictionary for recognition, IEEE Trans. Pattern Anal. Mach. Intell. 35 (2013) 2651–2664.

[58] Q. Zhang, B. Li, Discriminative K-SVD for dictionary learning in face recognition, in: Proceedings of the Conference on CVPR, 2010.

[59] M. Yang, L. Zhang, J. Yang, et al. Metaface learning for sparse representation based face recognition, in: Proceedings of the ICIP, 2010.

Yong Xu was born in Sichuan, China, in 1972. He received his B.S. degree and M.S. degree at the Air Force Institute of Meteorology (China) in 1994 and 1997, respectively. He received the Ph.D. degree in Pattern recognition and Intelligence System at the Nanjing University of Science and Technology (NUST) in 2005. Now, he works at Shenzhen Graduate School, Harbin Institute of Technology. His current interests include pattern recognition, biometrics, machine learning and video analysis.

Zheng Zhang received the B.S. degree from Henan University of Science and Technology and M.S. degree from Shenzhen Graduate School, Harbin Institute of Technology (HIT) in 2012 and 2014, respectively. Currently, he is pursuing the Ph.D. degree in computer science and technology at Shenzhen Graduate School, Harbin Institute of Technology, Shenzhen, China. His current research interests include pattern recognition, machine learning and computer vision.

Guangming Lu received the B.S. degree in electrical engineering, master's degree in control theory and control engineering, and the Ph.D. degree in computer science and engineering from the Harbin Institute of Technology (HIT), Harbin, China, in 1998, 2000, and 2005, respectively. He was a Post-Doctoral Fellow with Tsinghua University, Beijing, China, from 2005 to 2007. He is currently a Professor with the Biocomputing Research Center, Shenzhen Graduate School, HIT. His current research interests include pattern recognition, image processing, and automated biometric technologies and applications.

Jian Yang received the B.S. degree in mathematics from the Xuzhou Normal University in 1995. He received the M.S. degree in applied mathematics from the Changsha Railway University in 1998 and the Ph.D. degree from the Nanjing University of Science and Technology (NUST), on the subject of pattern recognition and intelligence systems in 2002. In 2003, he was a postdoctoral researcher at the University of Zaragoza. From 2004 to 2006, he was a Postdoctoral Fellow at Biometrics Centre of Hong Kong Polytechnic University. From 2006 to 2007, he was a Postdoctoral Fellow at Department of Computer Science of New Jersey Institute of Technology. Now, he is a professor in the School of Computer Science and Technology of NUST. He is the author of more than 80 scientific papers in pattern recognition and computer vision. His journal papers have been cited more than 1600 times in the ISI Web of Science, and 2800 times in the Web of Scholar Google. His research interests include pattern recognition, computer vision and machine learning. Currently, he is an Associate Editor of Pattern Recognition Letters and IEEE Transaction On Neural Networks and Learning Systems, respectively.