Neurocomputing 147 (2015) 307–314

Contents lists available at ScienceDirect

Neurocomputing

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

Noise-free representation based classification and face recognition experiments

Yong Xu [a,b], Xiaozhao Fang [a], Jane You [c], Yan Chen [a,d,*], Hong Liu [e]

[a] Bio-Computing Research Center, Shenzhen Graduate School, Harbin Institute of Technology, Shenzhen, China
[b] Key Laboratory of Network Oriented Intelligent Computation, Shenzhen, China
[c] Department of Computing, The Hong Kong Polytechnic University, Hong Kong
[d] Shenzhen Sunwin Intelligent Corporation, Shenzhen, China
[e] Engineering Lab on Intelligent Perception for Internet of Things, Shenzhen Graduate School, Peking University, Shenzhen, China

ARTICLE INFO

Article history:
Received 1 July 2013
Received in revised form
16 June 2014
Accepted 23 June 2014
Communicated by D. Tao
Available online 30 June 2014

Keywords:
Noise elimination
Representation based classification

ABSTRACT

The representation based classification has achieved promising performance in high-dimensional pattern classification problems. As we know, in real-world applications the samples are usually corrupted by noise. However, representation based classification can take only noise in the test sample into account and is not able to deal with noise in the training sample, which causes side-effect on the classification result. In order to make the representation based classification more suitable for real-world applications such as face recognition, we propose a new representation based classification method in this paper. This method can effectively and simultaneously reduce noise in the test and training samples. Moreover, the proposed method can reduce noise in both the original and virtual training samples and then exploits them to determine the label of the test sample. The virtual training sample is generated from the original face image and shows possible variation of the face in scale, facial pose and expression. The experimental results show that the proposed method performs very well in face recognition.

1. Introduction

As we know, noise extensively exists in the data [1]. In order to better model, predict and classify the data, we should well deal with noise. For pattern classification problems, it is significant to devise a way to resist noise and to improve the robustness of the classifier [2–4].

As shown in literature [5], noise has great influence on the face recognition accuracy. Moreover, the image of a face usually varies with the illumination, pose and facial expression [6–11]. This is indeed a great challenging problem in face recognition. We can treat the difference between the images of the same face as generalized noise. For pattern classification problems, if we can identify and reduce noise, a better result can be obtained. A number of methods have been proposed for noise-free face recognition. For example, it is conceived that not-necessarily orthogonal basis which may reconstruct the data is better than principal component analysis (PCA) in the presence of noise and independent component analysis (ICA) is proposed for face recognition in the presence of noise [12]. A unified framework of subspace is proposed for robust face recognition [13]. The enhanced fisher linear discriminant (EFLD) model is also proposed for overcoming noise in face images [14]. In recent years, the low-rank decomposition is also applied to eliminate noise [15–17].

The recently proposed representation based classification has performed very well in high-dimensional pattern classification problems. Especially, the representation based classification proposed by Wright et al. [18,19], i.e. sparse representation classification (SRC) is viewed as a breakthrough of face recognition. Besides SRC, there have also been a number of other representation based classification methods. For example, representation based classification with the l_2 norm minimization constraint on the solution vector is not only able to obtain a high accuracy but also computationally very efficient. For example, collaborative representation [20], the two-phase test sample representation (TPTSR) method [21], the feature space representation method [22–24] have achieved satisfactory results in face recognition. The recently proposed linear regression classification (LRC) is also a representation based classification with the l_2 norm minimization constraint [25,26]. LRC is closely related with the previously proposed nearest intra-class space (NICS) method [27]. The pattern recognition community has paid much attention to the theoretical foundation of representation based classification and to design new representation based classification algorithms [28–30].

Rationales of representation based classification have been demonstrated from different aspects. Wright et al. considers that the

* Corresponding author at: Bio-Computing Research Center, Shenzhen Graduate School, Harbin Institute of Technology, Shenzhen, China
E-mail address: jadechenyan@gmail.com (Y. Chen).

http://dx.doi.org/10.1016/j.neucom.2014.06.058
0925-2312/

"sparsity" of the representation is very helpful for achieving the high classification accuracy [18,19]. However, Zhang et al. claimed that for representation based classification it is not the "sparsity" but the way to represent and classify the test sample, i.e. collaborative representation that contributes the most to the face recognition performance [20]. Moreover, Yang et al. considered that for SRC "locality" is more significant than "sparsity" because in representation based classification "locality" always leads to "sparsity" but not vice visa [28]. Our studies on representation based classification with the l_2 norm minimization constraint show that the "sparsity" can be achieved by using a simple scheme which is very helpful for identifying the class that the test sample is truly from [21,31,32]. The idea of representation based classification has been applied to improve various methods such as tensor discriminant analysis and eigen-subspace methods [30,33–35]. A number of studies on manifold learning also made notable contributions in designing methods that preserve localities of samples. For example, Laplacian faces [36], semisupervised multiview distance metric learning [37], elastic manifold embedding [38] and adaptive hypergraph learning [39] all address the problem of preserving locality structures of samples from different viewpoints. These methods can be applied to various issues such as object correspondence construction in animation [40] and Cartoon character retrieval [41]. Sparse representation was also integrated with other methods such as the wavelet decomposition for face recognition [42].

It seems that it is crucial to properly model noise in data [43]. Though noise exists in both the training and test samples, the algorithm of conventional representation based classification is established on the basis of the conventional least squares algorithm and it cannot take noise in the training samples into account. This will cause side-effect to the classification accuracy. For face recognition, besides noise from the acquisition stage, the variation of the facial pose and expression [44,45] of the same face can also be viewed as generalized noise.

In this paper, we use the following scheme to improve the representation based classification: we first perform matrix decomposition for the matrix consisting of all the training samples and obtain the approximations of all the training samples, referred to as approximation training samples (ATRS). Then we exploit the ATRS to obtain an approximation of the test sample, referred to as approximation test sample (ATES). Finally, conventional representation based classification (CRBC) is applied to ATES and ATRS and the classification result of the test sample is obtained. Moreover, motivated by the fact the face usually has an axis-symmetrical structure, the proposed method also exploits the original face images to generate virtual symmetrical face images, which are helpful for showing possible variation of the face in scale and pose. Simply speaking, an original face image will generate two virtual symmetrical face images. The left half of the first virtual symmetrical face image is the same as the left half of the original face image and the right half of the first virtual symmetrical face image is just the mirror image of its left half. The right half of the second virtual symmetrical face image is the same as the right half of the original face image and the left half of the second virtual symmetrical face image is just the mirror image of its right half. The proposed method uses both the original and virtual training samples to represent and classify the test sample and outperforms the state-of-art face recognition methods. The main contributions of this work are as follows: (1) it proposes a simple and reasonable way to simultaneously reduce noise in the training and test samples. (2) The designed algorithm can lead to very accurate recognition of faces by properly integrating the original and virtual training samples.

2. The proposed method

In this section we describe the proposed method in detail. Suppose that there are C classes and each class has n training samples. Let $x_1, x_2, \ldots, x_N (n = nC)$ be all the training samples from the first, second,…, and C-th classes, respectively. In other words, $x_1, x_2, \ldots, x_n$ are the n training samples from the first class. $x_{n(i-1)+1}, x_{n(i-1)+2}, \ldots, x_{ni}$ are the n training samples from the i-th class. Let y be the test sample. $x_1, x_2, \ldots, x_N$ and y are all M dimensional column vector. The following context describes the main steps of the algorithm of approximation representation (AAR).

Step 1. Let $x_1, x_2, \ldots, x_N$ be all available training samples. Let $X = [x_1 x_2 \ldots x_N]$ and perform singular value decomposition (SVD) to obtain $X = U \Lambda V^T$, $U = [u_1 \ldots u_d]$, $V = [v_1 \ldots v_d]$, Λ is a diagonal matrix and $diag(\Lambda) = [\lambda_1 \ldots \lambda_d]$, $\lambda_1 \geq \lambda_2 \ldots \geq \lambda_d$. Take $\hat{X} = \sum_{i=1,\ldots,K} \lambda_i u_i v_i^T (K \leq d)$ as the reconstruction of X. The j-th column of $\hat{X}$ i.e. $\hat{x}_j$ is the j-th approximation training sample (ATRS) and is also referred to as reconstruction of the j-th training sample x_j.

Step 2. Establish equation $y = \hat{X}A$ and solve it using $\hat{A} = (\hat{X}^T \hat{X} + \mu I)^{-1} y$. μ is a small positive constant and I is the identity matrix. Treat $\hat{y} = \hat{X}\hat{A}$ as the approximation test sample (ATES).

Step 3. Use the ATES and ATRS to establish equation $\hat{y} = \hat{X}B$ and solve it using $\tilde{B} = (\hat{X}^T \hat{X} + \mu I)^{-1} \hat{y}$.

Step 4. Let $\tilde{B} = [\tilde{b}_1 \ldots \tilde{b}_N]$ and $\hat{X} = [\hat{x}_1 \ldots \hat{x}_N]$. Define $S_j = ||\hat{y} - \sum_{k=1,\ldots,n} \tilde{b}_{n(j-1)+k} \hat{x}_{n(j-1)+k}||_2$ as the distance between the test sample and the j-th class.

The proposed method is implemented for test sample y as follows:

(1) Implement the algorithm of approximation representation (AAR) for the original training samples. The score of the test sample with respect to the j-th class is denoted by S_j^1.
(2) Generate virtual training samples and perform AAR for them. The score of the test sample with respect to the j-th class is denoted by S_j^2.

Virtual training samples are generated using the procedure below, which was first proposed in [31]. Let $x_{i0} \in \Re^{p \times q}$ be the i-th training sample in the form of image matrix. Let y_i^1 and y_i^2 respectively stand for the first and second virtual training samples generated from x_{i0}. The left half columns of y_i^1 is set to the same as that of x_{i0} and the right half columns of y_i^1 is the mirror image of the left half columns of y_i^1. However, the right half columns of y_i^2 is set to the same as that of x_{i0} and the left half columns of y_i^2 is the mirror image of the right half columns of y_i^2. y_i^1 and y_i^2 are also referred to as the first and second 'symmetrical face' training samples. The mirror image S of an arbitrary image R is defined as $S(i,j) = R(i, H-j+1)$, $i = 1, \ldots, G$, $j = 1, \ldots, H$. G and H stand for the numbers of the rows and columns of R, respectively. $S(i,j)$ denotes the pixel located at the ith row and jth column of S. For more details, please refer to [31].

AAR is performed for the set consisting of all the first and second virtual training samples. The score of the test sample with respect to the j-th class is denoted by S_j^2.

(3) The scores generated from the original and virtual training samples are integrated using $d_j = uS_j^1 + vS_j^2$, where u, v are the weights and $u + v = 1$. If $q = \arg\min_j d_j$, then test sample y is assigned to the q-th class.

We summarize the algorithms below. First, the algorithm framework of the proposed method is presented by the algorithm table named "Algorithm 1".

Algorithm 1.

Input: X consisting of all original training samples; test sample y;
Output: The classification result of y.

1. Implement the algorithm of approximation representation (AAR) for X and y to obtain the score of y with respect to the j-th class S_j^1
2. Generate the first virtual training samples $y_1^1, y_2^1, \ldots, y_N^1$ and the second virtual training samples $y_2^2, y_2^2, \ldots, y_N^2$. Construct matrix $Y = [y_1^1 \quad y_1^2 \quad y_2^1 \quad y_2^2 \ldots \quad y_N^1 \quad y_N^2]$.
3. Implement the algorithm of approximation representation (AAR) for Y and y to obtain the score of y with respect to the j-th class S_j^2.

Let $d_j = uS_j^1 + vS_j^2$. If $q = \arg\min_j d_j$, then test sample y is assigned to the q-th class.

Second, the algorithm of approximation representation (AAR) is presented by the algorithm table named "Algorithm 2". For simplicity of presentation, only the algorithm of X and y is described. The algorithm of the algorithm of Y and y has the same procedures.

Algorithm 2.

Input: X and y (or Y and y);
Output: The score of the test sample with respect to the j-th class S_j^1 (or S_j^2).

1. Perform SVD to obtain $X = U\Lambda V^T$, $U = [u_1 \ldots u_d]$, $V = [v_1 \ldots v_d]$, Λ is a diagonal matrix and $diag(\Lambda) = [\lambda_1 \ldots \lambda_d]$, $\lambda_1 \geq \lambda_2 \ldots \geq \lambda_d$. Let $\hat{X} = \sum_{i=1,\ldots,K} \lambda_i u_i {v_i}^T$ $(K \leq d)$.
2. Solve $y = \hat{X}A$ using $\hat{A} = (\hat{X}^T\hat{X} + \mu I)^{-1} y$. Treat $\hat{y} = \hat{X}\hat{A}$ as the approximation test sample (ATES).
3. Solve equation on ATRS and ATES $\hat{y} = \hat{X}B$ using $\tilde{B} = (\hat{X}^T\hat{X} + \mu I)^{-1}\hat{y}$.
   Let $\tilde{B} = [\tilde{b}_1 \ldots \tilde{b}_N]$ and $\hat{X} = [\hat{x}_1 \ldots \hat{x}_N]$. Define $S_j^1 = ||\hat{y} - \sum_{k=1,\ldots,n} \tilde{b}_{n(j-1)+k} \hat{x}_{n(j-1)+k}||_2$.

3. Insight into the proposed method

3.1. Analysis on the method

In complex real-world applications, it seems that both the training and test samples contain noise. As a result, to reduce noise in both the training samples and test sample will be beneficial for recognition of faces. However, the conventional representation based classification (CRBC) is based on the conventional least squares and can take only noise in the test sample into account. Specifically, CRBC assumes that the true test sample can be approximated by a weighted sum of all the training samples. The deviation between the original test sample and weighted sum can be viewed as noise in the original test sample. Because the classification procedure of CRBC is almost not influenced by this deviation and is affected by noise in the original training sample, we say that CRBC takes only noise in the test sample into account.

The proposed method exploits the following strategy to simultaneously reduce noise in the training and test samples: first, SVD is performed for the matrix consisting of all the available training samples and the SVD based reconstruction of the training sample is viewed as noise-free training sample. Then, a weighted sum of all noise-free training samples is used to obtain noise-free test sample. This weighted sum is obtained under the assumption that the weighted sum can best approximate the true test sample.

It seems that CRBC can obtain quite satisfactory performance when the test sample is very similar with the training samples that are from the same class as the test sample. However, in the real-world problem such as face recognition the test sample may be so different from the training sample of the same class that the test sample might be erroneously classified by CRBC. As we know, for face recognition, the difference between the training sample and test sample from the same class can also be viewed as generalized noise and usually reflects the variation in facial pose and expression of the face. By simultaneously reducing noise in the training and test samples, the proposed method is able to greatly eliminate the side-effect of the above difference and to obtain a higher accuracy.

The proposed method is somewhat similar with and different from the combination of the low-rank decomposition and representation based classification. The first step of "Algorithm 2" indeed exploits SVD to perform a low-rank decomposition for the training samples and uses the low-rank components to reconstruct the training samples. However, the proposed method significantly differs from the low-rank decomposition as follows. After the SVD based low-rank decomposition is performed and approximation training samples (ATRS) are obtained, our method also exploits ATRS to generate the approximation test sample (ATES), which allows noise in the test sample to be reduced and will somewhat enable the test sample to be more similar with the training samples from the same class. Fig. 1 shows three test samples (shown in the first row) that are from the ORL face database and correctly and erroneously classified by our method

Fig. 1. Three test samples correctly and erroneously classified by our method and the original collaborative representation (CR) respectively. For each column, the image in the first row shows a test sample and the image in the second row shows one training sample of the subject to which this test sample is erroneously assigned by CR. The third row shows training samples of the subject present in the first row. The first five face images of each subject and the remaining images are used as training and test samples, respectively.

and the original collaborative representation (CR), respectively. In other words, our method assigns all of these three test samples into correct classes, whereas CR fails in correctly classifying all of these three test samples. The first five face images of each subject in the ORL face database and the remaining images are used as training and test samples, respectively. For each column, the image in the first row is a test sample and the image in the second row is a training sample of the subject to which this test sample is erroneously assigned by CR. The third row shows training samples of the subject present in the first row.

3.2. Effect of the virtual training samples

The use of the virtual training sample is significant. The virtual training sample is indeed motivated by the fact that the face usually has a symmetrical structure. These virtual training samples show possible variation of the face in scale and pose. We use the same way shown in [31] to obtain the virtual training sample. For our work and the method proposed in [31], there is the following great difference: in [31] the virtual training sample is directly used to represent and classify the test sample, whereas our method proposed in this paper first reduces noise in virtual training samples and then exploits them to obtain representation and classification of the test sample.

In order to show the difference, in representing the test sample, of the original and virtual training sample, we use Fig. 2 to show six test samples from the FERET face database and the reconstruction face images of the test samples obtained using our method on the original and virtual training samples. The reconstruction face image of a test sample obtained using the original training samples is defined as $Z = matrix(\hat{z})$, $\hat{z} = \tilde{b}_1\hat{x}_1 + \cdots + \tilde{b}_N\hat{x}_N$. $\hat{x}_1, \ldots, \hat{x}_N$ are the approximation training samples (ATRS). Z is converted from $\hat{z}$ and is an image with the same size as the original face image. The reconstruction face image of a test sample obtained using the virtual training samples is defined in the same way except that approximation virtual training samples rather than approximation original training samples are used. The first four face images of each subject are used as original training samples and the remaining are used as original test samples. From Fig. 2, we see that the reconstruction face image of a test sample obtained using the virtual training samples is different from that obtained using the original training samples. As a result, we say that the virtual training samples are helpful in reflecting possible variation of the faces. Though the original test sample is not a symmetrical image, the reconstruction face image of a test sample obtained using the virtual training samples seems to be always a symmetrical image and well reflect the symmetrical structure of the face. On the other hand, the reconstruction face image of a test sample obtained using the original training samples is usually not a symmetrical image. As we know, the method devised in the paper finally exploits the reconstruction face image of the test sample to perform classification, so we say that the original and virtual training samples are very complementary in both representing and classifying the test sample. As a result, to simultaneously use the original and virtual training samples can lead to a better recognition result.

Fig. 3 shows the scores obtained using the original and virtual training samples corresponding to the first test sample from the FERET face database. From this figure, we clearly see that the two kinds of scores are complementary and the correlation coefficient between them does not seem to be high. Thus, to combine the

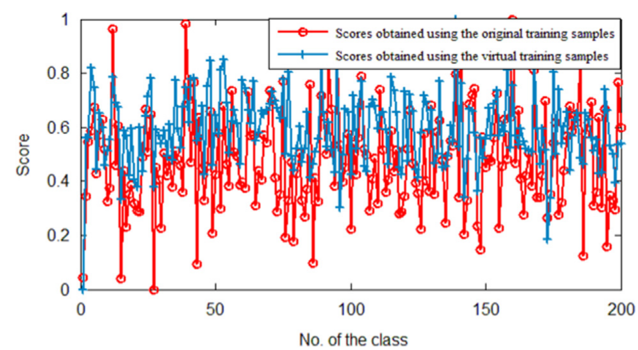

Fig. 3. Scores obtained using the original and virtual training samples corresponding to the first test sample. The first four face images of each subject in the FERET database are used as original training samples and the remaining are used as original test samples.

Fig. 2. Six test samples from the FERET face database and the reconstruction face images of the test samples obtained using our method on the original and virtual training samples. The first row shows the six test samples. The second and third rows show the reconstruction face images of the test samples obtained using the original and virtual training samples, respectively.

original and virtual training samples is helpful for better representing and correctly classify the test sample.

4. Experimental results

We used the ORL and AR face databases as well as a subset of the FERET face database to test our method. We also tested collaborative representation classification (CRC) in [20], coarse to fine face recognition (CFFR) in [32], the improvement to the nearest neighbor classifier (INNC) in [46] and coarse to fine k nearest neighbor classifier (CFKNNC) in [47], feature space-based human face image representation and recognition (FSHFRR) in [24]. When we implemented CFKNNC, we set its parameters n and K to $n=N/2$ and $K=1$. N stands for the total number of all training samples. When FSHFRR was implemented, its parameter σ was set to 10^6. When our method was implemented, we set $K=d-800$. The code of the proposed will be available at http://www.yongxu.org/lunwen.html.

4.1. Experimental results on the ORL face database

The ORL database [48] includes 400 face images taken from 40 subjects each providing 10 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 a 56 by 46 image matrix by using the down-sampling algorithm. We respectively took the first 1, 2 and 3 face images of each subject as original training samples and treated the remaining face images as test samples. Fig. 4 shows face images of three subjects in the ORL database. The experimental results were shown in Table 1. We see that our method can outperform other methods.

4.2. Experimental results on the 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 subjects each providing seven images [49]. 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 2, 3 and 4 face images of each subject as original training samples and treated the remaining face images as test samples. Fig. 5 shows face images of three subjects in subset of the FERET database. Table 2 shows the rates of classification errors of different methods. From this table, we also see that our method can obtain a lower rate of classification errors than other methods.

4.3. Experimental results on the AR face database

For the AR face database [50], the images of each subject have different facial expressions, and were acquired under lighting conditions and with and without occlusions. These images were taken in two sessions. From the AR face database, we used 3120 Gy images from 120 subjects with each subject providing 26 images. We note that 12 face images of each subject are occluded with sunglasses or a scarf. We adopted the 50×40 normalized face image which were manually cropped in [51]. Fig. 6 shows some face images of three subjects in the AR database. From Table 3, we also see that our method outperforms other methods.

4.4. More analysis on the experimental results

Superficially, our proposed method should first properly set its two parameters for obtaining good results. However, from the experimental results, we see that the accuracy of our method does not greatly vary with the parameters when they vary from 0.4 to 0.6. More importantly, if our method simply sets both of the two parameters to 0.5, it can also obtain satisfactory experimental results. This means that in real-world applications, our method can be easily implemented by simply setting each of its two parameters to 0.5.

In order to show how the use of the virtual training samples improves the accuracy of face recognition, we use Tables 4–6 to show the rates of classification errors of AAR using the original training

Table 1
The rate of classification errors of different methods on the ORL face database.

Number of the training samples per subject	u	v	3	2	1
Our method	0.6	0.4	10.00%	12.81%	29.44%
Our method	0.5	0.5	8.93%	13.44%	29.17%
Our method	0.4	0.6	10.71%	14.06%	28.33%
CRC			13.93%	16.56%	31.94%
CFFR			11.07%	14.06%	26.94%
FSHFRR			14.64%	17.50%	29.44%
INNC			17.86%	18.44%	28.33%
CFKNNC			19.29%	17.81%	26.39%

Fig. 4. Face images of three subjects in the ORL database.

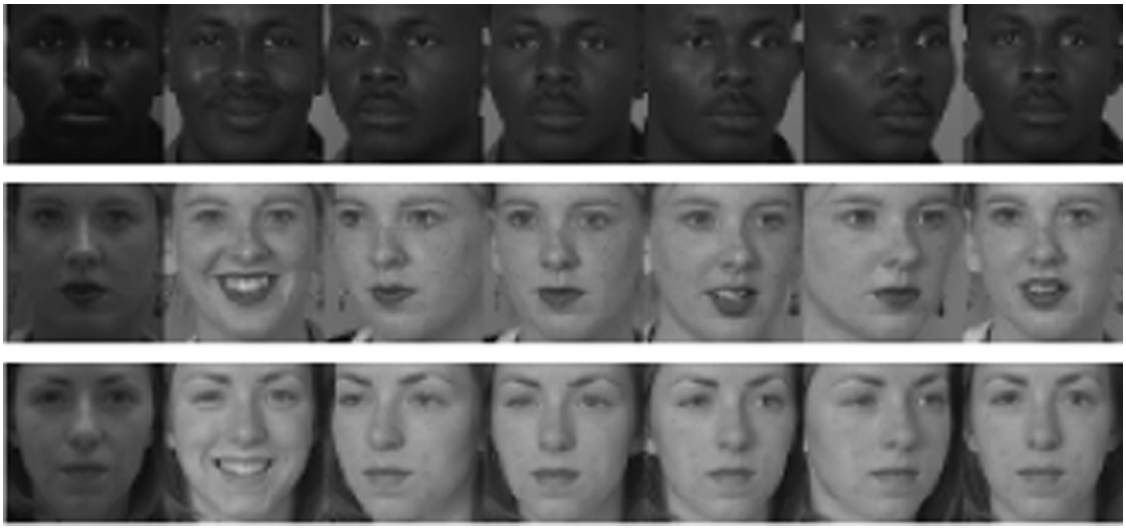

Fig. 5. Face images of three subjects in the FERET database.

Table 2
The rate of classification errors of different methods on the FERET face database.

Number of the training samples per subject	u	v	4	3	2
Our method	0.6	0.4	38.50%	45.00%	37.50%
Our method	0.5	0.5	37.67%	41.25%	36.40%
Our method	0.4	0.6	38.17%	39.50%	36.10%
CRC			44.67%	55.63%	41.60%
FSHFRR			41.67%	50.50%	43.40%
INNC			42.67%	49.50%	41.70%
CFKNNC			38.50%	45.12%	36.70%

Fig. 6. Some face images of three subjects in the AR database.

samples and AAR using the virtual training samples. The rate in the bracket stands for the result of subtracting the rate of classification errors of our method from that of AAR. These tables show that for non-occluded face images, the rate of classification errors might be greatly reduced owing to the use of the virtual training samples. For example, for the FERET face database, the use of the virtual training samples can lead to reduction of the rate of classification errors more than 10% (as shown in Table 5, when the rate of classification errors of our method is 41.25%, the rate of classification errors of AAR on the original training samples is 52.63%).

5. Conclusions

The proposed method uses a simple and feasible way to reduce noise in the test and training samples and exploits the obtained

Table 3
The rate of classification errors of different methods on the AR face database.

Number of the training samples per subject	u	v	4	3	2	1
Our method	0.6	0.4	30.34%	29.60%	29.79%	28.40%
Our method	0.5	0.5	30.57%	30.04%	30.87%	28.97%
Our method	0.4	0.6	31.67%	30.91%	31.74%	30.60%
CRC			32.46%	31.41%	31.67%	30.17%
FSHFRR			48.33%	46.38%	45.42%	43.27%
INNC			34.13%	33.01%	32.99%	30.17%
CFKNNC			34.51%	33.26%	32.15%	29.40%

Table 4
The rates of classification errors of our method and AAR on the ORL face database. $u=0.5$, $v=0.5$.

Number of the training samples per subject	3	2	1
Our method	8.93%	13.44%	29.17%
AAR on the original training samples	13.57% (4.64%)	15.31% (1.87%)	30.56% (1.39%)
AAR on the virtual training samples	16.79% (7.86%)	17.50% (4.06%)	31.67% (2.50%)

Table 5
The rates of classification errors of our method and AAR on the FERET face database. $u=0.5$, $v=0.5$.

Number of the training samples per subject	4	3	2
Our method	37.67%	41.25%	36.40%
AAR on the original training samples	43.50% (5.83%)	52.63% (11.38%)	42.80% (6.40%)
AAR on the virtual training samples	42.50% (4.83%)	40.13% (−1.12%)	37.10% (0.70%)

Table 6
The rates of classification errors of our method and AAR on the AR face database. $u=0.6$, $v=0.4$.

	4	3	2	1
Our method	30.34%	29.60%	29.79%	28.40%
AAR on the original training samples	31.63% (1.29%)	30.43% (0.83%)	30.83% (1.04%)	28.93% (0.53%)
AAR on the virtual training samples	37.88% (7.54%)	37.86% (8.26%)	38.54% (8.75%)	36.97% (8.57%)

"uncontaminated" samples to perform representation based classification. The proposed method is very suitable for complex real-world applications such as face recognition, in which it is almost impossible for conventional representation based classification to completely eliminate the side-effect on classification of noise. However, the proposed method simultaneously reduces noise in the test and training samples by using a way compatible with the nature of representation based classification. The method is not only easy to implement but also achieves very good performance.

Acknowledgments

This article is partly supported by NSFC under Grant nos. 61370163, 61300032 and 61332011, as well as the Shenzhen Municipal Science and Technology Innovation Council (Nos. JC201005260122A, CXZZ20120613141657279, JCYJ201206131533-52732, JCYJ20130329151843309 and JCYJ20130401152508661).

References

[1] S.V. Vaseghi, Advanced Digital Signal Processing and Noise Reductions, second ed., John Wiley & Sons Ltd., 2000.
[2] B. Conrey, J.M. Gold, Pattern recognition in correlated and uncorrelated noise, J. Opt. Soc. Am. A Opt. Image Sci. Vis. 26 (11) (2009) B94–109, http://dx.doi.org/10.1364/JOSAA.26.000B94.
[3] K. Fukushima, Increased robustness against background noise: pattern recognition by a neocognitron, in: Proceedings of the International Conference on Neural Information Processing (ICONIP), vol. 2, 2010, pp. 574–581.
[4] B. Fan, G. Zhang, H.-X. Li, Multiple models fusion for pattern classification on noise data, in: Proceedings of the 2012 International Conference on System Science and Engineering, June 30–July 2, 2012, Dalian, China.
[5] W.-Y. Zhao, R. Chellappa, P.J. Phillips, A. Rosenfeld, Face recognition: a literature survey, ACM Comput. Surv. 35 (4) (2003) 399–458.
[6] W. Yang, C. Sun, L. Zhang, A multi-manifold discriminant analysis method for image feature extraction, Pattern Recognit. 44 (8) (2011) 1649–1657.
[7] Y. Xu, A. Zhong, J. Yang, D. Zhang, LPP solution schemes for use with face recognition, Pattern Recognit. 43 (12) (2010) 4165–4176.
[8] H. Han, S. Shan, X. Chen, W. Gao, A comparative study on illumination preprocessing in face recognition, Pattern Recognit. 46 (6) (2013) 1691–1699.
[9] S. Li, X. Liu, X. Chai, H. Zhang, S. Lao, S. Shan, Morphable displacement field based image matching for face recognition across pose, in: Proceedings of the European Conference on Computer Vision (ECCV), vol. 1 2012, pp. 102–115.
[10] H. Han, S. Shan, X. Chen, S. Lao, W. Gao, Separability oriented preprocessing for illumination-insensitive face recognition, in: Proceedings of the European Conference on Computer Vision (ECCV), vol. 7, 2012, pp. 307–320.
[11] Y. Xu, D. Zhang, J.-Y. Yang, A feature extraction method for use with bimodal biometrics, Pattern Recognit. 43 (3) (2010) 1106–1115.
[12] M.S. Bartlett, J.R. Movellan, T.J. Sejnowski, Face recognition by independent component analysis, IEEE Trans. Neural Netw. 13 (6) (2002) 1450–1464.
[13] X. Wang, X. Tang, A unified framework for subspace face recognition, IEEE Trans. Pattern Anal. Mach. Intell. 26 (9) (2004) 1222–1228.
[14] C. Liu, H. Wechsler, Gabor feature based classification using the enhanced fisher linear discriminant model for face recognition, IEEE Trans. Image Process. 11 (4) (2002) 467–476.
[15] T. Zhou, D. Tao, GoDec: randomized low rank & sparse matrix decomposition in noisy case, in: Proceedings of ICML, 2011, pp. 33–40.

[16] M. Fazel, E.J. Cand'es, B. Recht, P. Parrilo, Compressed sensing and robust recovery of low rank matrices, in: Proceedings of the 42nd Asilomar Conference on Signals, Systems and Computers, 2008.
[17] Y. Peng, A. Ganesh, J. Wright, RASL: robust alignment by sparse and low-rank decomposition for linearly correlated images, IEEE Trans. Pattern Anal. Mach. Intell. 34 (11) (2012) 2233–2246.
[18] J. Wright, Y. Ma, J. Mairal, et al., Sparse representation for computer vision and pattern recognition, in: Proceedings of IEEE, 2009, pp. 1–8.
[19] J. Wright, A.Y. Yang, A. Ganesh, et al., Robust face recognition via sparse representation, IEEE Trans. Pattern Anal. Mach. Intell. 31 (2) (2009) 210–227, http://dx.doi.org/10.1109/TPAMI.2008.79.
[20] L. Zhang et al., Sparse representation or collaborative representation: which helps face recognition? in: Proceedings of the ICCV, 2011.
[21] Y. Xu, D. Zhang, J. Yang, J.-Y. Yang, A two-phase test sample sparse representation method for use with face recognition, IEEE Trans. Circuits Syst. Video Technol. 21 (9) (2011) 1255–1262.
[22] J. Yin, Z. Liu, Z. Jin, W. Yang, Kernel sparse representation based classification, Neurocomputing 77 (1) (2012) 120–128.
[23] Q. Zhu, Y. Xu, J. Wang, Z. Fan, Kernel based sparse representation for face recognition, in: Proceedings of the International Conference on Pattern Recognition (ICPR), 2012.
[24] Y. Xu, Z. Fan, Q. Zhu, Feature space-based human face image representation and recognition, Opt. Eng. 51 (1) (2012) 017205.
[25] I. Naseem, R. Togneri, M. Bennamoun, Linear regression for face recognition, IEEE Trans. Pattern Anal. Mach. Intell. 32 (11) (2010) 2106–2112.
[26] I. Naseem, R. Togneri, M. Bennamoun, Robust regression for face recognition, Pattern Recognit. 45 (1) (2012) 104–118.
[27] W. Liu, Y. Wang, S.Z. Li, T. Tan, Nearest intra-class space classifier for face recognition, in: Proceedings of the International Conference on Pattern Recognition (ICPR), vol. 4, 2004, pp. 495–498.
[28] J. Yang, L. Zhang, Y. Xu, J.-Yu Yang, Beyond sparsity: the role of L1-optimizer in pattern classification, Pattern Recognit. 45 (2012) 1104–1118.
[29] M. Yang, L. Zhang, J. Yang, D. Zhang, Regularized robust coding for face recognition, IEEE Trans. Image Process 22 (5) (2013) 1753–1766.
[30] J. Yang, D. Chu, L. Zhang, Y. Xu, J. Yang, Sparse representation classifier steered discriminative projection with applications to face recognition, IEEE Trans. Neural Netw. Learn. Syst. S 24 (7) (2013) 1023–1035.
[31] 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.
[32] Y. Xu, Q. Zhu, Z. Fan, D. Zhang, J. Mi, Z. Lai, Using the idea of the sparse representation to perform coarse to fine face recognition, Inf. Sci. 238 (2013) 138–148.
[33] Z. Lai, Y. Xu, J. Yang, J. Tang, D. Zhang, Sparse tensor discriminant analysis, IEEE Trans. Image Process 22 (10) (2013) 3904–3915.
[34] Y. Xu, Q. Zhu, Z. Fan, Y. Wang, J.-S. Pan, From the idea of "sparse representation" to a representation-based transformation method for feature extraction, Neurocomputing 113 (2013) 168–176.
[35] Z. Lai, W.K. Wong, Z. Jin, J. Yang, Y. Xu, Sparse approximation to the eigensubspace for discrimination, IEEE Trans. Neural Netw. Learn. Syst. 23 (12) (2012) 1948–1960.
[36] X. He, S. Yan, Y. Hu, P. Niyogi, H. Zhang, Face recognition using Laplacian faces, IEEE Trans. Pattern Anal. Mach. Intell. 27 (3) (2005) 328–340.
[37] J. Yu, M. Wang, D. Tao, Semisupervised multiview distance metric learning for cartoon synthesis, IEEE Trans. Image Process. 21 (11) (2012) 4636–4648.
[38] X. Deng, X. Liu, M. Song, J. Cheng, J. Bu, C. Chen, LF-EME: local features with elastic manifold embedding for human action recognition, Neurocomputing 99 (2013) 144–153.
[39] J. Yu, D. Tao, M. Wang, Adaptive hypergraph learning and its application in image classification, IEEE Trans. Image Process. 21 (7) (2012) 3262–3272.
[40] J. Yu, D. Liu, D. Tao, H.S. Seah, Complex object correspondence construction in two-dimensional animation, IEEE Trans. Image Process. 20 (11) (2011) 3257–3269.
[41] J. Yu, D. Liu, D. Tao, H.S. Seah, On combining multiple features for cartoon character retrieval and clip synthesis, IEEE Trans. Syst. Man Cybern. Part B 42 (5) (2012) 1413–1427.
[42] L. Zhou, Y. Xu, Z.-M. Lu, T. Nie, Face recognition based on multi-wavelet and sparse representation, J. Inf. Hiding Multimed. Signal Process. 5 (3) (2014) 399–407. ⟨http://bit.kuas.edu.tw/~jihmsp/2014/vol5/JIH-MSP-2014-03-007.pdf⟩.
[43] N. Guan, D. Tao, Z. Luo, J. S.-Taylor, MahNMF: Manhattan Non-negative Matrix Factorization, CoRR /1207.3438 (2012).
[44] B. Jiang, K. Jia, Semi-supervised facial expression recognition algorithm on the condition of multi-pose, J. Inf. Hiding Multimed. Signal Process. 4 (3) (2013) 138–146.
[45] J.-B. Li, S.-C. Chu, J.-S. Pan, L.C. Jain, Multiple viewpoints based overview for face recognition, J. Inf. Hiding Multimed. Signal Process. 3 (4) (2012) 352–369.
[46] Y. Xu, Q. Zhu, Y. Chen, J.-S. Pan, An improvement to the nearest neighbor classifier and face recognition experiments, Int. J. Innov. Comput. Inf. Control 9 (2) (2013) 543–554.
[47] Y. Xu, Q. Zhu, Z. Fan, M. Qiu, Y. Chen, H. Liu, Coarse to fine K nearest neighbor classifier, Pattern Recognit. Lett. 34 (2013) 980–986.
[48] ⟨http://www.cl.cam.ac.uk/research/dtg/attarchive/facedatabase.html⟩.
[49] ⟨http://www.itl.nist.gov/iad/humanid/feret/feret master.html⟩.
[50] A.M. Martinez, R. Benavente, The AR face database CVC Technical Report, No. 24, June 1998.
[51] J. Yang, D. Zhang, A.F. Frangi, J.Y. Yang, Two dimensional PCA: a new approach to appearance-based face representation and recognition, IEEE Trans. Pattern Anal. Mach. Intell. 24 (1) (2004) 131–137.

Yong Xu received his B.S. and M.S. degrees at Air Force Institute of Meteorology (China) in 1994 and 1997, respectively. He then received his Ph.D. degree in pattern recognition and intelligence system at the Nanjing University of Science and Technology (NUST) in 2005.

From May 2005 to April 2007, he worked at Shenzhen graduate school, Harbin Institute of Technology (HIT) as a postdoctoral research fellow. Now he is a professor at Shenzhen graduate school, HIT. He also acts as a research assistant researcher at the HongKong Polytechnic University from August 2007 to June 2008. His current interests include pattern recognition, biometrics, and machine learning. He has published more than 40 scientific papers.

Xiaozhao Fang received the M.S. degree in computer science from Guangdong University of Technology, Guangzhou, China, in 2008. He is currently pursuing the Ph.D. degree in computer science and technology at Shenzhen Graduate School, Harbin Institute of Technology, Shenzhen, China.

He has published more than 9 journal papers. His current research interests include pattern recognition and machine learning.

Jane You obtained her B.Eng. in Electronic Engineering from Xi'an Jiao tong University in 1986 and Ph.D. in Computer Science from La Trobe University, Australia in 1992.

She was a lecturer at the University of South Australia and senior lecturer at Griffith University from 1993 till 2002. Currently she is a professor at the Hong Kong Polytechnic University. Her research interests include image processing, pattern recognition, medical imaging, biometrics computing, multimedia systems and data mining.

Yan Chen received her B.E. and M.E. degree in computer science from Northeastern University, China in 1997 and 2000 respectively, and her Ph.D. in 2010 from University of Technology, Sydney (UTS), Australian. Currently, she is a Post-Doctoral Researcher with Harbin Institute of Technology (HIT) at Shenzhen, China. Her research interests include computer vision and pattern recognition.

Hong Liu received Bachelor degree in computer science in 1990, Master degree in computer science in 1993 and Doctor degree in electrical control and automation engineering in 1996, post-doctoral in computer science and technology in 1996. Professor Liu is currently the supervisor of doctoral students, the director of Research Department of Shenzhen Graduate School and director of Intelligent Robot Laboratory of Peking University. He is also an IEEE member, an executive director and vice secretary – Intelligent Automation Committee of Chinese Automation Association (IACAA). His expertise is in the areas of image processing and pattern recognition, intelligent robots and computer vision, intelligent microsystems hardware and software co-design. He has published more than 100 papers in the important scholarly journals and international conferences, and access to subsidized Korean Academy of Jung-geun Awards, Department of Space Science and Technology Progress Award, and Peking University Teaching Excellence Award, Aetna award and candidates of Peking University Top Ten Teachers. He has done exchange visits in many famous universities and research institutions in several countries and regions, including the United States, Canada, France, the Netherlands, Japan, Korea, Singapore, Hong Kong and so on.