Pattern Recognition 46 (2013) 1151–1158

Contents lists available at SciVerse ScienceDirect

Pattern Recognition

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

Using the original and 'symmetrical face' training samples to perform representation based two-step face recognition

Yong Xu [a,b,*], Xingjie Zhu [a], Zhengming Li [a,c], Guanghai Liu [d], Yuwu Lu [a], 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] Industry Training Center, Guangdong Polytechnic Normal University, Guangzhou, China
[d] College of Computer Science and Information Technology, Guangxi Normal University, Guilin541004, China
[e] Engineering Lab on Intelligent Perception for Internet of Things, Shenzhen Graduate School, Peking University, Shenzhen, China

ARTICLE INFO

Article history:
Received 9 August 2012
Received in revised form
30 October 2012
Accepted 2 November 2012
Available online 12 November 2012

Keywords:
Pattern recognition
Face recognition
Sparse representation method

ABSTRACT

A limited number of available training samples have become one bottleneck of face recognition. In real-world applications, the face image might have various changes owing to varying illumination, facial expression and poses. However, non-sufficient training samples cannot comprehensively convey these possible changes, so it is hard to improve the accuracy of face recognition. In this paper, we propose to exploit the symmetry of the face to generate new samples and devise a representation based method to perform face recognition. The new training samples really reflect some possible appearance of the face. The devised representation based method simultaneously uses the original and new training samples to perform a two-step classification, which ultimately uses a small number of classes that are 'near' to the test sample to represent and classify it and has a similar advantage as the sparse representation method. This method also takes advantages of the score level fusion, which has proven to be very competent and usually performs better than the decision level and feature level fusion. The experimental results show that the proposed method outperforms state-of-the-art face recognition methods including the sparse representation classification (SRC), linear regression classification (LRC), collaborative representation (CR) and two-phase test sample sparse representation (TPTSSR).

1. Introduction

As one of the most active branches of biometrics, face recognition is attracting more and more attention [1–5]. In the past, various face recognition algorithms have been devised [6–9]. However, up to now, face recognition is still faced with a number of challenges such as varying illumination, facial expression and poses [10–16]. It seems that more training samples are able to reveal more possible variation of the illumination, facial expression and poses and are consequently beneficial for correct classification of the face. However, in real-world applications, there are usually only a limited number of available training samples. This is mainly because a face recognition system usually has limited storage space and captures training samples in a short time. In some special cases such as the personal identity card based face recognition, there is even only one training sample per subject. Non-sufficient training samples indeed have become one bottleneck of face recognition [17–20].

In order to obtain better face recognition result, the literatures have proposed to synthesize new samples from the true face images. For example, Tang et al. [21] used prototype faces and an optic flow and expression ratio image based method to generate 'virtual' facial expression. Thian et al. [22] used simple geometric transformations to generate virtual samples. Ryu et al. [23] exploited the distribution of the given training set to generate virtual training samples. Beymer et al. [24] and Vetter et al. [25] synthesized new face samples with virtual views. Jung et al. [26] exploited the noise to synthesize new face samples. Sharma et al. [27] synthesized multiple virtual views of a person under different poses and illumination from a single face image and exploited extended training samples to classify the face. In order to overcome the small sample size problem of face recognition, Liu et al. [28] represented each single image as a subspace spanned by its synthesized (shifted) samples. From the viewpoint of applications, the ways to generate virtual face images can be categorized into two kinds, i.e. the way to generate two-dimensional virtual face images and the way to construct three-dimensional virtual face images [29].

* Corresponding author at: Bio-computing Research Center, Shenzhen Graduate School, HIT Campus of Shenzhen University Town, Shenzhen 518005, China. Tel.: +86 755 26032458; fax: +86 755 26032461.
E-mail address: laterfall2@yahoo.com.cn (Y. Xu).

0031-3203/$ - see front matter © 2012 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.patcog.2012.11.003

The one training sample issue has also emerged as an active research sub-area of face recognition in recent years, and many ad hoc techniques have been proposed for this problem [30]. For example, previous literatures have used probabilistic matching [31–33] and neural network methods [34–36] for the one training sample issue. Qiao et al. [37] and Liu et al. [28] proposed sparsity preserving discriminant analysis and the single image subspace method for the one training sample issue respectively. Moreover, synthesizing virtual samples has also been used as a means to cope with the one training sample issue [23,24,38,39].

We note that the face has a symmetrical structure. Not only the facial structure but also the facial expression is symmetry [39]. The symmetry property has been successfully applied to face detection [40–42]. In face detection, the symmetry property of the human face is very useful to quickly locate the candidate faces [41].

In this paper, we propose to exploit the symmetry of the face to generate new training samples and devise a representation based method to perform face recognition. The new training samples indeed reflect some possible appearance of the face. The devised representation based method simultaneously uses the original and new training samples to perform a two-step classification. This method also takes advantages of the score level fusion, which has proven to be very competent and is usually better than the decision level and feature level fusion.

The two-step face recognition (TSFR) algorithm proposed in this paper is indeed a representation based classification (RBC) algorithm. A number of works have shown that the RBC algorithm can obtain a very high accuracy for image classification such as face recognition [43–49]. However, the conventional RBC algorithm such as the one proposed in [43,44] has a relatively high computational cost. Our TSFR algorithm is mathematically tractable and computationally efficient. The first step of our algorithm aims at identifying and discarding the classes whose training samples are 'far' from the test sample and the second step then exploits the training samples of the remaining classes to determine the class label of the test sample. The experimental results show that the proposed method can classify the face with a high accuracy and outperforms the state-of-the-art face recognition methods.

2. The proposed method

In this section we present the main steps of the proposed method in detail. Suppose that there are c classes and each class has n training samples. Let $x_1,\ldots,x_N$ represent all the N training samples ($N=nc$).

2.1. Main steps of the proposed method

The proposed method includes the following main steps. The first step generates 'symmetrical face' training samples. The second and third steps use the original and 'symmetrical face' training samples to perform two-step face recognition, respectively. The algorithm of two-step face recognition will be described in Section 2.2. The fourth step combines the scores obtained using the second and third steps to conduct weighted score level fusion, getting the ultimate classification result. We present these steps as follows:

Step 1. Use every original training sample to generate two 'symmetrical face' training samples. Let $x_i \in \Re^{p\times q}$ 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 'symmetrical face' training samples generated from x_i. The left half columns of y_i^1 is set to the same as that of x_i 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_i and the left half columns of y_i^2 is the mirror image of the right half columns of y_i^2. The mirror image S of an arbitrary image R is defined as $S(i,j)=R(i,V-j+1), i=1,\ldots,U, j=1,\ldots,V$. U and V stand for the numbers of the rows and columns of R, respectively. $S(i,j)$ denotes the pixel located in the ith row and jth column of S.

Step 2. Use the original training samples to perform two-step face recognition. Let s_j^1 denote the score of test sample z with respect to the j-th class. For the algorithm, please see Section 2.2.

Step 3. Use the 'symmetrical face' training samples to perform two-step face recognition. Let s_j^2 denote the score of test sample z with respect to the j-th class. This step shares the same algorithm as Step 2.

Step 4. Combine the scores obtained using the second and third steps to conduct weighted score level fusion. For test sample$z$, we use $s_j = w_1 s_j^1 + w_2 s_j^2$ to calculate the ultimate score with respect to the j-th class. w_1 and w_2 are the weights. Let $w_1+w_2=1$ and w_2 be smaller than w_1.

2.2. The algorithm of two-step face recognition

In this subsection we present the algorithm of two-step face recognition in detail. This algorithm first coarsely determines a small number of candidate classes of the test sample and then finely identifies the class that the test sample is the most similar to.

Before this algorithm is employed, all the samples should be converted into one-dimensional column vectors. We use $\tilde{x}_i, \tilde{y}_i^1, \tilde{y}_i^2$, $\tilde{z}$ to denote the one-dimensional column vectors of x_i, y_i^1, y_i^2, z respectively.

For simplicity of presentation, we describe only the algorithm on the original training samples. The algorithm on the 'symmetrical face' training samples is the same except that the original training samples are replaced with the 'symmetrical face' training samples. The algorithm first assumes that the following equation is approximately satisfied:

$$\tilde{z} = a_1\tilde{x}_1 + \ldots + a_N\tilde{x}_N, \tag{1}$$

a_i is referred to as the coefficient. Let $X=[\tilde{x}_1 \ldots \tilde{x}_N]$ and $A=[a_1 \ldots a_N]^T$. We rewrite Eq. (1) as

$$\tilde{z} = XA, \tag{2}$$

A is calculated using $\overline{A} = \left(X^T X + \mu I\right)^{-1} X^T \tilde{z}$. $\overline{A} = [\overline{a}_1,\ldots,\overline{a}_N]^T$. μ is a small positive constant and I is the identity matrix.

Eq.(1) implies that the effect on representing the test sample of the k-th class can be evaluated using

$$d_k = \left|\left|\tilde{z} - \sum\nolimits_{i=(k-1)n+1}^{kn} \overline{a}_i \tilde{x}_i\right|\right|. \tag{3}$$

It is clear that $\tilde{x}_{(k-1)n+1},\ldots,\tilde{x}_{kn}$ are training samples of the$k$-th class and $\overline{a}_{(k-1)n+1},\ldots,\overline{a}_{kn}$ are the corresponding coefficients. We would like to point out that the effect on representing the test sample of the k-th class is somewhat similar to the distance between the test sample and the k-th class. As a result, the effect on representing the test sample of the k-th class can be also evaluated by others means such as the sum of the Euclidean distances between the test sample and each training sample from the k-th class.

If $d_{r_1} \le d_{r_2} \ldots \le d_{r_c}$, then we say that the r_1-th, r_2-th,..,r_t-th classes are the first t candidate classes of the test sample. In other words, we can consider that the ultimate class label of the test sample should be one element of $D = c_{r_1}, c_{r_2}, \ldots, c_{r_t}$.

$c_{r_1}, c_{r_2}, \ldots, c_{r_t}$ are the class labels of the first t candidate classes, respectively. As the above steps roughly determine that the test sample is from a small number of classes, we refer to them as coarse classification.

The algorithm then uses a linear combination of the training samples of the first t candidate classes to represent the test sample. In other words, if the training samples of the first t candidate classes are denoted by $\tilde{x}'_1, \ldots, \tilde{x}'_{tn}$, respectively, then the algorithm assumes that the following equation is approximately satisfied:

$$\tilde{z} = f_1 \tilde{x}'_1 + \ldots + f_{tn} \tilde{x}'_{in}, \tag{4}$$

where f_i is the coefficient. We rewrite Eq.(4) as

$$\tilde{z} = X'F, \tag{5}$$

where $F = [f_1 \ldots f_{tn}]^T$, $X' = [\tilde{x}'_1, \ldots, \tilde{x}'_{tn}]$. F is calculated using

$$\bar{F} = \left(X'^T X' + \gamma I\right)^{-1} X'^T \tilde{z} \tag{6}$$

$\bar{F} = [\bar{f}_1 \ldots \bar{f}_{tn}]^T$. γ is a small positive constant and I also denotes the identity matrix.

Suppose that $\tilde{x}'_g, \ldots, \tilde{x}'_h$ stand for all the training samples of the r-th class $\left(r \in c_{r_1}, c_{r_2}, \ldots, c_{r_t}\right)$ and the coefficients are $\bar{f}_g, \ldots, \bar{f}_h$, respectively. The ultimate effect on representing the test sample of the r-th class can be evaluated using

$$u_r = \left|\left|\tilde{z} - \sum_{i=g}^{h} \bar{f}_i \tilde{x}'_i\right|\right|. \tag{7}$$

If $k = \operatorname{argmin} u_r$, then test sample $\tilde{z}$ is ultimately assigned to the k-th class, which is also referred to as the result of fine classification. The code of the proposed method can be downloaded at (http://www.yongxu.org/lunwen.html).

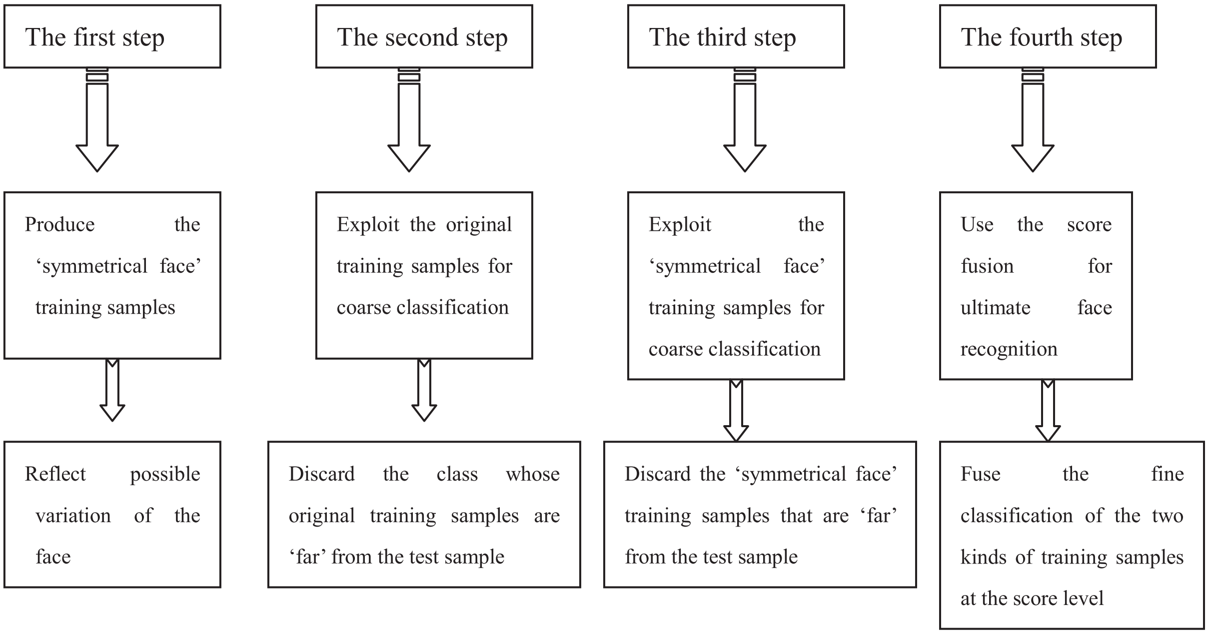
3. Analysis of the proposed method

In this section we show the rationales of the proposed method. First, the 'symmetrical face' training samples in the proposed method indeed reflect some possible appearance of the face, which are not shown by the original training samples. Fig. 1 shows some original training samples from the ORL face database and the 'symmetrical face' training samples generated from the original training samples. Fig. 2 shows some original training samples from the FERET face database and the corresponding 'symmetrical face' training samples. Fig. 3 shows some original training samples from the AR face database and the corresponding 'symmetrical face' training samples. We see that the 'symmetrical face' training samples not only seem to be different from the original training samples, but also indeed somewhat reflect the possible variation of the face in image scale, pose and illumination. Thus, 'symmetrical face' training samples are very useful to overcome the issue of non-sufficient training samples. In the real world, the image scale is variable owing to the various distances between the person and camera. Fig. 4 shows a test sample (4(a)) that is erroneously and correctly classified by the collaborative representation (CR) method proposed in [47] and our method, respectively. As the collaborative representation method exploits only the original training samples and our method uses both the original training samples and 'symmetrical face' training samples, the 'symmetrical face' training sample is really beneficial for correct classification of the test sample. During the register phase in real-world applications, after the face image is obtained by the face detection procedure, the 'symmetrical face' training samples can be easily and efficiently generated. Since the 'symmetrical face' training samples are complementary for the original training samples, the system can capture only a few original training samples and can still obtain enough information of the face. To capture only a few original training samples will also allow the system to take only a short time to complete the register phase.

The second rationale of the proposed method is that it respectively uses the 'symmetrical face' training samples and the original training samples to obtain the scores of the test sample with respect to different classes and properly exploits a weighted fusion scheme to combine them for ultimate face recognition. As the 'symmetrical face' training samples contain less information than the original training samples, it is very reasonable for the proposed method to assign a smaller weight to the 'symmetrical face' training samples.

The third rationale of the proposed method is that it uses the two-step face recognition, which is able to reduce the side-effect on classification, of the test sample, of the classes that are very dissimilar to the test sample. Actually, the literature has shown that the test sample is usually not from these classes [45]. As a result, by eliminating these classes the fine recognition can increase the probability of the test sample being correctly classified and can achieve a higher accuracy. The significance of every step of the proposed method can be briefly described below.

Fig. 1. Some original training samples from the ORL face database and the corresponding 'symmetrical face' training samples. The first row shows the original training samples. The second and third rows respectively show the first and second 'symmetrical face' training samples generated from the original training sample.



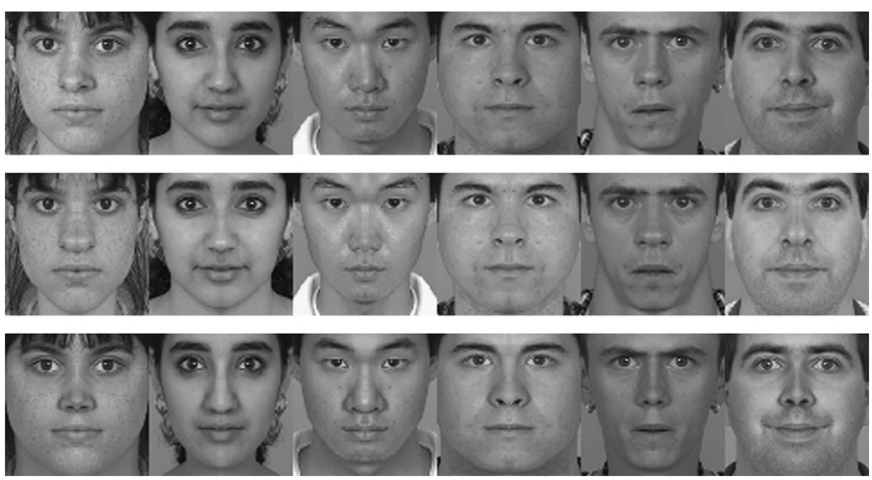


Fig. 2. Some original training samples from the FERET face database and the corresponding 'symmetrical face' training samples. The first row shows the original training samples. The second and third rows respectively show the first and second 'symmetrical face' training samples generated from the original training sample.

In order to assess the capability, to represent the test sample, of our method, we define representation error as follows. Let z be the test sample. For a RBC method, let P and s be the matrix consisting of all the available training samples and the solution vector of the corresponding RBC, respectively. $d_z = ||z - Ps||$ is referred to as the representation error of z. Ps is referred to as representation result of$z$. It is clear that the smaller the d_z is the more the representation result approximates the test sample. The analysis also shows that in our method, a lower representation error usually does not mean a lower rate of classification errors. This is true for both the RBC on the original training samples and the RBC on the 'symmetrical face' training samples. Figs. 5 and 6 show the representation errors of the test sample obtained using the coarse classification and fine classification based on the original training samples and 'symmetrical face' training samples, respectively. These figures are obtained under the condition that the first 14 face images of each subject in the AR face database are used as original training samples and the remaining face images are taken as test samples. We see that the coarse classification based on the original training samples and 'symmetrical face' training samples always leads to a lower representation error than the corresponding fine classification. However, these two figures show that the fine classification produces fewer classification errors than the coarse classification. Actually, as shown in the experimental section, the fine classification always obtains a lower rate of classification errors. This tells us that for a RBC, the key is not to produce a low representation error but to make the class that the test sample is truly from the most similar with the test sample, having the minimum deviation from the test sample among all the classes. Though the fine classification is not optimal for providing a good representation for the test sample, it performs very well in making the test sample more similar with the test sample. As shown early, it achieves this by discarding the classes that are 'far' from the test sample. Usually, there is also a large probability of the test sample being not truly from these classes.

4. Experimental results

We used the FERET, AR and ORL face databases to conduct experiments. For the FERET and ORL databases, we show the experimental results under the conditions that the weights of the

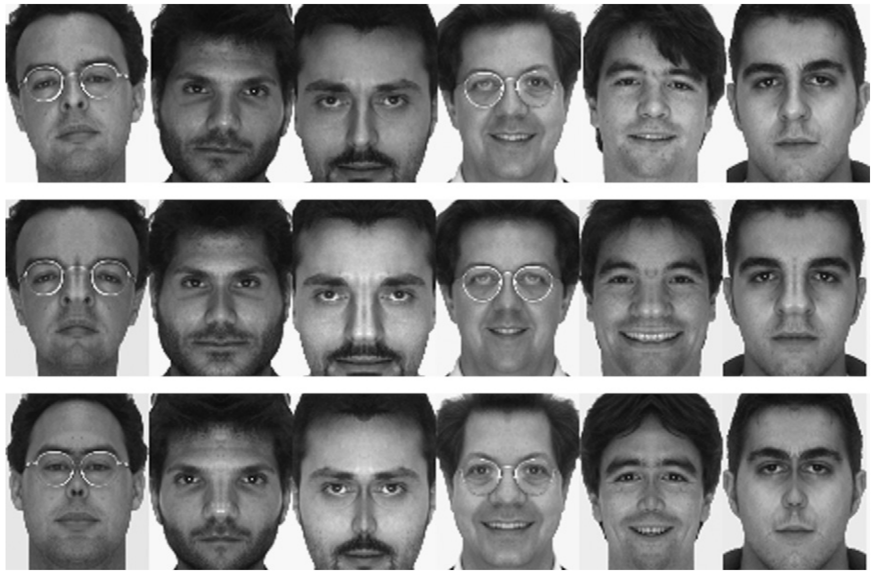


Fig. 3. Some original training samples from the AR face database and the corresponding 'symmetrical face' training samples. The first row shows the original training samples. The second and third rows respectively show the first and second 'symmetrical face' training samples generated from the original training sample.

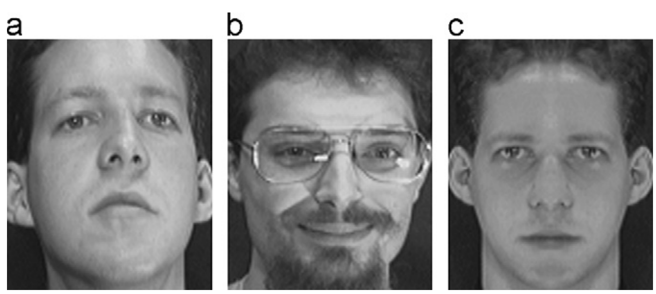


Fig. 4. A test sample that is erroneously and correctly classified by the collaborative representation and our method, respectively. (a) Test sample. (b) Original training sample which is the nearest to the test sample. (c) 'Symmetrical face' training sample which is the nearest to the test sample among all the 'symmetrical face' training samples.

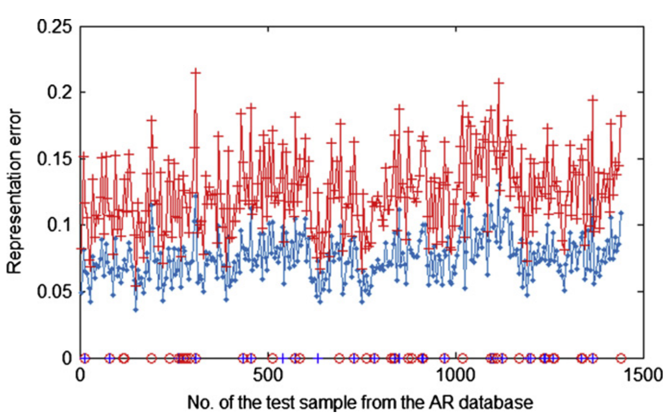


Fig. 5. Representation errors of the test sample obtained using the coarse classification and fine classification based on the original training samples. The blue and red lines respectively depict the representation errors obtained using the coarse classification and fine classification. In the horizontal axis, the red circle and blue '+' respectively means that the corresponding test sample is erroneously classified by the coarse classification and fine classification based on the original training samples. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

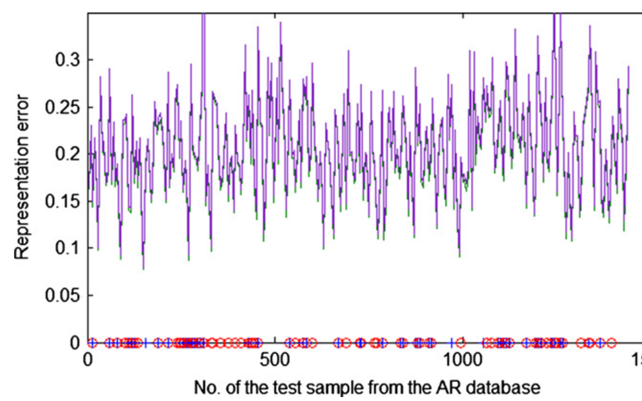


Fig. 6. Representation errors of the test sample obtained using the coarse classification and fine classification based on the 'symmetrical face' training samples. The green and purple lines respectively depict the representation errors obtained using the coarse classification and fine classification. In the horizontal axis, the red circle and blue '+' respectively means that the corresponding test sample is erroneously classified by the coarse classification and fine classification based on the 'symmetrical face' training samples. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 1
Experimental results on the AR database.

Number of the original training samples per class	13	14	15	16
The proposed method (w_1=0.75) (%)	23.72	9.93	8.03	6.83
CR on the original training samples (%)	28.97	14.86	11.82	9.83
CR on the 'symmetrical face' training samples (%)	36.15	27.15	22.27	18.83
Coarse score fusion (%)	28.33	14.44	11.74	9.58
SRC (%)	32.37	16.74	16.82	17.42
TPTSSR (%)	24.81	9.72	8.33	7.17
LRC (%)	34.36	13.13	12.58	13.25
Feature space-based representation method (σ=1.0e6) [49] (%)	39.36	25.07	24.47	25.42
Combination of two-step classification and the feature space-based representation method (σ=1.0e6) (%)	39.23	25.07	24.47	25.33

score fusion were set to $w_1=0.85$ and $w_2=0.15$,$w_1=0.75$ and $w_2=0.25$, $w_1=0.65$ and $w_2=0.35$, respectively. For the AR database, because the experiment is very time consuming, we conduct only the experiment on $w_1=0.75$ and $w_2=0.25$. Both μ and γ were set to 0.01. Each sample was converted into a unit vector with length of 1 in advance. Tables 1, 2 and 3 show the rate of classification errors under the condition that 0.3^*c candidate classes were used. Besides our proposed method was tested, several state-of-the-art face recognition methods such as two-phase test sample sparse representation (TPTSSR) [45], collaborative representation (CR) proposed in [47], the feature space-based representation method proposed in [49], sparse representation classification (SRC) proposed in [50] and linear regression classification (LRC) [51]were also tested. As the performance of the feature space-based representation method is directly related to parameter σ, we set it to different values and show the best classification result and the corresponding value in the table.

Table 2
Experimental results on the ORL database.

Number of the original training samples per class	1	2	3
The proposed method ($w_1=0.85$) (%)	24.72	10.31	10.00
The proposed method ($w_1=0.75$) (%)	24.17	10.31	8.93
The proposed method ($w_1=0.65$) (%)	24.72	10.63	9.29
CR on the original training samples (%)	31.39	16.56	15.00
CR on the ‘symmetrical face’ training samples (%)	34.72	23.13	20.71
Coarse score fusion (%)	30.28	16.25	13.57
SRC (%)	26.67	15.00	14.29
TPTSSR (%)	26.39	13.44	11.43
LRC (%)	32.50	20.94	18.21
Feature space-based representation method($\sigma=1.0e7$) [49] (%)	27.22	11.87	10.71
Combination of two-step classification and the feature space-based representation method ($\sigma=1.0e7$) (%)	27.22	11.25	10.71

Table 3
Experimental results on the FERET database.

Number of the original training samples per class	1	2	3
The proposed method ($w_1=0.85$) (%)	50.42	35.30	41.88
The proposed method ($w_1=0.75$) (%)	48.75	34.80	41.63
The proposed method ($w_1=0.65$)	48.08	32.80	38.25
CR on the original training samples (%)	55.67	41.60	55.63
CR on the ‘symmetrical face’ training samples (%)	58.33	39.20	41.38
Coarse score fusion (%)	55.67	40.00	52.25
SRC (%)	50.25	35.20	40.00
TPTSSR (%)	52.17	38.70	46.88
LRC (%)	55.08	36.80	42.88
Feature space-based representation method($\sigma=1.0e6$) [49] (%)	56.25	43.40	50.50
Combination of two-step classification and the feature space-based representation method ($\sigma=1.0e6$) (%)	54.25	41.70	48.50

In the experiments, we performed ‘coarse score fusion’ as follows: CR was respectively first implemented for the original and ‘symmetrical face’ training samples and then the scores generated from the original and ‘symmetrical face’ training samples were fused for ultimate face recognition by using the same weight fusion scheme as our proposed method. When implementing TPTSSR, we set parameter $M=N/2$, where N is the number of all the original training samples. We also tested the combination of two-step classification and the feature space-based representation method proposed in [49]. Specifically, the feature space-based representation method first respectively uses the original and ‘symmetrical face’ training samples to perform coarse classification and then respectively uses them to conduct fine classification. The fine classification was also implemented under the condition that 0.3^*c candidate classes were exploited. Finally, the weighted fusion scheme ($w_1=0.75$ and $w_2=0.25$) in our method was used to combine the scores generated from the fine classification on the original and ‘symmetrical face’ training samples and the combined score was used to ultimately classify the test sample.

4.1. Experiments the AR face database

From the AR face database [52], we used 3120 Gray images from 120 subjects, each providing 26 images. These images were taken in two sessions. Every image was resized to a 50×40 image. We respectively took the first 13, 14, 15 and 16 face images of each subject as the original training samples and treated the remaining face images as the test samples. The experimental results were shown in Table 1. We see that our method obtains a lower rate of classification errors than all the other methods. For example, when the first 13 face images of each subject were used as the original training samples and the remaining face images were taken as the test samples, the rates of classification errors of our method, CR on the original training samples, SRC, TPTSSR, LRC and the feature space-based representation method are 23.72%, 28.97%, 32.37%, 24.81%, 34.36% and 39.36%, respectively. The fact that coarse score fusion also obtains a higher rate of classification errors than our method also means that to discard the classes that are ‘far’ from the test sample is really beneficial for the correct classification of the test sample. Moreover, we see that the combination of two-step classification and the feature space-based representation method also performs worse than our method.

4.2. Experiments the ORL face database

The ORL database [53] 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 an image with one half of the original size by using the down-

Table 4
Variation of the rate of classification errors of our method ($w_1=0.75$ and $w_2=0.25$) with the number of candidate classes.

The ORL database (3 training sample per subject)				
Number of candidate classes	10	20	30	40
The rate of classification errors (%)	10.00	9.29	8.93	12.50
The FERET database (3 training sample per subject)				
Number of candidate classes	*50*	100	150	200
The rate of classification errors (%)	40.75	45.25	47.13	51.25
The AR database (13 training sample per subject)				
Number of candidate classes	30	60	90	120
The rate of classification errors (%)	23.40	24.42	25.64	28.21

sampling algorithm. We respectively took the first 1, 2 and 3 face images of each subject as the original training samples and treated the remaining face images as the test samples. The experimental results were shown in Table 2. It shows again that our method performs better than all the other methods. For example, when the first face image of each subject and the remaining face images were respectively used as the original training samples and the test samples, the rates of classification errors of our method ($w_1 = 0.75$ and $w_2 = 0.25$), CR on the original training samples, SRC, TPTSSR, LRC and the feature space-based representation method are 24.17%, 31.39%, 26.67%, 26.39%, 32.50% and 27.22% respectively. We also see that the combination of two-step classification and the feature space-based representation method obtains a higher rate of classification errors than our method.

4.3. Experiments the FERET face database

We used a subset of the FERET face database [54] to test our method. This subset consists of 1400 images from 200 individuals each providing seven images. 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 1, 2 and 3 face images of each subject as the original training samples and treated the remaining face images as the test samples. Table 3 shows the rate of classification errors. This table shows that in most cases our method outperforms all the other methods. Table 3 also shows that the combination of two-step classification and the feature space-based representation method obtains a higher rate of classification errors than our method.

4.4. Variation of the performance of our method with the number of candidate classes

In order to comprehensively show the performance of our method, we use Table 4 to briefly indicate the variation of the rate of classification errors of our method with the number of candidate classes. We see that when the number of candidate classes is smaller than that of all the classes, our method almost always obtains a lower rate of classification errors. This clearly shows that to discard the classes that are 'far' from the test sample is really beneficial for the improvement of the accuracy of our method.

5. Conclusions

We propose a very promising method to exploit limited training samples for two-step face recognition. The new training samples generated in this paper can well exploit the symmetry structure of the face. The face almost always has a symmetrical structure and the real face image is usually not an axial symmetry image owing to the non-frontal poses, so to use the left (or right) half of the face image to produce the right (or left) half of the face allows us to obtain a possible new sample of the right (or left) half of the face. This is very helpful for overcoming the drawback of limited training samples in the real-world face recognition system. The two-step classification used in the paper can greatly eliminate the side-effect, on the classification of the test sample, of the classes that are 'far' from it (the test sample is usually not truly from these classes). The proposed weighted score level fusion can use a very proper way to integrate the original training samples and 'symmetrical face' training samples for ultimate face recognition by assigning a larger weight to the original training samples. The experimental results show that the proposed method can outperform a number of state-of-the-art representation based methods.

Acknowledgments

This article is partly supported by Program for New Century Excellent Talents in University (Nos. NCET-08–0156 and NCET-08–0155), NSFC under Grants nos. 61071179, 61173086, 61020106004, 61001037 and 61173086, as well as the Fundamental Research Funds for the Central Universities (HIT.NSRIF. 2009130).

References

[1] S.Z. Li, A. Jain, Encyclopedia of biometrics, Springer, US, 2009.
[2] J. Lu, X. Yuan, T. Yahagi, A method of face recognition based on fuzzy c-means clustering and associated sub-NNs, IEEE Transactions on Neural Networks 18 (1) (2007) 150–160.
[3] L. Zhang, S. Chen, L. Qiao, Graph optimization for dimensionality reduction with sparsity constraints, Pattern Recognition 45 (3) (2012) 1205–1210.
[4] Y.-H. Liu, Y.-T. Chen, Face recognition using total margin-based adaptive fuzzy support vector machines, IEEE Transactions on Neural Networks 18 (1) (2007) 178–192.
[5] Z. Fan, Y. Xu, D. Zhang, Local linear discriminant analysis framework using sample neighbors, IEEE Transactions on Neural Networks 22 (7) (2011) 1119–1132.
[6] Ting Shan, Brian C. Lovell, Shaokang Chen: Face recognition robust to head pose from one sample image, ICPR 1 (2006) 515–518.
[7] Dalong Jiang, Yuxiao Hu, Shuicheng Yan, Lei Zhang, HongJiang Zhang, Wen Gao: Efficient 3D reconstruction for face recognition, Pattern Recognition 38 (6) (2005) 787–798.
[8] Terence Sim, Takeo Kanade, Combining Models and Exemplars for Face Recognition: An Illuminating Examples, in: Proceedings of the CVPR 2001 Workshop on Models versus Exemplars in Computer Vision, December, 2001.
[9] Q. Feng, J.-S. Pan, L. Yan, Restricted nearest feature line with ellipse for face recognition, Journal of Information Hiding and Multimedia Signal Processing 3 (3) (2012) 297–305.
[10] S.-J. Wang, J. Yang, M.-F. Sun, et al., Sparse tensor discriminant color space for face verification, IEEE Transactions on Neural Networks and Learning Systems 23 (6) (2012) 876–888.
[11] D. Zhang, F. Song, Y. Xu, Z. Lang, Advanced pattern recognition technologies with applications to biometrics, New York, Medical Information Science Reference, 2009.
[12] J. Wang, J. You, Q. Li, Y. Xu, Orthogonal discriminant vector for face recognition across pose, Pattern Recognition 45 (12) (2012) 4069–4079.
[13] X. Zhang, Y. Gao, Face recognition across pose: A review, Pattern Recognition 42 (11) (2009) 2876–2896.
[14] S.N. Kautkar, G.A. Atkinson, M.L. Smith, Face recognition in 2D and 2.5D using ridgelets and photometric stereo, Pattern Recognition 45 (9) (2012) 3317–3327.
[15] L. Pishchulin, T. Gass, P. Dreuw, H. Ney, Image warping for face recognition: From local optimality towards global optimization, Pattern Recognition 45 (9) (2012) 3131–3140.
[16] B. Jun, D. Kim, Robust face detection using local gradient patterns and evidence accumulation, Pattern Recognition 45 (9) (2012) 3304–3316.
[17] I. Naseem, R. Togneri, M. Bennamoun, Robust regression for face recognition, Pattern Recognition 45 (1) (2012) 104–118.
[18] W. Yang, C. Sun, L. Zhang, A multi-manifold discriminant analysis method for image feature extraction, Pattern Recognition 44 (8) (2011) 1649–1657.
[19] J. Yang, D. Zhang, Y. Xu, J.-Y. Yang, Two-dimensional discriminant transform for face recognition, Pattern Recognition 38 (7) (2005) 1125–1129.
[20] J. Yang, D. Zhang, A.F. Frangi, J.-Y. Yang, Two-dimensional PCA: A new approach to appearance-based face representation and recognition, IEEE Transactions on Pattern Analysis and Machine Intelligence 26 (1) (2004) 131–137.
[21] B. Tang, S. Luo, H. Huang, High performance face recognition system by creating virtual sample, in: Proceedings of the International Conference on Neural Networks and Signal Processing, 2003, pp. 972–975.
[22] N.P.H. Thian, S. Marcel, S. Bengio. Improving face authentication using virtual samples, in: Proceeding of the IEEE International Conference on Acoustics, Speech, and Signal Processing, 2003, pp. 6–10.
[23] Y.-S. Ryu, S.-Y. Oh, Simple hybrid classifier for face recognition with adaptively generated virtual data, Pattern Recognition Letters 23 (7) (2002) 833–841.
[24] D. Beymer, T. Poggio. Face recognition from one example view, in: Proceedings of the Fifth International Conference on Computer Vision, 1995, pp. 500–507.
[25] T. Vetter, Synthesis of novel views from a single face image, International Journal of Computer Vision 28 (2) (1998) 102–116.

[26] H.-C. Jung. Authenticating corrupted face image based on noise model, in: Proceedings of the Sixth IEEE International Conference on Automatic Face and Gesture Recognition, 2004, pp. 272–277.
[27] A. Sharma, A. Dubey, P. Tripathi, V. Kumar, Pose invariant virtual classifiers from single training image using novel hybrid-eigenfaces, Neurocomputing 73 (10–12) (2010) 1868–1880.
[28] J. Liu, S. Chen, Z.-H. Zhou, X. Tan. Single image subspace for face recognition, in: Proceedings of the 3rd International Conference on Analysis and Modeling of Faces and Gestures, 2007, pp. 205–219.
[29] Y. Hu, D. Jiang, S. Yan, L. Zhang, H. Zhang, Automatic 3D Reconstruction for Face Recognition, FGR (2004) 843–850.
[30] X. Tan, S. Chen, Z.-H. Zhou, F. Zhang, Face recognition from a single image per person: A survey, Pattern Recognition 39 (9) (2006) 1725–1745.
[31] A.M. Martinez, Recognizing imprecisely localized, partially occluded, and expression variant faces from a single sample per class, IEEE Transactions on Pattern Analysis and Machine Intelligence 25 (6) (2002) 748–763.
[32] A.M. Martinez. Recognizing expression variant faces from a single sample image per class, in: Proceedings of the IEEE Computer Vision and Pattern Recognition, 2003, pp. 353–358.
[33] A.M. Martinez, Matching expression variant faces, Vision Research 43 (9) (2003) 1047–1060.
[34] S. Lawrence, C.L. Giles, A. Tsoi, A.D. Back, Face recognition: A convolutional neural-network approach, IEEE Transaction on Neural Networks 8 (1) (1997) 98–113.
[35] X. Tan, S. Chen, Z.-H. Zhou, F. Zhang, Recognizing partially occluded, expression variant faces from single training image perperson with SOM and soft kNN ensemble, IEEE Transaction on Neural Networks 16 (4) (2005) 875–886.
[36] X. Tan, J. Liu, S. Chen, Recognition from a single sample per person with multiple SOM fusion, Lecture Notes in Computer Science 3972 (2006) 128–133.
[37] L. Qiao, S. Chen, X. Tan, Sparsity preserving discriminant analysis for single training image face recognition, Pattern Recognition Letters 31 (5) (2010) 422–429.
[38] P. Niyogi, F. Girosi, T. Poggio, Incorporating prior information in machine learning by creating virtual examples, in: Proceedings of the IEEE, 86, 11, 1998, 2196–2209.
[39] P. Ekman, J.C. Hager, W.V. Friesen, The symmetry of emotional and deliberate facial actions, Psychophysiology 18 (2) (1981) 101–106.
[40] S. Saha, S. Bandyopadhyay. A symmetry based face detection technique, in: Proceedings of the IEEE WIE National Symposium on Emerging Technologies, 2007, pp. 1–4.
[41] M.-C. Su, C.-H. Chou. Application of associative memory in human face detection, in: Proceedings of the International Joint Conference on Neural Networks, vol. 5, 1999, pp. 3194–3197.
[42] E. Saber, A. Murat Tekalp, Frontal-view face detection and facial feature extraction using color, shape and symmetry based cost functions, Pattern Recognition Letters 19 (8) (1998) 669–680.
[43] J. Wright, Y. Ma, J. Mairal, et al., Sparse representation for computer vision and pattern recognition, in: Proceeding of the IEEE, 98, 6, 2010, 1031–1044.
[44] J. Wright, A.Y. Yang, A. Ganesh, et al., Robust face recognition via sparse representation, IEEE Transactions on Pattern Analysis and Machine Intelligence 31 (2) (2009) 210–227.
[45] Y. Xu, D. Zhang, J. Yang, J.-Y. Yang, A two-phase test sample sparse representation method for use with face recognition, IEEE Transactions on Circuits and Systems for Video Technology 21 (9) (2011) 1255–1262.
[46] X.-T. Yuan, S. Yan. Visual classification with multi-task joint sparse representation, in: Proceedings of the 2010 IEEE Conference on Computer Vision and Pattern Recognition, 2010, pp. 3493–3500.
[47] L. Zhang, M. Yang, X. Feng. Sparse representation or collaborative representation: Which helps face recognition? in: Proceedings of the 2011 IEEE International Conference on Computer Vision, 2011, pp. 471–478.
[48] Y. Xu, W. Zuo, Z. Fan, Supervised sparse presentation method with a heuristic strategy and face recognition experiments, Neurocomputing 79 (2011) 125–131.
[49] Y. Xu, Z. Fan, Q. Zhu, Feature space-based human face image representation and recognition, Optical Engineering 51 (1) (2012) 017205-017205-7.
[50] K. Koh, S. Kim, S. Boyd. l1_ls: A matlab solver for large-scale L1-regularized least squares problems, 〈http://www.stanford.edu/~boyd/l1_ls/〉.
[51] I. Naseem, R Togneri, M. Bennamoun, Linear regression for face recognition, IEEE Transactions on Pattern Analysis and Machine Intelligence 32 (11) (2010) 2106–2112.
[52] 〈http://www2.ece.ohio-state.edu/~aleix/ARdatabase.html〉.
[53] 〈http://www.cl.cam.ac.uk/research/dtg/attarchive/facedatabase.html〉.
[54] 〈http://www.itl.nist.gov/iad/humanid/feret/feret_master.html〉.

Yong Xu was born in Sichuan, China, in 1972. He received his B.S. degree, M.S. degree in 1994 and 1997, respectively. He received the Ph.D. degree in Pattern Recognition and Intelligence system at NUST (China) in 2005. Now he works at Shenzhen Graduate School, Harbin Institute of Technology. His current interests include pattern recognition, biometric, face recognition, machine learning and image processing.

Xingjie Zhu received his B.S. degree in Anyang Institute of Technology in 2010. Now, He is studying in Shenzhen Graduate School, Harbin Institute of Technology. His current interests include Pattern recognition, Video analysis and Machine learning.

Zhengming Li was born in Henan, China, in 1982. He received his B.S. degree, M.S. degree in 2004 and 2007, respectively. He is currently 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 and machine learning.

Guang-Hai Liu is currently an Associate Professor in the College of Computer Science and Information Technology, Guangxi Normal University in China. He received Ph.D. degree from the School of Computer Science and Technology, Nanjing University of Science and Technology (NUST). His current research interests are in the areas of image processing, pattern recognition and artificial intelligence.

Yuwu Lu was born in Hebei, China, in 1985. He received his B.S. degree, M.S. degree in 2008 and 2011, respectively. 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 2 scientific papers. His current research interests include pattern recognition and machine learning.

Hong Liu received his Ph.D. in Mechanical Electronics and Automation in 1996, and serves as a full professor in School of EE&CS, Peking University. He is also the director of Engineering Lab on Intelligent Perception for Internet of Things. His research fields include computer vision and robotics, image processing and pattern recognition. Prof. Liu has published more than 100 papers and gained Chinese National Aero-space Award, Excellence Teaching Award, and Candidates of Top Ten Outstanding Professors, Peking University. He is an IEEE member, vice chair of Intelligent Robotics Society of Chinese Association for Artificial Intelligent (CAAI), and also the President of National Youth Committee of CAAI. He has served as co-chairs, session chairs or PC members of many important international conferences, such as IEEE/RSJ IROS, IEEE ROBIO and IEEE SMC, and also serves as reviewers for many international journals such as Pattern Recognition, IEEE Trans. on Signal Processing and IEEE Trans. on PAMI.