

A hierarchical face recognition algorithm based on humanoid nonlinear least-squares computation

Zhendong Wu^{1,2} · Jie Yuan¹ · Jianwu Zhang¹ · Huaxin Huang²

Received: 7 June 2015 / Accepted: 2 September 2015 / Published online: 10 September 2015
© Springer-Verlag Berlin Heidelberg 2015

Abstract Face recognition is a critical component in many computer vision applications. Although now big data computing could bring high face recognition rate, it needs strong computing power, and normally working in the cloud. However, in many computer vision applications, especially a lot of front-end application, it needs to quickly and efficiently recognize faces. Inspired by human rapid and accurate identification of familiar faces, we think that there may be a class of fast computing mechanisms that play a role in human face recognition and thus improve the accuracy of recognition. In this paper, we study the nonlinear least-squares calculation in face recognition application, and find that it really can improve the recognition rate, and more importantly, it can deal with any combination of face features, such as “detail” and “holistic” features, obtaining a high recognition rate. Further more, we study Sparse Representation-based Classification in depth and find that some “detail” features, such as mouth, eyes, could be accurately identified by Sparse Representation. Then we propose a hierarchical face recognition

algorithm by the use of nonlinear least-squares computation named HSRC. HSRC combines the components of face features using nonlinear least-squares and reduces the requirement of alignment and integrity and so on. And the results of these experiments prove that the face recognition rate can be considerably improved.

Keywords Face recognition · Facial components · Nonlinear least-squares · Humanoid computing

1 Introduction

Face recognition is a critical component in many computer vision applications, such as access control, video surveillance and public security. In real world, face images are often corrupted by many unknown factors, namely, illumination, occlusion and expression, so that face recognition is still a challenging task. What’s more, in some computer vision front-end application, it is necessary to quickly and efficiently recognize faces, which is still an open problem.

Among the proposed face recognition method, deep learning (DL), manifold learning and sparse representation-based method are three main research branches. Hinton proposed the deep learning method, which converts high-dimensional data to low-dimensional codes by training a multilayer neural network (Hinton and Salakhutdinov 2006). Deep learning method has made great success in computer vision field (Farabet et al. 2013; Kavukcuoglu et al. 2010; Ciresan et al. 2012). In the field of face recognition, the deep learning method has also achieved very good results. Taigman et al. (2014) proposed a deep learning (DL) frame-work to deal with the mission of unconstrained face recognition. Researches have presented

✉ Zhendong Wu
wzd@hdu.edu.cn

Jie Yuan
0-year-old@163.com

Jianwu Zhang
jwzhang@hdu.edu.cn

Huaxin Huang
rw211@zju.edu.cn

¹ School of Communication Engineering, Hangzhou Dianzi University, Hangzhou 310018, China

² Center for the Study of Language and Cognition, Zhejiang University, No. 148 Tian Mu Shan Road, Hangzhou 310028, China

a system (DeepFace) that is close to the human visual system. Despite the high recognition rate, the deep learning method needs to handle massive amounts of data and multilayer neural network, and at present, it is very difficult to study the relationship between human face recognition and multilayer neural network recognition.

Tenenbaum et al. (2000) introduced the manifold learning, which is built on classical multidimensional scaling (MDS) but aim to preserve the intrinsic geometry of the data. Silva and Tenenbaum (2003) proposed the mathematical description of manifold learning in detail. Based on that, manifold learning obtains the development continuously. Ataer-Cansizoglu et al. (2014) proposed a distance order preserving manifold learning algorithm that extends the basic mean-squared error cost function used mainly in multidimensional scaling (MDS)-based methods. Bhatia et al. (2014) presented a novel method of hierarchical manifold learning, which aims to automatically discover regional properties of medical image datasets. It extends conventional techniques by additionally examining local variations, in order to produce spatially-varying manifold embeddings to as a characterisation of a given dataset. Lu et al. (2013) proposed a multi manifold analysis technology for face recognition. Researches assume that the samples from the same face class would define a single manifold in the feature space and seek better separated feature dimensions in the low-dimensional feature spaces. Manifold learning has the noise sensitive problem, which affects the robustness of recognition, and it will be further studied.

Sparse representation is first introduced into face recognition by Wright et al. (2009). In this method, a testing image is first sparsely represented over all training images, and then the classification result is obtained by finding the class, which leads to the minimal representation error. Sparse representation has been widely used in both computer vision and signal processing applications. However, the SRC algorithm has its limitation in some aspects considering the problem of position, illumination and occlusions of the actual database and so on. Hence, a number of researchers have been attracted to tackle these difficulties and many effective methods were proposed. To overcome the problem of illumination variations and occlusions, Yang and Zhang (2010) introduced Gabor features, which make the occlusion dictionary compressible. Yang et al. (2011) and He et al. (2011) studied the robust sparse representation to avoid the utilization of identity occlusion dictionary, so that the speed of sparse representation can be accelerated and the higher recognition rate can be reached, especially when face image is occluded by disguise. Wang et al. (2015) proposed a Manifold Regularized Local Sparse Representation (MRLSR) model, which guarantees that all coding vectors

in sparse representation should be group sparse. It means that each representation meets both individual sparsity and local similarity. Zhuang et al. (2013) proposed a method to train a sparse illumination dictionary to represent the different illumination patterns between the training and testing image named SIT technique. To overcome the limitation of requiring large amounts of samples per class, Deng et al. (2012, 2013) proposed the use of additional dictionaries, constructing it by the use of available data, which refines face recognize performance. A testing sample is represented as a sparse linear combination of the training and intra-class dictionary samples. The additional intra-class dictionary of subject samples can effectively model face variations. The method is called Extended SRC (ESRC). Iliadis et al. (2014) combined the sparsity-based approaches with additional least-squares steps, mitigating the need for a large number of training images since it proves robust to vary number of training samples. Liu et al. (2014) proposed a local structure based sparse representation classification (LSSRC), which divides the face into local blocks, classifies each local block, and then integrates all the classification results by voting. It successfully alleviates the dilemma of high data dimensionality and small samples. To overcome the problem of emphasizing the sparsity too much and over-looking the correlation information, Rigamonti et al. (2011) proposed correlation representation-based classifier (CRC) in consideration of the collaborative representation. Wang et al. (2014) proposed a framework called adaptive sparse representation-based classification (ASRC) in which sparsity and correlation are jointly considered. To address the alignment problem, Liao et al. (2013) developed an alignment-free face representation method based on Multi-Key point Descriptors (MKD). The approach represents each face image with a set of key point descriptors (GTP and SIFT), and constructs a large dictionary from all the gallery descriptors.

Although now big data computing could bring high face recognition rate, it needs strong computing power, and normally works in the cloud. However, in many computer vision applications, especially a lot of front-end application, it needs to quickly and efficiently recognize faces. Inspired by human rapid and accurate identification of familiar faces, we think that there may be a class of fast computing mechanisms that play a role in human face recognition and thus improve the accuracy of recognition. Bonnen et al. (2013) considered this problem from another point of view. They proposed a component-based recognition method. Referring to some psychological research results (Richler et al. 2011; Schwaninger et al. 2006; Piepers and Robbins 2012), this method suggests that the facial features are composed of the “detail” features and the “holistic” features, and the “detail” features occupy

the main position. Then, face recognizing could be realized through extracting “detail” features and discriminating them. Continuing in-depth thinking about this idea, we study the nonlinear least-squares calculation in face recognition application, and find that it really can improve the recognition rate, and more importantly, it can deal with any combination of face features, such as “detail” and “holistic” features, obtaining a high recognition rate. we also study the algorithm of SRC in depth, and find that the SRC algorithm has acceptable robustness for the facial rigid region (e.g. eyes, nose and mouth). What’s more, there is an interesting result that the recognition rate of single facial component using SRC algorithm in actual database is close to the human face recognition. Then we propose a hierarchical face recognition algorithm using nonlinear least-squares named HSRC. HSRC has three steps: the first step is multichannel analysis of the holistic features of one face; the second step is to analyze the components of facial rigid region; the third step is to judge the face by using nonlinear least-squares.

This paper continues as follows. In Sect. 2, we introduce the basic principle and main process of SRC. In Sect. 3, we describe the main contents of HSRC and display the whole process of HSRC algorithm. In Sect. 4, we introduce the experimental results of facial components and judgment by nonlinear least-squares. In Sect. 5, we summarize the whole paper and put forward a direction of further research.

2 Sparse representation-based classification (SRC)

Sparse Representation-based Classification (SRC) is proposed by Yang and John Wright 2009, which is based on the theory of compressed sensing (CS). The main process of SRC is as follows:

1. Form the training set:

$$A = [A_1, A_2, \dots, A_k], \quad A_i (1 \leq i \leq k)$$

A_i are the sample images of the i -th person, and k is the total number of classes. $A_i = [v_{i,1}, v_{i,2}, \dots, v_{i,n}]$, $v_{i,n}$ is the n -th sample image of the i -th person.

2. Use the training set to present the testing image y :

$$y = \sum_{i=1}^k \sum_{j=1}^n \alpha_{i,j} v_{i,j} = Ax$$

$x = [\alpha_1, \alpha_2, \dots, \alpha_k]$ is the sparse solver of the testing image and $\alpha_i = [\alpha_{i,1}, \alpha_{i,2}, \dots, \alpha_{i,n}] (1 \leq i \leq k)$ is the sparse coefficient of the person i . If the class of the testing image is i , y could be sparsely represented only by the sample images of person i in some degree.

3. Get the sparse representation via ℓ_1 norm minimization:

$$\hat{x} = \min \|x\|_1$$

4. Calculate the residual:

$$r_i(y) = \|y - A\delta_i(\hat{x})\|_2, \quad i = 1, \dots, k$$

Confirm the class of the testing image:

$$Identity(y) = \min(r_i)$$

If we want to identify the class of the testing image intuitively, we can calculate the residuals between y and $A \cdot \delta_i(\hat{x})$. $\delta_i(\hat{x})$ denotes the coefficients sequence, where $\alpha_j = 0, (j \neq i)$, and $\alpha_i \neq 0$.

3 A hierarchical face recognition algorithm based on nonlinear least-squares (HSRC)

SRC algorithm has been widely used in both computer vision and signal processing applications, and the recognition rate is really high in some databases such as ORL. However, in other databases, such as actual environment databases, SRC algorithm encounters severe performance degradation, just as Wu et al. (2014) analyzed. To deal with the performance degradation problem, SRC algorithm needs to relax the confirm conditions while recognizing holistic face pictures, and to enhance the accuracy rate while recognizing the face details (Wu et al. 2014). Along this technical route, we further suggest that, in recognition of the top floor, people recognize the face through combining the “detail” and “holistic” features. And the “detail” features, such as mouth, eyes, may be accurately identified by Sparse Representation. So, we design a new hierarchical face recognition algorithm named as HSRC. Hierarchy is embodied by the analysis of holistic face features and facial “detail” features, and we use nonlinear least-squares to combine these features in order to enhance the recognition performance.

HSRC has three steps:

Step 1: Multichannel analysis. We use bidirectional PCA, linear discriminant analysis and GradientFace to extract more “holistic/configural” face features, then select some face images close to the testing image to generate 1-step face database, which will be used as sample images for the next operation.

Step 2: Extract facial components. After handling the 1-step face database with Self Quotient Image (SQI) (Wang et al. 2004), we extract the facial components by using a new proposed algorithm inspired by ESRC to generate 2-step face database, which includes four sub-databases: left eye database, right eye database, nose database and mouth database.

Step 3: Judge by nonlinear least-squares. We use SRC algorithm in each sub-database, and combine these

recognition results with nonlinear least-squares method, which promotes the recognition rate.

The flow diagram of HSRC is shown in Fig. 1.

3.1 Multichannel analysis

The first step of HSRC follows the first step of T_SRC (Wu et al. 2014). We select three algorithms as the basic algorithms for the multichannel analysis. They are bidirectional PCA (Zuo et al. 2006), linear discriminant analysis and GradientFace (Zhang et al. 2009). Considering that each basic algorithm can capture part of the “holistic” facial information, combining multi parts can get more complete “holistic” facial information. We propose multichannel analysis. Three algorithms, BDPCA, PCA + LDA and GradientFace + PCA + LDA are used respectively in each channel to recognize the face. Then we combine the candidates as the 1-step face database, which are the most likely 1–3 face images in each channel. By relaxing the candidate constraints, HSRC introduces more “holistic” facial information as recognition conditions.

3.2 Extract facial components

Mainly based on the results of psychology research, HSRC extracts facial components as facial “detail” features. The existing methods may not have the ideal effect of locating facial components. The Core Image Programming of Apple includes two parts: one is detecting face in an image and the other is getting face components. It has an ideal result in getting face components on the premise of recognizing the whole face-contour correctly. Otherwise it cannot work well. Gabor filter is used to detect edges in image processing, and it could be used to detect facial components too. But the accuracy of detection would be affected by

eyebrows, spectacle frames or cheekbones, which will result in the deviation of detecting. Adaboost algorithm with Haar features is usually used in face detection, but it does not have an ideal result when it is used in facial components detection. In this paper, we propose a new method to extract facial components, which is multi-adaptive and robust.

The motivation for our proposing algorithm comes from the fact that different images of the same subject share a lot of similarities. Since facial components such as eyes, noses and mouths from different people look alike in some degree. They have the similar structure and contour line. On the basis of this point, we can select existing facial components to form an average template to locate the new facial components by using SRC.

The model of extracting facial components is showed as below:

$$y = Bx_0 + D\xi_0 + \varepsilon \tag{1}$$

Bx_0 denotes the average facial components. $D\xi_0$ denotes the variant individual deviation and is the small dense noise. B and D are dictionaries of the average facial components and the individual deviation respectively. x_0 and ξ_0 are the sparse representations, which could be recovered simultaneously by ℓ_1 minimization. (Bx, Δ_1) and $(D\xi, \Delta_2)$ are two thresholds. If $|Bx_0 - Bx| \leq \Delta_1$ and $|D\xi_0 - D\xi| \leq \Delta_2$, the testing samples could be accepted as the facial component represented by Bx .

The algorithm steps are described as follows:

First, to remove the effect of varied illumination, we preprocess the images in 1-step face database by using SQI. Figure 2 is the result of SQI extracting illumination invariant before and after.

Second, after SQI dealing with the image, we cut the image into four big blocks, and each block includes left

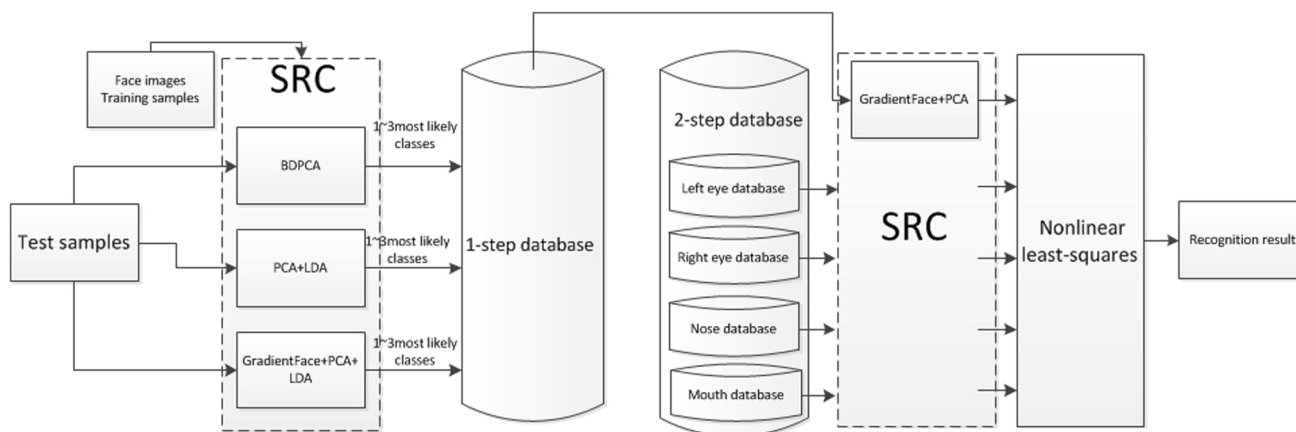


Fig. 1 The main process of HSRC algorithm



Fig. 2 Before and after handling with SQI

eye, right eye, nose and mouth respectively. Figure 3 shows the result. Then we use a rectangle with the size of $M \times N$ to traverse these big blocks every pixels from x axis and y axis and extract many small blocks. These small blocks are used as training samples. In this paper, $M = N = 31$ Pixels.

Third, we select some possible facial component images with the size of $M \times N$ as testing samples, and find out the most likely sample by using training samples and SRC.

The algorithm process is expressed as follows:

1. Input: a matrix of training samples

$$A = [B_1 + D_1, B_2 + D_2, \dots, B_k + D_k]$$

B_i is the average facial feature; $D_i(1 \leq i \leq k)$ is individual deviation; y is a testing sample.

2. Normalize the columns of A and y .
3. Solve the ℓ_1 minimization problem:

$$\hat{x} = \arg \min \|x\|_1 \quad s.t. \quad y = Ax$$

4. Compute the residuals:

$$r_i(y) = \|y - A\delta_i(\hat{x})\|_2, \quad i = 1, \dots, k$$

where

$$\delta_i(\hat{x}) = [0, 0, \dots, x_i, 0, \dots, 0], \quad x_i = x_{i1}, \dots, x_{im}$$

coefficient m is corresponds to the m training samples coming from the No. i individual.

5. Output:

$$Identity(y) = \arg \min_i (r_i(y))$$

According to this algorithm, we locate the facial components successfully and use these images as 2-step face

Fig. 3 Four big blocks



Table 1 The accuracy of the algorithm of extracting facial components (%)

	Left	Right	Nose	Mouth	Average
Actual	88.90	86.70	96.40	96.10	92.03
ORL	95.80	97.00	94.30	97.80	96.23
YaleB	85.80	85.00	99.50	93.40	90.90

database. The accuracy of this algorithm is shown in Table 1, which is judged by people.

3.3 Judge by nonlinear least-squares

How to combine some facial components into a holistic one is still an open research problem. We use nonlinear least-squares method to combine the facial components judgements, and get a good holistic recognition rate. There is an interesting phenomenon that although each facial component judgement is not so accurate, just like the human beings, the holistic judgement using nonlinear least-squares method is in high accuracy, close to human face recognition.

3.3.1 Nonlinear least-squares method

Considering a curve $y = f(x, \beta)$ that relies on variable and parameter β , we hope the curve fits the data in the meaning of least squares:

$$Q = \sum_{k=1}^m [y_k - f(x_k, \beta)]^2 \tag{2}$$

x_k is the No. k factor; y_k is the actual score of No. k factor; $f(x_k, \beta)$ is the target score of No. k factor; β is the weight value of each factor. Q is the value of nonlinear least-squares error.

3.3.2 Judge by nonlinear least-squares

2-step face database involves four subdatabases (left eye, right eye, nose, mouth), which only include the facial components information, but lack the “configural” face information. So we introduce GradientFace to supplement the holistic information of face. We respectively apply SRC algorithm to these five factors mentioned above and

get the recognition results. First of all, according to our actual experience, the weights β of the factors of eyes, nose, mouth and Gradientface are assigned with 0.2 0.3, 0.1 0.2, 0.1 0.2, 0.2 0.25 in turn. And then, for one testing sample, we extract 5 component factors, namely left eye, right eye, nose, mouth and Gradientface. For No. $k(1 \leq k \leq 5)$ factor of each testing sample, we sort the similarity of all training samples according to the residual errors, and assign a “similar score” to each class, which is represented by the absolute differences of $y_k - f(x_k, \beta)$. We stipulate that the more similarity corresponds to lower scores. The specified rule is described as follows:

1. Initialize each class:

$$\text{similar_score}_{class_j} = S; \quad 1 \leq j \leq n$$

where n denotes the total number of classes, and $class_j$ denotes the j -th class.

2. Each factor of the testing sample, such as left eye, will be compared with all training samples and according to the residual errors, reassign the “similar score” of the corresponding class. Specific steps are as follows:

```
Sort the residual errors from low to high;
Extract each training sample corresponding to each residual error sequentially;
repeat
  if the current training sample belong to the  $j$ -th class
  then
     $\text{similar\_score}_{class_j} = \text{similar\_score}_{class_j} + i - N$ 
  end if
   $i++$ 
until Traverse all training samples
Where  $i$  is the serial number of residual errors sorted from low to high,  $N$  is constant.
```

Now for a testing sample, each class corresponding to each factor is assigned to a “similar score”. Then we use Eq. (2) to calculate the scores for each class, and the class corresponding to minimum value Q_{\min} is the final recognition result.

4 Experimental results and analysis

4.1 Experiment databases

For the purpose of evaluating the performance of the proposed scheme HSRC, we conduct the experiments on three typical face image databases: ORL database, extended YaleB database and the actual environment database. There are 40 people in total and each person has ten images in ORL database. The size of the image is 112×92 . In the Extended YaleB database, there are 39 people in total and 25 images per person, and the size of an image is 192×168 . In order to compare them conveniently, we

select ten images (the same images as ORL) from 25 images per person randomly. The actual environment database is collected by our computer front-camera via OpenCV. There are 36 persons and 10 images per person in the actual environment database. And the size of an image is 128×128 . The face databases are shown in Fig. 4.

4.2 Experimental results of HSRC 1-step and 2-step

Multichannel analysis is the 1-step process of HSRC algorithm, and the result is 1-step database, which includes several most likely classes. No. 1–5 images are selected from all databases as the training set and the rest of the five images are taken as testing images. The actual environment database uses three most likely classes in each channel. Then we make these classes a union set to generate the 1-step database. The 1-step recognition rate in actual database is 95.5 %.

Although the 1-step HSRC has high recognition rate, it is not the real recognition rate. It just means that some most likely classes include the right class of the testing image. So the next work is to choose the correct class from these most likely classes through the perspective of facial components.

Next we elaborate the process of extracting facial components by using actual environment database as an example. First, we use SQI to handle the 1-step database. The size of images are 128×128 and we cut them into four big blocks with the size of 64×64 , each of which includes left eyes, right eyes, noses and mouths respectively. Through using 31×31 rectangles to traverse these big blocks, we pick out many small blocks as training samples. According to the training samples, we could locate the facial component in testing samples by using the method referred in Sect. 3.2. Figure 5 shows the part results of locating partial components of face in actual environment database. The facial components extracted by the above method are collected and stored in the 2-step database.

the 2-step database involves database of left eye, database of right eye, database of nose and database of mouth. In each subdatabase, we respectively select 5, 6, 7, 8 images as training samples, with the rest as testing samples. Tables 2, 3 and 4 illustrate the accuracy rates of the eyes components recognition in the 2-step databases by using SRC algorithm. The first lines of Tables 2, 3, 4, 5, 6 and 7 denote the different number of the training sample images 5, 6, 7, and 8.

Tables 3 and 4 show that the recognition rates of facial components are really high. The reason is that when we use three basic algorithms for multichannel analysis, the recognition rates in each channel are high enough for ORL database and Extended YaleB database. The most likely

Fig. 4 Part images of three databases

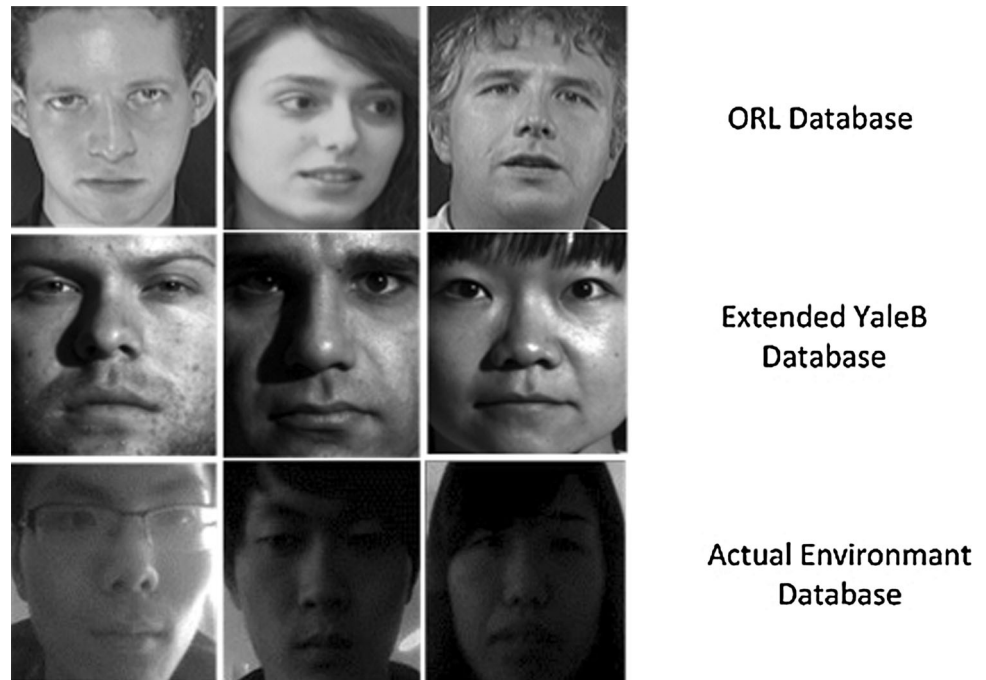


Fig. 5 Part facial components located

Table 2 The accuracy rate of the components recognition in actual environment database (%)

	5	6	7	8
Left eye	66	69	75	78
Right eye	58	63	73	69

Table 3 The accuracy rate of the components recognition in ORL database (%)

	5	6	7	8
Left eye	85	86	90	96
Right eye	87	89	93	96

Table 4 The accuracy rate of the components recognition in Extended YaleB database (%)

	5	6	7	8
Left eye	91	90	91	100
Right eye	91	90	91	100

classes that each channel select concentrate on the correct class. It means that the alternatives of the 2-step database are very small and always correct. Thus, the components recognition can get high accuracy rate. Table 2 shows that the recognition rate of eyes by using SRC algorithm in actual database ranges from 65 to 75 %, which is close to

Table 5 The recognition rate of actual environment database (%)

	5	6	7	8
SRC + PCA	48.3	53.5	59.3	63.9
SRC + BDPCA	60.6	61.1	72.2	75.0
SRC + LDA	71.7	78.5	77.8	81.9
SRC + GradientFace + LDA	68.3	74.3	75.9	73.6
T_SRC	83.3	84.0	86.1	90.3
HSRC	83.9	86.8	88.9	91.7

the accuracy of human facial component recognition (Bonnen et al. 2013).

4.3 Experimental results of HSRC 3-step

Next, we assign similar scores to all classes and distribute the weights for each factor referred to the rules in Sect. 3.3. Specifically, S is 15, N is 5, i is 1 to 5 and the weights for each factor are 0.25, 0.25, 0.15, 0.15, 0.2 in turn. Using nonlinear least-squares, the facial component recognitions can be combined into one result, which is the face recognition rate. We compare the recognition results of HSRC

Table 6 The recognition rate of ORL database (%)

	5	6	7	8
SRC + PCA	90.5	94.3	94.2	95.0
SRC + BDPCA	92.5	96.2	97.5	97.5
SRC + LDA	85.0	88.8	92.5	91.3
SRC + GradientFace + LDA	69.5	72.5	80.0	85.0
T_SRC	91.0	91.8	95.0	96.3
HSRC	90.5	92.0	93.6	97.0

Table 7 The recognition rate of Extended Yaleb database (%)

	5	6	7	8
SRC + PCA	80.5	88.8	97.4	100
SRC + BDPCA	73.2	59.5	77.2	100
SRC + LDA	90.5	90.2	91.2	100
SRC + GradientFace + LDA	98.0	100	100	100
T_SRC	97.9	99.4	98.3	100
HSRC	97.5	99.4	99.5	100

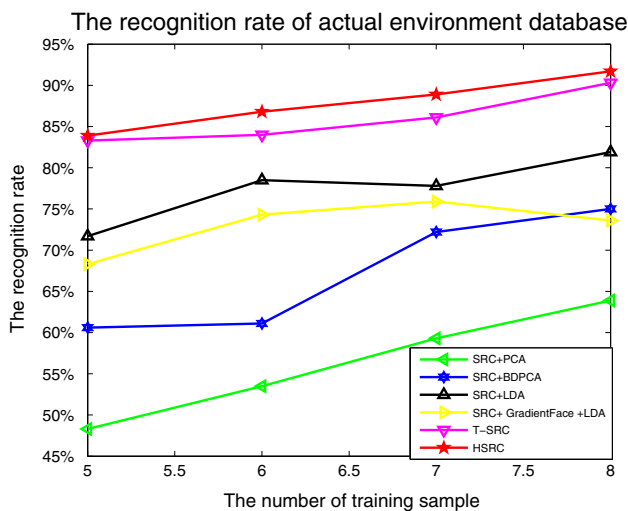


Fig. 6 Part facial components located

with that of some proposed algorithms, which are shown in Tables 5, 6 and 7 and Figures 6, 7 and 8.

As is shown to us, in actual database, HSRC does have promotion compared to the other algorithms due to its robustness. It can accept the low accuracy of one component recognition, but the accuracy of combining all components is high. If we could extract the facial component more accurate, the recognition rate of HSRC would be higher. Comparing with SRC+BDPCA and T_SRC, HSRC does not perform so well in ORL database, and the reason from our inference is that the accuracy of extracting facial components is not so ideal. More components identification errors will affect the robustness of HSRC. The accuracy of

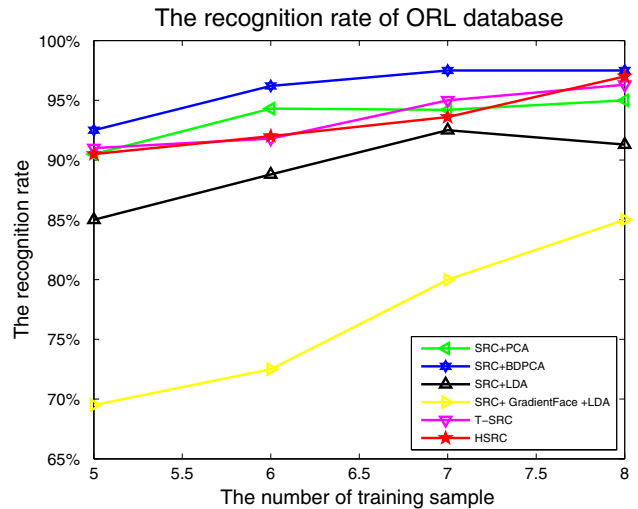


Fig. 7 Part facial components located

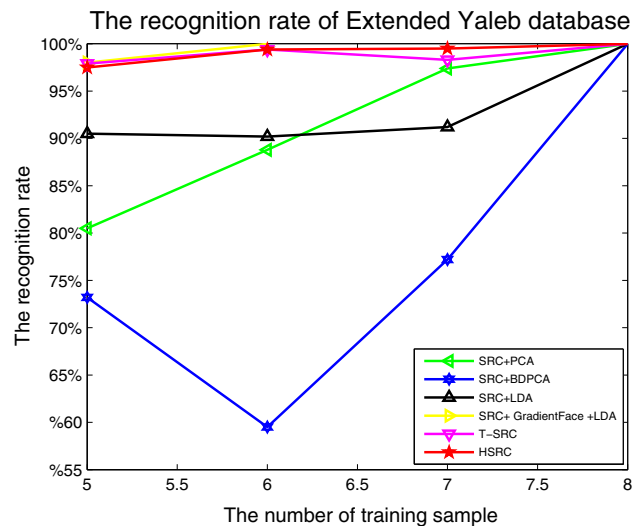


Fig. 8 Part facial components located

HSRC recognition is really high in Extended YaleB database, because the main problem of the database, varied illumination, could be solved by SQL. Summing up the above discussion, generally, HSRC algorithm has really good performance. Moreover, we further analyze the robustness of HSRC through extracting facial components more accurately. We manually locate the exact coordinates of facial components in actual database, and use HSRC to recognize the testing images. The results are shown in Table 8. When the training images are 5 or 6, the performance improvement with manual extracting is obvious, and the recognition rates using 5–8 training images are close. It can be inferred that if the number of training images reaches a certain threshold, HSRC algorithm can achieve the same accuracy of manual extraction of facial

Table 8 The recognition rates of HSRC and HSRC (manually) in actual database (%)

	5	6	7	8
HSRC	83.9	86.8	88.9	91.7
HSRC (manually)	87.8	89.6	88.9	91.7

components. Furthermore, it is generally believed that the location of the facial components is accurate when people recognize face image, and the recognizing results is robust when the face image is familiar. This effect can be achieved by HSRC algorithm to some extent. The crucial part of the algorithm is that nonlinear least-squares method can provide good robustness while recognizing face.

5 Conclusions

With the increasing face recognition rate, the technology of face recognition is widely used in our life, such as public security forensic and management of houses security. Recently, people do research and propose new methods to solve the problem of variant illumination, posture and incomplete face in face recognition. We focus on the problem whether proposed face recognition algorithm can establish correlation with human face recognition, and it may bring us more efficient algorithms to recognize familiar faces. We study the algorithm of SRC in depth and find out that the recognition rate of facial components by the use of SRC algorithm is close to the human recognition rate. Then we propose a hierarchical face recognition algorithm based on nonlinear least-squares named as HSRC. HSRC can reduce the effect of variant illumination, posture and incomplete face image and improve the recognition rate. Experiments prove that HSRC has a better performance than SRC and T_SRC have.

However, only combining the facial components features in HSRC may not be enough to get a higher recognition rate. In the next phase of research, we will try to introduce the “configural” information among the facial components into the face recognition. We believe there is still room for improvement in face recognition, and we will do better in further research.

Acknowledgments We gratefully acknowledge funding support from the Major Program of National Social Science Foundation of China (No. 11&ZD088) and the Zhejiang Province Science and Technology Innovation Program under grant number 2013TD03 and the Zhejiang province science and technology Project No. 2013C33056.

References

Ataer-Cansizoglu E, Akcakaya M, Orhan U, Erdogmus D (2014) Manifold learning by preserving distance orders. *Pattern Recogn Lett* 38:120–131

- Bhatia KK, Rao A, Price AN, Wolz R, Hajnal JV, Rueckert D (2014) Hierarchical manifold learning for regional image analysis. *IEEE Trans Med Imaging* 33(2):444–461
- Bonnen K, Klare BF, Jain AK (2013) Component-based representation in automated face recognition. *IEEE Trans Inf Forens Secur* 8(1):239–253
- Ciresan D, Meier U, Schmidhuber J (2012) Multi-column deep neural networks for image classification. In: 2012 IEEE conference on computer vision and pattern recognition (CVPR). IEEE, pp 3642–3649
- Deng W, Hu J, Guo J (2012) Extended SRC: undersampled face recognition via intraclass variant dictionary. *IEEE Trans Pattern Anal Mach Intell* 34(9):1864–1870
- Deng W, Hu J, Guo J (2013) In defense of sparsity based face recognition. In: 2013 IEEE conference on computer vision and pattern recognition (CVPR). IEEE, pp 399–406
- Farabet C, Couprie C, Najman L, LeCun Y (2013) Learning hierarchical features for scene labeling. *IEEE Trans Pattern Anal Mach Intell* 35(8):1915–1929
- He R, Zheng WS, Hu BG (2011) Maximum correntropy criterion for robust face recognition. *IEEE Trans Pattern Anal Mach Intell* 33(8):1561–1576
- Hinton GE, Salakhutdinov RR (2006) Reducing the dimensionality of data with neural networks. *Science* 313(5786):504–507
- Iliadis M, Spinoulas L, Berahas AS, Wang H, Katsaggelos AK (2014) Sparse representation and least squares-based classification in face recognition. In: Signal processing conference (EUSIPCO), 2013 proceedings of the 22nd European. IEEE, pp 526–530
- Kavukcuoglu K, Sermanet P, Boureau YL, Gregor K, Mathieu M, Cun YL (2010) Learning convolutional feature hierarchies for visual recognition. In: Advances in neural information processing systems, 23, Curran Associates, Inc. pp 1090–1098
- Liao S, Jain AK, Li SZ (2013) Partial face recognition: alignment-free approach. *IEEE Trans Pattern Anal Mach Intell* 35(5):1193–1205
- Liu F, Tang J, Song Y, Xiang X, Tang Z (2014) Local structure based sparse representation for face recognition with single sample per person. In: 2014 IEEE international conference on image processing (ICIP). IEEE, pp 713–717
- Lu J, Tan YP, Wang G (2013) Discriminative multimaniifold analysis for face recognition from a single training sample per person. *IEEE Trans Pattern Anal Mach Intell* 35(1):39–51
- Piepers DW, Robbins RA (2012) A review and clarification of the terms holistic, configural, and relational in the face perception literature. *Front Psychol* 3:559. doi:10.3389/fpsyg.2012.00559.
- Richler JJ, Cheung OS, Gauthier I (2011) Holistic processing predicts face recognition. *Psychol Sci* 20:464–471
- Rigamonti R, Brown MA, Lepetit V (2011) Are sparse representations really relevant for image classification? In: 2011 IEEE conference on computer vision and pattern recognition (CVPR). IEEE, pp 1545–1552
- Schwaninger A, Wallraven C, Cunningham DW, Chiller-Glaus SD (2006) Processing of facial identity and expression: a psychophysical, physiological, and computational perspective. *Progr Brain Res* 156:321–343
- Silva VD, Tenenbaum JB (2003) Global versus local methods in nonlinear dimensionality reduction. In: Advances in neural information processing systems, vol 15. MIT Press, Cambridge, USA, pp 721–728
- Taigman Y, Yang M, Ranzato M, Wolf L (2014) Deepface: closing the gap to human-level performance in face verification. In: 2014 IEEE conference on computer vision and pattern recognition (CVPR). IEEE, pp 1701–1708
- Tenenbaum JB, De Silva V, Langford JC (2000) A global geometric framework for nonlinear dimensionality reduction. *Science* 290(5500):2319–2323

- Wang J, Lu C, Wang M, Li P, Yan S, Hu X (2014) Robust face recognition via adaptive sparse representation. *IEEE Trans Cybern* 44(12):2368–2378
- Wang L, Wu H, Pan C (2015) Manifold regularized local sparse representation for face recognition. *IEEE Trans Circuits Syst Video Technol* 25(4):651–659
- Wang H, Li SZ, Wang Y, Zhang J (2004) Self quotient image for face recognition. In: 2004 international conference on image processing, 2004, ICIP'04, vol 2. IEEE, pp 1397–1400
- Wright J, Yang AY, Ganesh A, Sastry SS, Ma Y (2009) Robust face recognition via sparse representation. *IEEE Trans Pattern Anal Mach Intell* 31(2):210–227
- Wu Z, Yu Z, Yuan J, Zhang J (2014) A twice face recognition algorithm. *Soft Comput*. doi:[10.1007/s00500-014-1561-9](https://doi.org/10.1007/s00500-014-1561-9)
- Yang M, Zhang L (2010) Gabor feature based sparse representation for face recognition with gabor occlusion dictionary. In: *Computer vision-ECCV 2010*. Springer, Berlin, pp 448–461
- Yang M, Zhang D, Yang J (2011) Robust sparse coding for face recognition. In: 2011 IEEE conference on computer vision and pattern recognition (CVPR). IEEE, pp 625–632
- Zhang T, Tang YY, Fang B, Shang Z, Liu X (2009) Face recognition under varying illumination using gradientfaces. *IEEE Trans Image Process* 18(11):2599–2606
- Zhuang L, Yang AY, Zhou Z, Sastry SS, Ma Y (2013) Single-sample face recognition with image corruption and misalignment via sparse illumination transfer. In: 2013 IEEE conference on computer vision and pattern recognition (CVPR). IEEE, pp 3546–3553
- Zuo W, Zhang D, Yang J, Wang K (2006) Bdpca plus lda: a novel fast feature extraction technique for face recognition. *IEEE Trans Syst Man Cybern Part B: Cybern* 36(4):946–953