Journal of Computers Vol. 30 No. 4, 2019, pp. 240-252
doi:10.3966/199115992019083004023

Face Recognition Via Segmentation Processing Based on the Fusion of Residual

Jie Gao[1*], Liquan Zhang[1], Yujuan Hao[1]

[1] Department of Electrical and Information Engineering, Beijing University of Civil Engineering and Architecture, Beijing 100044, China
{gjinchina, Zhang_lq2000, 18810973259}@163.com

Received 7 March 2019; Revised 7 May 2019; Accepted 17 May 2019

Abstract. Sparse representation is an efficient image representation method, the face recognition algorithm based on sparse representation is widely used in the recognition and classification of face images because it has the advantages of being insensitive to feature extraction and easy to construction algorithm model. However, the original algorithm does not effectively discriminate non-face images. At the same time, the algorithm is based on the whole face, ignoring the positive influence of local features on the whole recognition process. Aiming at the problems existing in the classical sparse representation method, we propose the face recognition algorithm via segmentation processing based on the fusion of residual. First, the test image is identified by the improved ℓ_2-norm value detection method, then the test image which is judged to be legal is processed into non-superimposed blocks and finally to be classified. This method reduces the negative impact of the noise-containing sub-blocks on the whole face recognition process and enhances the positive influence of the feature significant sub-blocks. In addition, under the premise of ensuring the recognition rate and controlling the calculation cost, the optimal segmentation strategy is proposed. The comparison experiments were carried out in the standard Extended Yale B face database and AR face database, the experiment results show that the proposed algorithm achieves better recognition effect.

Keywords: face recognition, non-face image, residual, segmentation processing, sparse representation

1 Introduction

As a research hotspot in the field of machine vision, face recognition has always been a classic research topic in the field of image processing. Because of its advantages of low cost and easy operation, face recognition has been widely applied in security monitoring, identity authentication and other fields, and has attracted much attention for many years. In the past, the signal processing methods based on data have been well developed and applied. Such as principal component analysis (PCA) [1] is widely used due to its good performance, Yang, et al. [2-3] proposed two-dimensional PCA, which is directly based on two-dimensional image matrix rather than one-dimensional vector. Before feature extraction, it is no longer necessary to convert images into vectors, thus improving the classification efficiency. Linear Discriminate Analysis (LDA), a supervised classification method, is called "the most discriminating feature" [4-5]. The basic principle is to make the projected data have smaller dispersion within the class and the maximum dispersion among the classes.

Serre T. [6] proposed that human visual system showed sparse characteristics to images. Olshausen B, et al. [7] pointed out that the natural image itself has sparse representation. Therefore, the idea of sparse has become a hot issue in statistics and machine learning. It has been successfully applied in the fields of computer vision, pattern recognition and image analysis, such as image analysis [8-9], image denoising [10], image super-resolution processing [11], and visual saliency detection [12]. "Sparse" is an important

* Corresponding Author

concept in many scientific fields. It appears in statistics, signal processing and machine learning as a variable or a problem of feature selection.

John Wright. [13] firstly applied compressed sensing theory to the field of visual recognition and face recognition, and conducted experiments on several internationally recognized databases. The results show that in the case of more samples, this method not only provides a simple and intuitive new method for face image recognition, but also can solve the difficulties of test samples with partial occlusion to a certain extent, thus the robust face recognition is preliminarily realized.

In recent years, sparse representation theory has developed rapidly, and sparse representation has been successfully applied to face recognition classification [14-16]. The traditional classification method (SRC) based on sparse representation of ℓ_1-norm can effectively solve the problems of image illumination, occlusion and noise without feature extraction [17]. Because of its robustness, a series of improved face recognition methods based on sparse representation are proposed. Zhou et al. [18] combined sparse representation with Markov random field to identify partially camouflaged faces. Yang et al. [19] add Gabor features to sparse representation, which not only improves the distinguishing ability of the face but also reduces the dimension of over-dictionary. Deng et al. [20] proposed a representation model of Prototype plus Variation to solve the problem of single-sample face recognition. Georghiades et al. [21] proposed a face recognition method based on Illumination Cones model with multi-illumination and multi-pose conditions. Blanz and Vetter [22] proposed a face image analysis and recognition method based on 3D deformation model with multi-pose and multi-illumination conditions, which achieved high recognition rate in international face database. Wang [23] uses weighted norms for sparse constraints. Fand [24] proposed a general sparse regularization term modeling framework to design sparse regularization terms, and the Bayesian method was also used to represent sparse regularization conditions. Hao [25] proposed a sparse representation method (C-SRC) based on the fusion of sparse coefficients and residuals. The method proposed that the sparse coefficients of images should be taken as the intra-class mean, and the correlation between the inter-class features and the overall features of the sparse coefficients should be added to the residual values, and the method of calculating the residual is redefined to improve the fidelity of the residual.

In the field of face recognition research based on sparse representation, simulation experiments are performed in standard face databases of all categories. When a test face image is arbitrarily input, the recognition result necessarily corresponds to the training sample. However, in practical applications, the category of the picture participating in the recognition classification is unknown, there is a case where the input test image is a non-face image, or the image participating in the recognition does not belong to any type of face image in the database. However, according to the sparse algorithm identification, the system still calculates according to the original algorithm flow, and obtains the image belonging to the training dictionary according to the sparse reconstruction. The algorithm system does not detect the validity of the input image, and obviously it is wrong to determine the class by relying solely on the residual. It can be seen that in order to avoid non-human picture interference, when performing face recognition, we absolutely have to consider how to distinguish and exclude non-human face test images, which is also an indispensable key in face recognition algorithms.

At the same time, the existing algorithm basically achieves a better recognition effect in a controlled environment. However, when there are expression changes and occlusion in the test image, the recognition rate of the algorithm will be drastically reduced. The traditional algorithm based on sparse representation usually constructs the face dictionary into a column vector when the face dictionary is constructed. Such a deformation processing method necessarily causes some local feature information to be lost. However, people can correctly identify different faces, precisely because people’s facial features are regarded as important information to distinguish different faces, such as human eyebrows, eyes, nose, mouth and even scars on the face, etc. How to highlight these feature information will directly affect the success rate of recognition. In order to highlight the importance of local features, this paper considers an improved method of blocking. The improved algorithm highlights the local feature information of the face. After the picture is processed into blocks, the resolution of the sub-blocks with extremely uneven illumination or occlusion is relatively poor, but the recognition rate of other sub-blocks without noise interference is not affected, even these features are obvious. The recognition result of the sub-blocks can influence the final recognition accuracy rate. Finally, the recognition results of all the sub-blocks can be combined to improve the probability of obtaining the correct recognition result.

In this paper, the main novelty and contribution are summarized as follows:

1. In order to solve the problem that the non-face image is not judged in the classical face recognition algorithm, an improved ℓ_2-norm value detection method for distinguishing non-face images is proposed. By training non-face and face samples, the characteristics of the sparse coefficients are analyzed and the proposed classification threshold is obtained. Comparing the sparse coefficient ℓ_2-norm detection value with the classification threshold, when the detection value is greater than or equal to the threshold rule, it is a valid human face, otherwise, it is determined that the test image is a non-face image. This method achieves effective discrimination of non-face images and optimizes the classification effect of the face recognition system.

2. This paper analyzes the negative effects of using the entire face to participate in the face recognition process when the test sample contains noise such as uneven illumination and facial occlusion. Based on the C-SRC algorithm, a segmented face recognition algorithm based on the fusion of residual is proposed. This method segment the test image and the training image into blocks based on the C-SRC algorithm, and performs non-superimposed training in the overcomplete dictionary composed by each sub-block to obtain the classification residual of each sub-block. We use the new residuals as the classification criterion which is obtained by fusing the residuals of each sub-block. This method completely preserves the sub-block information without interference factors, and reduces the negative influence of the noise-containing sub-block on the entire face recognition process.

3. We proposed the optimal face segmentation strategy, by adopting an appropriate face image segmentation method, the integrity of the important organ information of the face can be preserved, and the feature block and the noise-containing block can be effectively segmented. The robustness simulation experiments are carried out in the AR face database and the Extended Yale B face database, the face samples are processed in multiple ways by using the block recognition algorithm, and finally acquired the optimal face segmentation strategy by their best recognition rate. The optimal face segmentation strategy optimizes the block face recognition algorithm, so its still has a good recognition rate in an uncontrollable environment.

2 Face Recognition Based on Sparse Representation

The key to the SRC algorithm lies in the sparsity of the signal, its core idea is to obtain most of the effective information of the target signal through the linear representation of several basic signals. In the process of signal representation and processing, sparse representation method is used to process the data, which not only reduces the cost of processing the data signal, but also improves the efficiency of signal processing.

Given a training set D, which is a concatenation of all k object class, is defined by using the entire training set, since all images of class i form a set $D_i=[v_{i,1},v_{i,2},\cdots,v_{i,n}]\in R^{m\times n_i}$, $v_{i,j}$ represents the j-th face image of the i-th subject, the image to be tested $y\in R^m$ can be expressed as:

$$y=\varphi_{i,1}v_{i,1}+\varphi_{i,2}v_{i,2}+\cdots+\varphi_{i,n_i}v_{i,n_i}, y\in R^m . \tag{1}$$

where the scalars $\varphi_{i,j}\in R, j=1,2,\cdots n_i$, the dictionary matrix composed of all the pictures in the library can be expressed as:

$$D=\left[D_1,D_2,...,D_k\right]=\left[v_{i,1},v_{i,2},...,v_{k,n_k}\right]. \tag{2}$$

Usually, the given sample data are accompany noisy, and hence the linear combination of y can be represented as:

$$y=D\alpha+z\in \mathrm{R}^m . \tag{3}$$

Where $\alpha=\left[0,...,0,\varphi_{i,1},\varphi_{i,2},...\varphi_{i,n_i},0,...,0\right]^T\in R^n$ denotes a coefficient vector. According to the sparse representation, the sparse solution can be obtained by solving the following ℓ_1-minimization problem:

$$(\ell_1): \hat{\alpha}_1 = \arg\min \|\alpha\|_1 \; s.t. D\alpha = y . \quad \textbf{(4)}$$

The object class is referred to the minimum value of residuals, which is defined as following:

$$\min_i r_i(y) = \|y - D\delta_i(\hat{\alpha}_i)\|_2 . \quad \textbf{(5)}$$

Where $\delta_i(\hat{\alpha}_i) \in R^n$ is the vector whose only nonzero terms are the terms in α that are associated with the i -th class.

The classical sparse representation method, the class with the smallest residual value as the final recognition result, achieved a good recognition effect under ideal conditions. However, when the number of samples in the face database is insufficient, or there are too many interfering factors in the test images to collect too little effective information, the residual values are very close, resulting in unsatisfactory classification effect. Hao [25] use the characteristics of the intra-class mean to neutralize the negative effects of the uncorrelated extreme values:

$$d_i = \sum_{j=1}^{n_i} \varphi_{i,j} / n_i . \quad \textbf{(6)}$$

Combine the intra-class mean with the original residual:

$$r_i = d^p / [\|y - D\delta_i(\hat{\alpha}_i)\|_2] = (\sum_{j=1}^{n_i} \varphi_{i,j} / n_i)^P / [\|y - D\delta_i(\hat{\alpha}_i)\|_2] . \quad \textbf{(7)}$$

The improved residual uses the principle that the maximum value is divided by the minimum value to obtain the maximum value, The maximum value of the residual is used as the final recognition result, that is $\max_i r_i(y) = d^p / [\|y - D\delta_i(\hat{\alpha}_i)\|_2]$, the C-SRC algorithm can well eliminate the influence of interference noise, and the residual contrast is more obvious, reflecting better robustness.

3 Improved Non-face Image Recognition Method

3.1 Characteristic Analysis of Non-face Image

As we all know, the face recognition system usually performed in the face database of the known category, and the recognition result corresponds to a certain type of person in the training sample. In practical, the categories of images participating in the recognition classification are unknown, when inputting non-face images, the system will still calculate according to the original algorithm. Obviously, in this case, it is wrong to obtain the recognition result by only relying on the value of the residual.

In order to analyze the influence of non-face images on recognition, we chose ORL database, randomly select half of samples as training sample, others as test samples. We chose a non-face image for testing, as shown in Fig. 1(a). By using C-SRC algorithm, the sparse reconstruction results are shown in Fig. 1(b), Fig. 1(c)is the corresponding sparse coefficient, and Fig. 1(d) is the corresponding residual.

When a non-face image is selected as the test sample, its sparse coefficient will no longer be sparse in Fig. 1(c). There are many positive coefficients with large values distributed irregularly in each category, and non-zero coefficients are also randomly distributed throughout the training sample. Even if a non-face image is input as a test image, the system will still select the category with the largest residual value as the final recognition result according to the algorithm., but the recognition result at this time must be wrong. Therefore, when performing face recognition, it is absolutely necessary to distinguish and exclude non-face images.

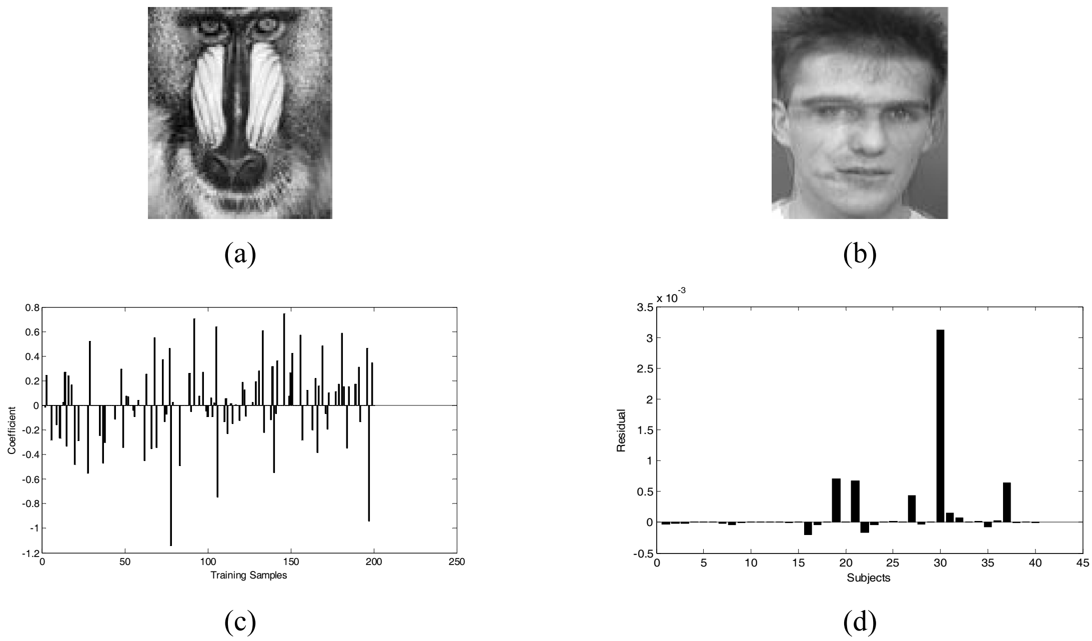


Fig. 1. Recognition of non-face images

3.2 Improved Non-face Image Discrimination Method

The classification recognition result of the C-SRC algorithm is determined by the residual: $r_i = d^p / [\|y - D\delta_i(\hat{\alpha}_i)\|_2] = (\sum_{j=1}^{n_i} \varphi_{i,j} / n_i)^P / [\|y - D\delta_i(\hat{\alpha}_i)\|_2]$, therefore, the residual value $r_i(y)$ only changes with the variation of the sparse coefficient $\delta_i(\hat{\alpha}_i)$, so the value of the residual is actually determined by the sparse coefficient vector α. Further research found the following rules: When the test image is a non-face image, the difference of the ℓ_2-norm value of the sparse coefficient vector corresponding to each class is not obvious, especially when the norm value is very close. On the contrary, when the test image is a face image, the ℓ_2-norm value of the sparse coefficient vector of a certain kind of people is always significantly larger than other classes.

Select the image of the first class of person in the ORL face database as a test sample, as show in Fig. 2(a). By using the C-SRC algorithm, we can obtain the sparse coefficient distribution diagram, shown in Fig. 2(b).

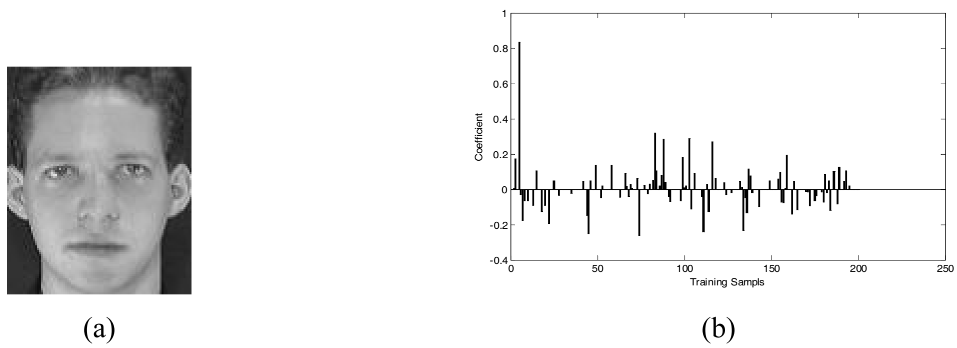
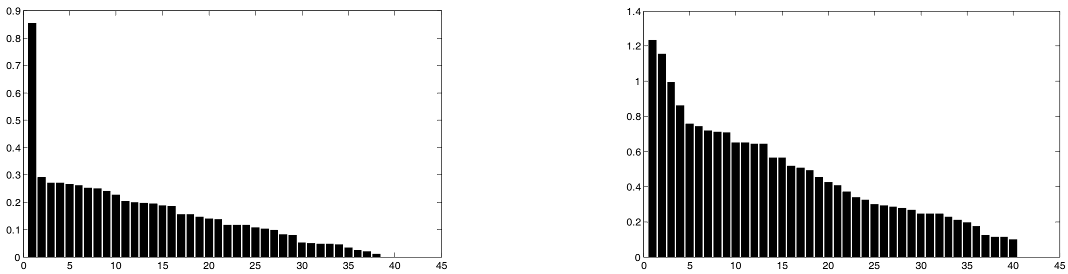


Fig. 2. Recognition of face images

In order to more intuitively observe the characteristics of the ℓ_2-norm values of various sparse coefficient vectors, the l-norm values are processed in descending order. Fig. 3(a) is a sorting diagram of the ℓ_2-norm value for the face image, and Fig. 3(b) is a sorting diagram of the ℓ_2-norm value for the non-face image. It can be seen from the distribution diagram, when the face image is input, the sparse coefficients are concentrated in the correlation class, and the class with the largest sparse coefficient vector ℓ_2-norm value is exactly the class of the input face image. At the same time, the input test samples are random, so this distribution feature has wide applicability.



(a) (b)

Fig. 3. Sorting of the ℓ_2-norm value

Aiming at the good characteristics of the ℓ_2-norm value of the sparse coefficient vector for non-face image discrimination, a method for discriminating non-face test images is proposed: the ℓ_2-norm value detection method. It basic rules are as follows: given the test image y, the ℓ_2-norm value of the sparse coefficient vector corresponding to each class is found in the training dictionary:

$$J_{y_i} = \left\| \delta_i(\hat{\alpha}_i) \right\|_2 . \tag{8}$$

Sort the obtained norm values from large to small:

$$J_a \geq J_b \geq J_c \cdots \geq J_z . \tag{9}$$

Where $a, b, c, \cdots, z$ represent the sorted category. The simulation results show that the ℓ_2-norm value of the first class after sorting is the largest, corresponding to the category of the test image, and the ℓ_2-norm values of other unrelated classes are small. According to this characteristic, Gao [56] used the ratio of the first two items after sorting as the basis for discrimination:

$$\varphi = \frac{J_a}{J_b} . \tag{10}$$

We select a threshold value τ, when $\varphi \geq \tau$, it is determined to be a non-face image, once a large threshold is be selected, the first category and other categories can be distinguished. However, this method has some limitations. First of all, it is difficult to give an accurate threshold value, for example, when the test image contains noise, the norm value for the second category will be larger, even the test image is a face image, but does not meet the threshold value, resulting in a discernment error. Secondly, from the characteristics of the ratio, regardless of whether or not the test image is a face image, the ratio φ must larger than 1, which cause the final discrimination is ambiguous, resulting in a decrease in the recognition rate.

n view of the defects of the above methods, we proposes an improved method of comparing the second and third terms:

$$M = \frac{J_a}{J_b + J_C} \geq \tau . \tag{11}$$

When the test image is a face image, the norm value of the first item after sorting J_a should be much larger than the value of the uncorrelated item. Considering the above characteristics, we sum the norm value of the second term J_b and the norm value of the third term J_c, and then divide with J_a, to get the new ratio M, and set the threshold value to $\tau = 1$. As show in Fig. 3, J_a should be much larger than other classes, compare J_a with $J_b + J_c$, the ratio M will be absolutely larger than 1, here we believe the values of J_b and J_c are generally less than half of J_a. On the contrary, when the test image is a non-face image, the values of the first three terms are very close, so the ratio M will be absolutely less than 1, that is, here we believe the values of J_b and J_c are generally larger than half of J_a.

Through the improved ℓ_2-norm value detection method, a clear threshold value is given, which increases the validity and accuracy of the recognition result. At the same time, after the algorithm determines that the test image is a non-face image, the recognition system will terminate the calculation, which saves the computational cost of the system and improves the robustness of the face recognition system.

4 Face Recognition Via Segmentation Processing Based on the Fusion of Residual

4.1 The Proposal of Segmentation Strategy

At present, most of the face recognition algorithms are based on the entire face image for calculation, such method is simple and easy to operate. However, when there are interference factors such as expressions, posstures, illumination changes, and even occlusions, the recognition effect will be unsatisfactory. At the same time, human eyebrows, eyes, nose, mouth and other organs are relatively different, which is the key to face recognition. In order to highlight the importance of local features, this paper proposed the improvement method of segmentation based on the C-SRC algorithm.

Select the second subject as the test image in the Extended Yale B face database, which contains the effects of uneven illumination. A 2×2 average blocking mode is adopted for the face image to obtain four sub-blocks with the same pixel, as shown in Fig. 4.

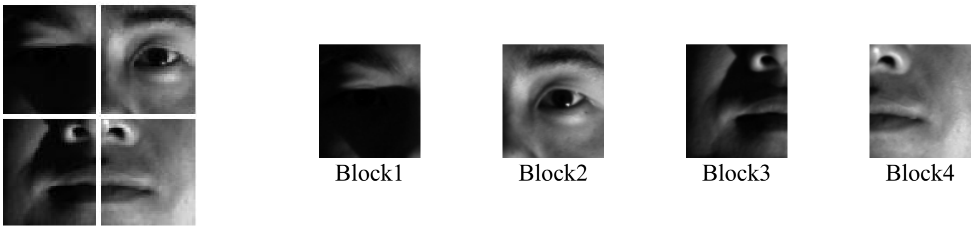


Fig. 4. Blocking the test image into four parts

In order to verify the robustness of the improved method, we performed simulation experiments in the Extend Yale B face database. This face database consists of 38 types of subjects, each of subject contains 64 images under different lighting conditions, randomly selected half images as the training samples and the rest for testing. We use the image which contains illumination interference, identified by the C-SCR algorithm, Fig. 5 shows the sparse coefficient and Fig. 6 shows the residual.

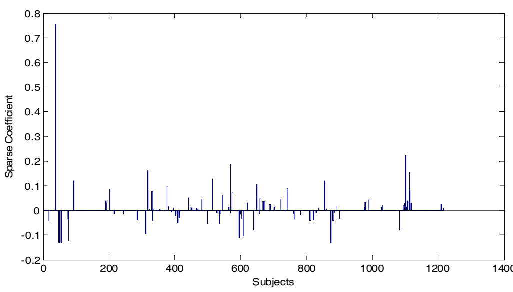


Fig. 5. Acquire the sparse coefficient via C-SCR algorithm

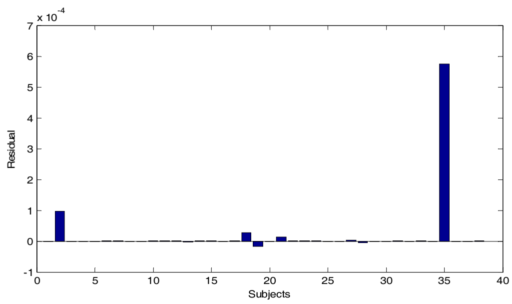
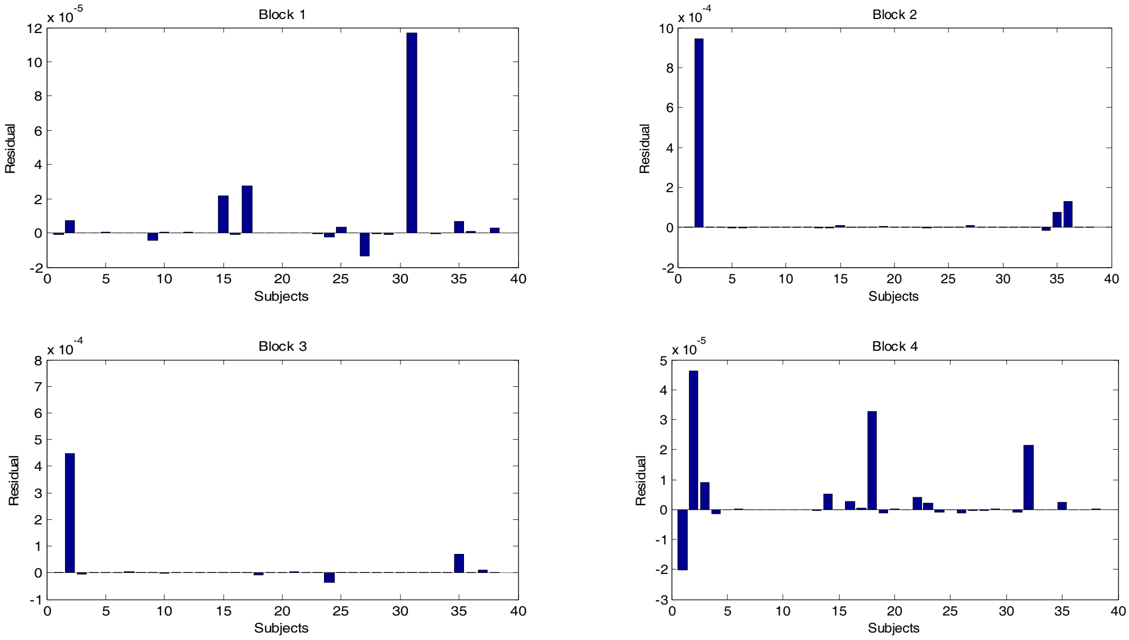


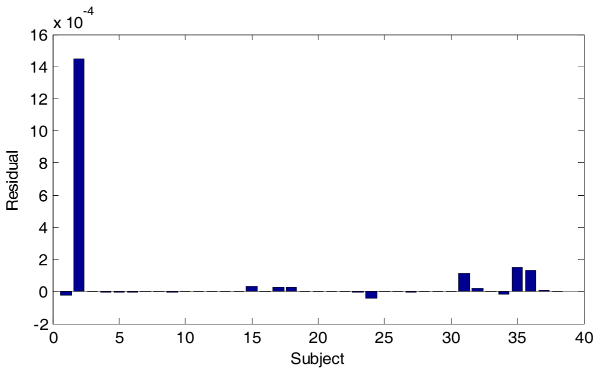
Fig. 6. Acquire the residual via C-SCR algorithm

It can be observed from Fig. 6 that the recognition result is wrong. In Fig. 5, in the interval [33-64] in which the sparse coefficient of the 2 person is present, the value with the largest coefficient appears, but there also include few values less than zero. In the interval where the sparse coefficient of the 35th class is located in [1089-1120], there are many coefficients that are not large but all positive. Since the sparse coefficient is the key to identification, the recognition result points to the wrong category 35.people. This is because the illumination in the test image is very uneven, and the important features of the face are in the shadow, and sufficient facial feature information cannot be extracted.

In order to optimize the problem of misidentification caused by uneven illumination, we propose an algorithm for segmentation processing. According to the pixel points of the image, the training library is formed by each sub-block, and each sub-block is calculated to obtain the sub-residual, then we sum the sub-residual values, the recognition result is output according to the sum of the residual sum. Through the segmentation processing method, even if some images have occlusion and uneven illumination, the part that is not affected by noise after block can still ensure the smooth progress of identification. The simulation experiment was carried out in accordance with the above method, which is show in Fig. 7.



(a)



(b)

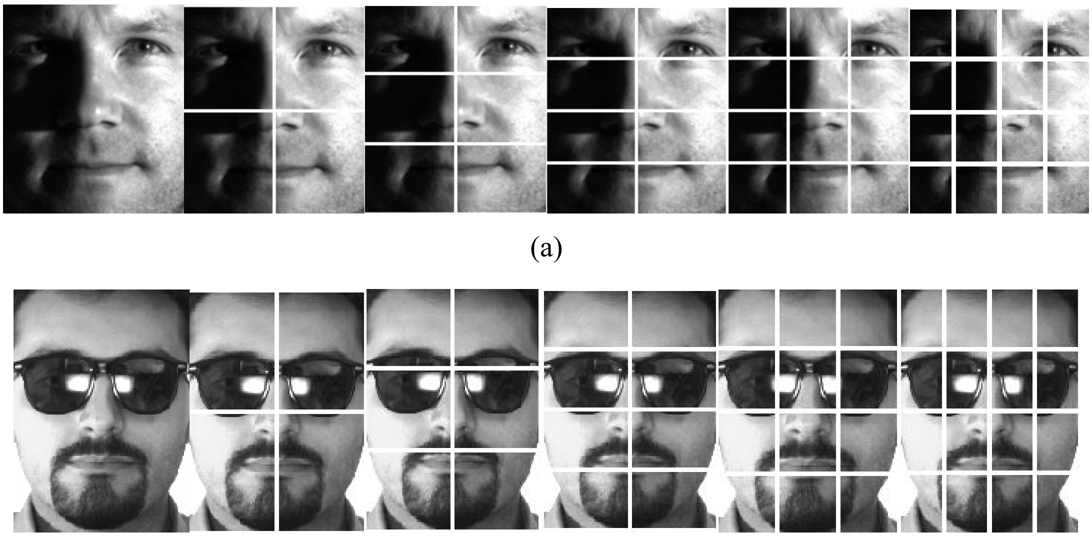
Fig. 7. Residual for each sub-block and the sum of the sub-residual

Comparing the residual graphs of each sub-block in Fig. 7(a), it can be seen that due to the uneven illumination, a large shadow is generated, and the feature of the left part is lost, which causes the recognition result of this part to be wrong. However, after the segmentation processing, other sub-blocks that are not affected by the illumination and have obvious features can be correctly identified, and the correct recognition result is obtained. Fig. 7(b) is the final recognition result obtained by adding the residuals of each sub-block, showing that after blocking, weakening the influence of the noise-containing block on the final recognition, and finally optimizing the classification effect.

4.2 Optimal Segmentation Method

The previous section uses 2×2 non-superimposed segmentation methods, but some sub-blocks still have a lot of adverse effects caused by noise, so we consider seeking a more appropriate segmentation method to optimize the recognition effect.

Try not to destroy the integrity of the feature blocks such as eyes, nose, mouth, etc., while separating the noise-containing areas. In this paper, the segmentation mode can be used such as: 2×2, 3×2, 4×2, 4×3, 4×4, show in Fig. 8, Fig. 8(a) shows the image segmentation effect in the Extend Yale B face database, and Fig. 8(b) shows the image segmentation effect in the AR face database.



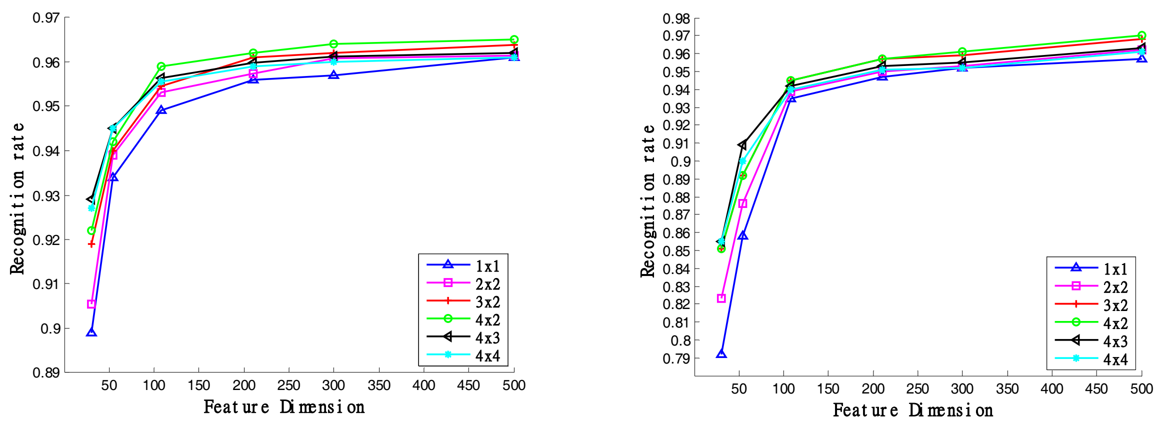
(a)

(b)

Fig. 8. Different face egmentation methods

The simulation experiment was carried out in the Extend Yale B face database, we randomly selected 32images as training samples and others as the test samples. The downsampling dimensionality reduction method was used t under different dimensions: 30, 54, 108, 210, 300, and 500, the recognition rate results are shown in Fig. 9(a).

In the AR face database, each subkject contains 26 pictures, including expressions, lighting and occlusion of a variety of changes. Randomly selected half as a training sample, but to ensure that each type of subject's training sample contains 7 unobstructed images, 3 images of wearing sunglasses and 3 images with scarves, a total of 13 images. The dimensions are selected as:30, 56, 108, 210, 300, and 500, and the recognition rate is as shown in Fig. 9(b).



(a) (b)

Fig. 9. Recognition rate via different segmentation method

In Fig. 9, the recognition rate gradually increases as the number of blocks increases, but when the number of blocks reaches 4×3 and 4×4, the recognition rate does not increase significantly or even decrease. The more blocks, the integrity of the feature is broken, such as splitting the eye into two parts. In addition, excessive blocks introduction of some unimportant features, such as cheeks, foreheads, etc., and these can also increases the computational cost, which has a negative impact on the face recognition system. It can be observed from the above figure that the result of identification is optimal when the number of blocks is 4×2.

4.3 Segmentation Classification Based on Fusion of Residual (SC-SRC)

The algorithm below summarizes the complete face recognition procedure.

Algorithm: Segmentation Classification based on fusion of residual (SC-SRC)

1. **Input**: a matrix of training samples $D=[D_1,D_2,\cdots,D_k]\in \mathrm{IR}^{m\times n}$ for k classes, a test sample $y\in \mathrm{R}^m$, (and an optional error $\varepsilon>0$).
2. Normalize the columns of D to have unit ℓ_2-norm.
3. Solve the ℓ_1- minimization problem:
   $\hat{\alpha}_1=\arg\min\|\alpha\|_1\ s.t.\|D\alpha=y\|_2\le\varepsilon$
4. Solve the ℓ_2-norm value of sparse coefficient vectors for each class:
   $J_{y_i}=\|\delta_i(\hat{\alpha}_i)\|_2$
5. Discrimination the validity of the test image:
   $J_a\ge J_b\ge J_c\cdots\ge J_z$, ($a,b,c,\cdots,z$ means sort from large to small)
   $M=\dfrac{J_a}{J_b+J_C}\ge\tau$
   When the discriminant is satisfied, the algorithm continues to run; when it is not satisfied, the algorithm terminates
6. Performs non-superimposed segmentation processing for face images:
   Training sample: $D_z=\left[D_{1z},D_{2z},...,D_{kz}\right]\in R^{m_p\times n}$, $z=1,2,\cdots,P$
   Test sample: $y_z=[y_1,y_2,\cdots y_z]\in R^{m_p}$, $z=1,2,\cdots,P$
7. Solve the ℓ_1- minimization problem:
   $\hat{\alpha}_{1z}=\arg\min\|\alpha_z\|_1\ s.t.\|D_z\alpha_z=y_z\|_2\le\varepsilon$
8. Compute the mean of coefficients:
   $\left(d_i\right)_z=(\sum_{i=1}^{n_i}\alpha_{ij}/n_{ij})_z,\left(a_{ij}\right)_z\in\left[\delta_i\left(x_i\right)\right]_z,\left[\delta_i\left(x_i\right)\right]_z$
   $=\left\{0,0,\cdots a_{i1},a_{i2}\cdots a_{in_i},0\cdots0,\right\}_z$
9. Compute the residuals of each sub-block:
   $\left(r_i\right)_z=\left(d_i\right)_z^3/\left[\left\|y_z-A_z\delta_i\left(\hat{x}_i\right)_z\right\|\right]_2, i=1,\cdots k$
10. Compute the sum of the residuals of each sub-block:
    $r_i(y)=\sum_{i=1}^{P}\left(r_i\right)_z$
11. **Output**: $\mathrm{identify}(y)=\arg\max r_i(y)$.

5 Simulation and Results

In order to verify the robustness of the proposed algorithm in this paper, simulation experiments were carried out in Matlab. Considering that the SC-SRC algorithm is mainly for the case of uneven illumination and facial occlusion, we chose the Extend Yale B face database and AR face database.

5.1 Recognition of Images with Uneven Illumination

In order to verify the effectiveness of the SC-SRC algorithm under the influence of non-uniform illumination noise in face images, this section selects Extend Yale B face data standard database for optimization verification experiments. In this face database, there are 38 subjects under different illumination conditions, each of the subject contains 64 images under different illumination conditions. Thirty-two samples were randomly selected as train samples and the rest as test samples, with dimensions of 30, 54, 108, 210 and 500, the recognition rate is shown in Fig. 10. The results of experiments show that the recognition rates of SRC, C-SRC and SC-SRC are respectively 95.5%, 96.1% and 96.5% under 500D feature space. The algorithm proposed in this paper has better recognition robustness in the non-ideal case with illumination changes.

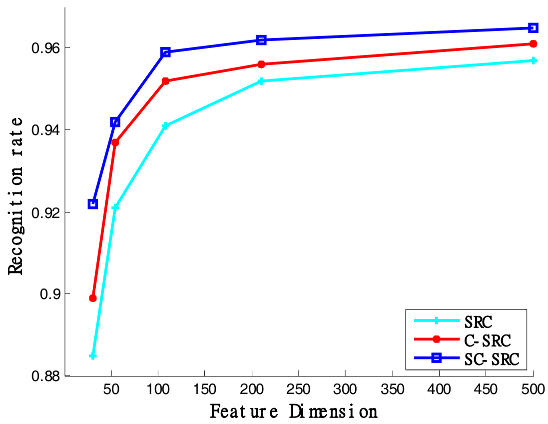


Fig. 10. Recognition rate in Extend Yale B database

5.2 Recognition of Images with Facial Occlusion

In order to verify the superiority of the SC-SRC algorithm in the case of interference with facial occlusion, AR face database is selected for optimization verification experiment. The images containing different noises are evenly distributed, and the composition of the training library is: seven pictures are normal faces, three pictures contain sunglasses, three pictures contain scarves, the total number of each subject for training thirteen. The dimension used are 30, 54, 108, 210 and 500, and the recognition rate is shown in Fig. 11. The results of experiments show that under 500D feature space, the recognition rates of SRC, C-SRC and SC-SRC are respectively 94.9%, 95.7% and 97.0%. It is proved that the SC-SRC proposed in this paper has better identification robustness under the condition of facial occlusion.

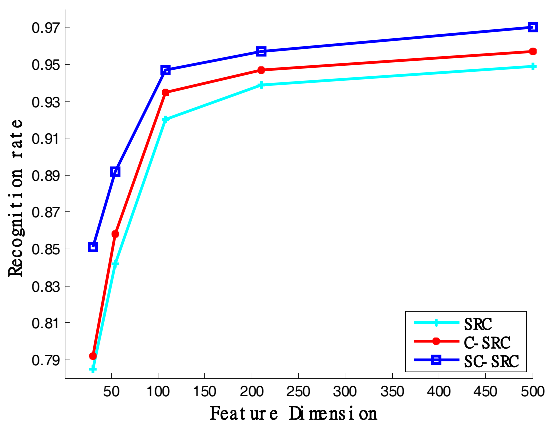


Fig. 11. Recognition rate in AR database

5 Conclusions

In this paper, we proposed a face recognition method via segmentation based on fusion of residual, according to the improved ℓ_2-norm value detection method, to discriminate the validity of the test image, and then the face image is reasonably segmented, and each sub-block is separately identified, after that we calculate the sum of residuals which is taken as the basis of final recognition. This method can effectively reduce the influence of the classification of the entire face when the noise area is contained in the face. Finally, the improved recognition algorithm is simulated in the standard face database, the result shows that the SC-SRC method has better recognition robustness under non-ideal conditions such as uneven illumination and facial occlusion.

An intriguing question for future work is whether this framework can be useful for object detection, in addition to recognition. From a practical standpoint, in the simulation experiment, the face images we selected are all from the international standard face database, the processed images exclude many factors such as unaligned images and uncontrollable images. However, in the practical application of the face recognition system, it often faces a variable and uncontrollable environment, which causes the recognition effect to decline sharply, the solution of such problems still needs further research. In addition, this paper only carried out simulation experiments under the premise of including illumination changes and occlusions, and did not explore the situation including posture and expression changes. Doing so in a principled manner remains an important direction for future work.

References

[1] M.A. Turk, A.P. Pentland, Face recognition using eigenfaces, in: Proc. IEEE Conf. Comput. Vis. Pattern Recognit., 1991.

[2] J. Neves, F. Narducci, S. Barra, H. Proença, Biometric recognition in surveillance scenarios: a survey, Artificial Intelligence Review 3(1)(2016) 1-27.

[3] H.K. Hoomod, A. abd ali, Fuzzy-cellular neural network for face recognition HCI authentication, Journal of Physics: Conference Series, 1003(1)(2018), 012033. doi :10.1088/1742-6596/1003/1/012033.

[4] Z. Zainuddin, A.S. Laswi, Implementation of the LDA algorithm for online validation based on face recognition, Journal of Physics: Conference Series 801(1)(2017), 012047. doi:10.1088/1742-6596/801/1/012047.

[5] F. Kuang, S. Zhang, Z. Jin, W. Xu, A novel SVM by combining kernel principal component analysis and improved chaotic particle swarm optimization for intrusion detection, Soft Computing 19(5)(2015) 1187-1199.

[6] T. Serre, Learning a dictionary of shape-components in visual cortex: comparison with neurons, human and machine, [dissertation] Boston: Massachusetts Institute of Technology, 2006.

[7] B. Olshausen, D. Field, Emergence of simple-cell receptive field properties by learning a sparse code for natural image, Nature 381(6583)(1996) 607-609.

[8] M.H. Siddiqi, R. Ali, A.M. Khan, Y.-T. Park, S. Lee, Human facial expression recognition using stepwise linear discriminant analysis and hidden conditional random fields, IEEE Transactions on Image Processing 24(4)(2015) 1386-1398.

[9] Z. Fan, Y. Xu, M. Ni, X. Fang, D. Zhang, Individualized learning for improving kernel Fisher discriminant analysis, Pattern Recognition 58(2016) 100-109.

[10] B. Song, S. Tan, H. Shi, Process monitoring via enhanced neighborhood preserving embedding, Control Engineering Practice 50(2016) 48-56.

[11] O. M. Parkhi, A. Vedaldi, A. Zisserman, Deep face recognition, in: Proc. British Machine Vision Conferenc, 2015.

[12] J. Yang, M.H. Yang, Top-down visual saliency via joint CRF and dictionary learning, IEEE Trans Pattern Anal Mach Intell. 39(3)(2017) 576-588.

[13] J. Wright, Y. Ma, J. Mairal, G. Sapiro, T.S. Huang, S. Yan, Sparse representation for computer vision and pattern recognition, Proceeding of the IEEE 98(6)(2010) 1031-1044.

[14] H. Li, C.Y. Suen, Robust face recognition based on dynamic rank representation, Pattern Recognition, 60(2016) 13-24.

[15] K. Hui, C. Li, L. Zhang, Sparse neighbor representation for classification, Pattern Recognition Letters 33(5)(2012) 661-669.

[16] J. Wright, A. Yang, A. Ganesh, S. Sastry, L. Yu, Robust face recognition via sparse representation, IEEE Transaction on Pattern Analysis and Machine Intelligence 31(2)(2009) 210-227.

[17] Y. Du, F. Guo, The analysis of face recognition based on BP artificial neural network, in: Proc. 7th International Conference on Mechatronics, Control and Materials, 2016.

[18] M. Yang, L. Zhang, S.C.K. Shiu, D. Zhang, Gabor feature based robust representation and classification for face recognition with Gabor occlusion dictionary, Pattern Recognition 46(7)(2013) 1865-1878.

[19] J. Zhang, Y.-Z. Zhan, Q.-R. Mao, Facial expression recognition based on Gabor wavelet and sparse representation, Computer Engineering 38(6)(2012) 207-209.

[20] W. Deng, J. Hu, J. Guo, In defense of sparsity based face recognition, in: Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2013.

[21] K. Huang, S. Li, X. Kang, L. Fang, Spectral-spatial hyperspectral image classification based on KNN, Sensing & Imaging, 17(1)(2016) 1-13.

[22] K. Feng, Z. Jiang, W. He, B. Ma, A recognition and novelty detection approach based on curvelet transform, nonlinear PCA and SVM with application to indicator diagram diagnosis, Expert Systems with Application: An International Journal 38(10)(2011) 12712-12729.

[23] J.J. Wang, J.C. Yang, K. Yu, F.J. Lv, T. Huang, Y.H. Gong, Locality-constrained linear coding for image classification, in: Proc. IEEE Conference on Computer Vision and Pattern Recognition, 2010.

[24] Z. Fan, Y. Xu, D. Zhang, Local linear discriminant analysis framework using sample neighbors, IEEE Trans. Neural Netw. 22(7)(2011) 1119-1132.

[25] Y. Hao, Research on robust face recognition based on compressed sensing, [dissetation] Beijing: Beijing University of Civil Engineering and Architecture, 2016.

[26] Z. Gao, C. Xiong, B. Da, Sparse representation based on residual weighting for face recognition, Journal of South-Central University for Nationalities (Natural Sciences) 31(3)(2012) 72-76.