Journal of Computers Vol. 29 No. 4, 2018, pp. 241-249
doi:10.3966/199115992018082904019

Robust Face Recognition Based on the Fusion of Sparse Coefficient and Residual

Liquan Zhang, Jie Gao*, Yujuan Hao

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

Received 1 December 2017; Revised 30 April 2018; Accepted 31 May 2018

Abstract. In the face recognition community, the research of sparse representation has seen a recent surge of interest. Even though the images are with varying expression and illumination, as well as occlusion, most of the algorithms still have a good recognition effect. However, when the test and training images contain both the changes of illumination and expression, the traditional sparse representation algorithm often performs the wrong face recognition. In sparse representation, the ℓ_1-norm was used to define the fidelity of sparse coding. In fact, the fidelity terms (sparse coefficients) can represent the testing samples as a sparse linear combination of the training dictionary, and hence they have a very important influence on the final classification. In this paper, we propose a simple and effective face recognition algorithm, in which the sparse coefficients and original residuals are fused effectively. The useful information of sparse coefficients can be fully reflected in the residuals. Hence the new residual values, that are obtained, can improve the fidelity of residuals. We exploit the fusion nature of sparse coefficients to redefine the computing method of residuals, and then perform classification. We conduct several experiments on publicly available database to verify the efficacy of the proposed approach and corroborate our claims.

Keywords: face recognition, facial expression, fusion, illumination, sparse coefficients and residuals, sparse representation

1 Introduction

Face recognition technology has become a hot topic in recent years. Since principal component analysis (PCA) [1] was applied successfully, much of this excitement centers around face recognition is focused on subspace analysis method due to its good properties [2-5]. Linear subspace analysis methods, such as linear discriminate analysis (LDA) [6], provide a solid foundation for various non-linear methods. The goal of non-linear subspace analysis methods was to extract the local structure of data per se. In this community, there have been several popular works such as local preserving projection (LPP) [7] and neighborhood preserving embedding (NPE) [8] algorithm. The two algorithms can not only solve the weakness, that is difficult to maintain the nonlinear flow of the original data, of traditional linear methods (e.g., PCA), but also overcome the shortcoming of the nonlinear methods which is difficult to acquire the low dimensional projection of the new sample points. Furthermore, Fan et al. [9] proposed an improved LDA framework (LLDA), which can effectively capture the local structure of the samples. Based on kernel principal component analysis (KPCA) and LDA, Yang et al. [10] presented a complete kernel Fisher discriminate analysis algorithm (CKFD), which can utilize the regular and non-regular discriminate information effectively. Similarly, Cevikalp et al. [11] proposed a kernel discriminative common vector approach. Recently, investigators have revealed that the face space is more likely to be presented in the low dimensional nonlinear manifold subspace, such as local tangent space alignment algorithm (LTSA) [12]). On the basis of LTSA, Zhang et al. [13] proposed an adaptive manifold learning

* Corresponding Author

approach to acquire the optimal neighborhood value. Furthermore, Chen et al. [14] presented a local discriminate embedding (LDE) method for pattern classification. At present, in the field of face recognition, particular interest has been put into sparse representation because of Wright et al.' pioneering study [15], in which a novel pattern classification method was proposed underlying sparse representation (SRC). SRC represents a test sample as a linear combination of all training samples, and this kind of representation is naturally sparse. According to the obtained sparse representation coefficients, the test sample will be divided into the class that the reconstruction error is minimal. Generally, the sparse representation can be described as:

$$\hat{x}_1 = \arg\min \|x\|_1 \text{ subject to } \|Ax - y\|_2 \le \varepsilon . \tag{1}$$

where y is any test sample vector, the matrix A is the dictionary of training samples, x is a sparse coefficient vector, and $\varepsilon > 0$ is a small constant.

Although the sparse representation in Eq. (1) has been used in a lot of literatures widely, there still exist two main issues. The first problem is whether the ℓ_1-norm constraint coefficient x can sufficiently characterize the sparse of signal. The second problem is whether the ℓ_2-norm term can effectively characterize the fidelity of signal effectively, even though real data y is noisy. An enormous volume of literature has been devoted to investigate the sparse constraint [16-17]. In this community, Liu et al. [18] proposed a non-negative curds and when (NNCW) method to seek sparse and non-negative representation. NNCW pointed out that the sparser the coefficients are, the easier the test sample is assigned to a correct class label. In addition, during the feature quantization process of sparse coding, some similar local features may be quantized into different visual words of the codebook. Aiming at this problem, Gao et al. [19] offered a Laplacian operator of sparse representation, which exploits the dependent relation among the local features.

In fact, the low-dimensional features of an object image are the most relevant or informative for classification generally. This dimensionality reduction issue underlying sparse representation has also been strongly supported by various studies of face recognition [20-22]. The main goal of these works was to project the high-dimensional test image into lower dimensional feature spaces. Zhang et al. [23] proposed a graph-based optimization for dimension reduction with sparsity constraints (GODRSC), which unifies the graph and the projection matrix into the same framework, and obtains the optimal graph by continuous iterations. However, most of studies do not explicitly treat the manifold structure of the data. In order to solve this problem, Qiao et al. [24] developed an unsupervised dimensionality reduction method called sparsity preserving projections (SPP). Unlike LPP and NPE, SPP aims to preserve the sparse reconstructive relationship of the data, and hence can be more easily realized in practice.

The original goal of mentioned-above works was to optimize the sparse representation from different aspects, and hence algorithm performance was measured in terms of sparsity of representation. However, the fidelity term has greater contributes to final classification because it will ensure that the given signal y can be represented by the dictionary A. Although the ℓ_1-norm was used to define the coding fidelity of sparse coefficients [25-26], it actually had limited the coding residuals to follow Gaussian or Laplace distribution. Nevertheless, the assumption may not work well in practice, especially with varying expression, occlusion and corruption. Generally, the residuals are used to determine the identity of test sample, and hence it is very important to improve the fidelity of the residuals. In this paper, we propose a new face recognition approach based upon the fusion of sparse coefficients and residual (C-SRC), which works on improving the classification ability of the residual. Through computing the mean of ℓ_1-minimization recovered sparse coefficients for each class, we can reflect the relativity, that is the characteristics among all the training samples and testing sample sparse coefficients, correlation among intra-class and inter-class sparse coefficients, but also highlight the characteristic of object class associated with the test image. We conduct extensive experiments on publicly available database to verify the efficacy of the proposed algorithm and support the above claims.

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

(1) Based on the AR face database and Extended Yale B face database, we use the data on the sparse coefficient characteristics to make analyzed. First, we propose to verify the correlation between the sparse coefficient and residual during the classification. By fusion the sparse coefficient and residual, we

can effectively weaken the extreme value which existing between the sparse coefficients, and also reduce the influence of extreme value of residual classification. This method can strengthen the sparse coefficient of overall indicators affect classification strategy.

(2) Based on analyzing the correlation between the sparse coefficient and residual, proposed the method that fusion the sparse coefficient and residual, a new method of face recognition called C-SRC. The purpose of this method is introduced in the SRC residuals average coefficients in class in order to reach optimization effect. This article has redefined the classification strategy, which apply the maximum residual instead of using residual minimum value to judge the SRC category. According to the results of experiment, the results show that the method under changing expression and illumination conditions have strong recognition robustness.

2 Robust Face Recognition Based on Sparse Representation

The essence of sparse representation is to decompose signals under the constraint of sparse regularization. In SRC, each $w \times h$ gray scale image can be regarded as a vector $v \in R^m (m = w \times h)$. Given in training samples of the i-th object class, a matrix $A_i = \left[v_{i,1}, v_{i,2}, ..., v_{i.n_i} \right] \in R^{m \times n_i}$ will be constructed, and any test sample $y \in R^m$ from the same class will approximately lie in the linear span of the training samples:

$$y = a_{i,1} v_{i,1} + a_{i,2} v_{i,2} + \cdots + a_{i,n_i} v_{i,n_i} . \quad \textbf{(2)}$$

where the scalars $a_{i,j \in R}$, $j = 1, 2, ..., n_i$.

Since the identity i of the test sample is initially unknown, a matrix $A = \left[A_1, A_2, ..., A_k \right] = \left[v_{1,1}, v_{1,2}, ..., v_{k,n_k} \right]$, which is a concatenation of n training samples of all k object class, is defined by using the entire training set. In fact, given sample data are usually noisy, and hence the linear combination of y can be represented as:

$$y = Ax_0 + z \in \mathrm{R}^m , \quad \textbf{(3)}$$

where $x_0 = \left[0, ..., 0, a_{i,1}, a_{i,2}, ..., a_{i,n_i}, 0, ..., 0 \right]^T \in R^n$ denotes a coefficient vector, and non-zero terms are corresponding to the i-th class, and $z \in R^m$ is a noise term with bounded energy $\|z\|_2 \in \varepsilon$, ε is a small enough value but not zero.

According to the sparse representation and compressed sensing, the sparse solution x can be obtained by solving the following ℓ_1-minimization problem:

$$\widehat{x}_1 = \arg\min \|x\|_1 \text{ subject to } \|Ax - y\|_2 \le \varepsilon . \quad \textbf{(4)}$$

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

$$\min_i r_i(y) = \|y - A\delta_i(\widehat{x}_1)\|_2 , \quad \textbf{(5)}$$

where $\delta_i(\widehat{x}_1) \in \mathrm{R}^n$ is the vector whose only nonzero terms are the terms in x that are associated with the i-th class.

3 Sparse Representation via Fusion of Coefficients and Residual

3.1 Fusion of Sparse Coefficients and Residual

As we all know, the sparse representation coefficients themselves reflect the relativity between the testing input and the object class under the influence of all training samples. It includes the relationship not only among the features of the testing sample and the features of the entire training samples, but also among the features of the testing sample and all the classes. According to traditional SRC algorithm, most of sparse coefficients associated with the object class are positive terms generally, include one maximum value, the other classes have smaller positive and negative sparse coefficients.

However, this is not the case. It has been observed that the sparse coefficients of SRC are not always sparse under varying lighting and posture, as well as occlusion and corrupt. Some larger positive sparse coefficients will also appear in non-object classes, and there probably exist multi-group negative sparse coefficients with larger absolute values. That is, if we minimize the residual by using only these coefficients, it may be possible to result in an error classification. Obviously, the value of residual $r_i(y)$ varies only with the vector $\delta_i(\hat{x}_1)$, that is, the coefficients vector x determines the residual uniquely. Generally, the values of sparse coefficients are smaller, and hence the obtained residual does not efficiently represent the relative information of negative coefficients based on the ℓ_2-norm. As a result, for each class, the difference of residual is not distinctive.

To resolve this difficulty, we firstly calculate mean value d_i of sparse coefficients for each object class, and then define a new residual by dividing the SRC residual by the cube of mean value d_i. Finally, the test sample can be assigned to the object class that maximizes the new residual. The intention of this work is to enhance the positive relativity between the testing sample and its corresponding class, and also weaken the negative correlation between them at the same time. In addition, the larger positive coefficient of any non-corresponding class can be weaken by these negative coefficients of the same class, and the cube can ensure to maintain the positive and negative properties of sparse terms. Of course, the higher the odd power is, the better the positive relativity will be.

We choose the Extended Yale B database to verify the robustness of our new residual to varying illumination. In Fig. 1 show: (a) Left: the testing sample from subject 24. Middle: reconstructed image by SRC. Right: reconstructed image by our method. (b) Sparse coefficients calculated by SRC. (c) 38 residuals by SRC,the smallest residual belongs to subject 10. (d) 38 residuals by our method,the largest residual is associated with subject 24.

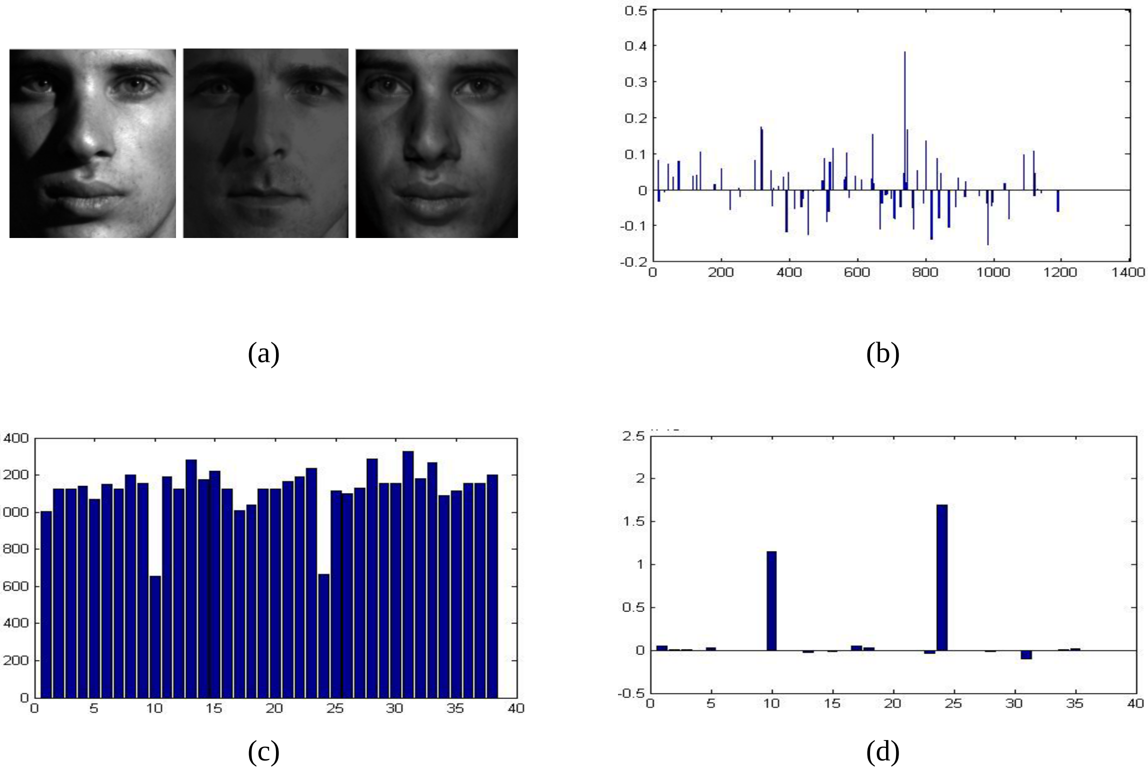


Fig. 1. Comparison between two classification algorithms with varying illumination

The database includes the frontal views of 38 subjects, and each subject has 64 images under different illumination conditions. We randomly select half of all the images in the database as the training samples and the rest for testing. Fig. 1(b) shows the coefficients recovered by the SRC algorithm for the testing image from subject 24. From the distribution of sparse coefficients in Fig. 1(b), the subject 10 contains some less negative coefficients and a larger positive term, whereas the largest coefficient is corresponding to the subject 24. Fig. 1(c) shows the corresponding residuals with respect to the 38 subjects. According to SRC algorithm, the smallest residual and reconstructed image (the middle image in Fig. 1(a) should be associated with subject) 10. To illustrate the contrast between SRC and our new residual, Fig. 1(d) plots the residuals of the same test image given by our new method. Obviously, the largest residual in Fig. 1(d) should be assigned to the correct subject 24, and the disturbance between two larger residuals is much less than those given by SRC algorithm (Fig. 1(c)). Obviously, our method has stronger robustness in a variety of illumination.

To illustrate the validation of our method to varying express and occlusion, we choose to use AR database. In Fig. 2.: (a) Left: the test sample from subject 5. Middle: reconstructed image by SRC. Right: reconstructed image by our method. (b) Sparse coefficients calculated by SRC. (c) 50 residuals by SRC, the smallest residual belongs to subject 19. (d) 50 residuals by our method, the ratio between the two largest residuals 4:1.The database preserves the frontal views of 50 subjects, and each subject has 26 images with varying express and occlusion. We randomly select the 14 images as the training set and the rest for testing. Fig. 2(b) shows the sparse coefficients recovered by the SRC algorithm for the testing image from subject 5. Because the change of facial express is larger for the testing image, the extracted features from eyes and mouth are not too obvious. As a result, this solution favored by ℓ_1-norm is not sufficiently sparse. Fig. 2(c) plots the residuals of a test image of subject 5 by the SRC algorithm. From this figure, the smallest residual should be corresponding to subject 19. Obviously, this SRC algorithm executes an error classification. Fig. 2(d) plots the corresponding residuals with respect to the 50 subjects by our new algorithm. Compared to those in Fig. 2(c), the largest residual in Fig. 2(d) is too evident, and the testing image should be assigned to subject 5. It follows that our method can enhances the positive relativity between the testing sample and its corresponding class, and weaken the negative correlation between them at the same time.

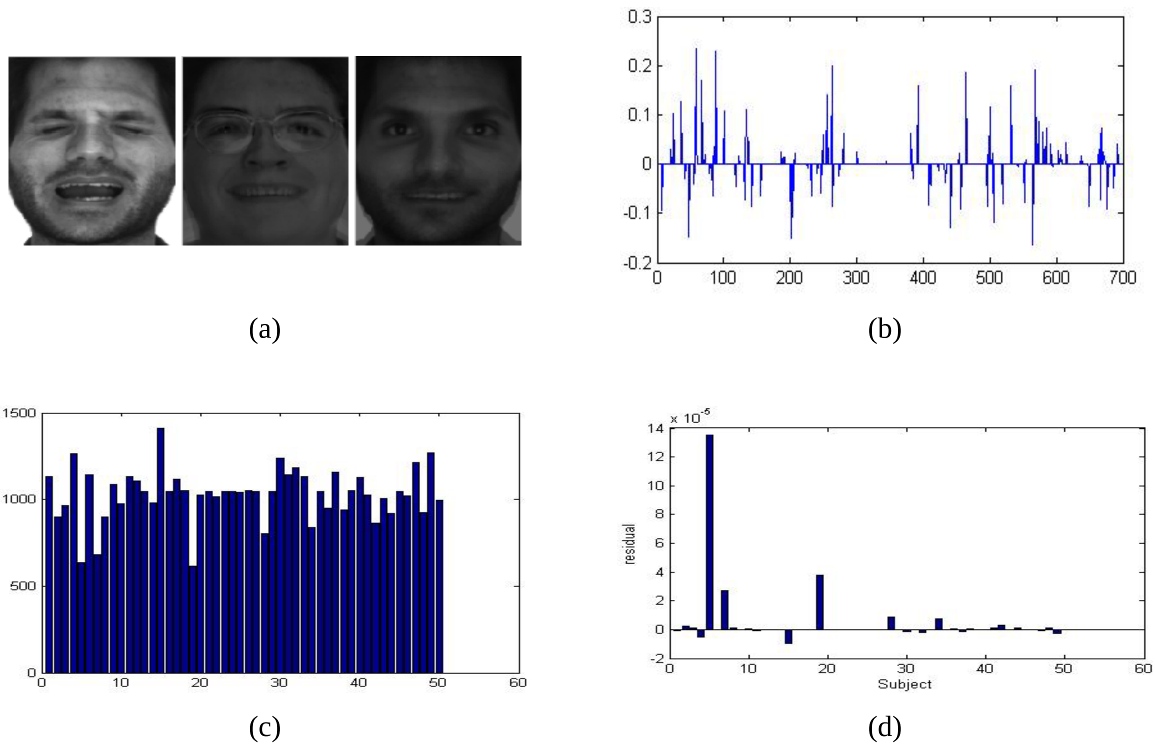


Fig. 2. Comparison between two classification algorithms with varying expression

3.2 Classification Based on Fusion of Coefficients and Residual (C_SRC)

The algorithm below summarizes the complete face recognition procedure.

Algorithm: *Classification based on fusion of coefficients and residual (C_SRC)*

1. Input: a matrix of training samples $A = [A_1, A_2, \cdots, A_k] \in \mathrm{IR}^{m\times n}$ for k classes, a test sample $y \in \mathrm{R}^m$, (and an optional error tolerance $\varepsilon > 0$).

2. Normalize the columns of A to have unit ℓ_2-norm.

3. Solve the ℓ_1- minimization problem:

$\hat{x}_1 = x \arg\min \|x\|_1$ subject to $Ax = y$

(Or alternatively, solve: $\hat{x}_1 = x \arg\min \|x\|_1$ subject to $\|Ax - y\|_2 \le \varepsilon$)

4. Compute the mean of coefficients

$$d_i = \sum_{j=1}^{n_i} \frac{a_{i,j}}{n_i}, \quad a_{i,j} \in \delta_i(x_i), \quad \delta_i(x_i) = [0, 0, \cdots, a_{i,1,} a_{i,2,} \cdots, a_{i,n_i,} 0, \cdots 0]$$

5. Compute the new residuals:

$$r_i(y) = \frac{\|y - A\delta_i(\hat{x}_i)\|_2}{d_i^3}.$$

6. Output: $\mathrm{identify}(y) = \arg\max r_i(y)$.

Generally, the sparse coefficients are utilized to verify the validation of testing image, whereas the residual is regarded as the classification criterion. Our definition of residual for classification differs significantly from that given by SRC algorithm. The motivation of fusion mainly considers the correlation among positive and negative coefficients. Hence the proposed algorithm can not only exclude some interfering images with larger positive values effectively, but also improve the recognition rate. Instead of using the minimum residual, we assign the maximum residual to the object class.

4 Simulation and Results

We conduct experiments on the Extended Yale B and AR databases for face recognition, which serve both to illustrate the efficacy of the proposed algorithm and to verify our claims of the previous sections. We also compare our algorithm with two classical algorithms, namely, SVM and SRC.

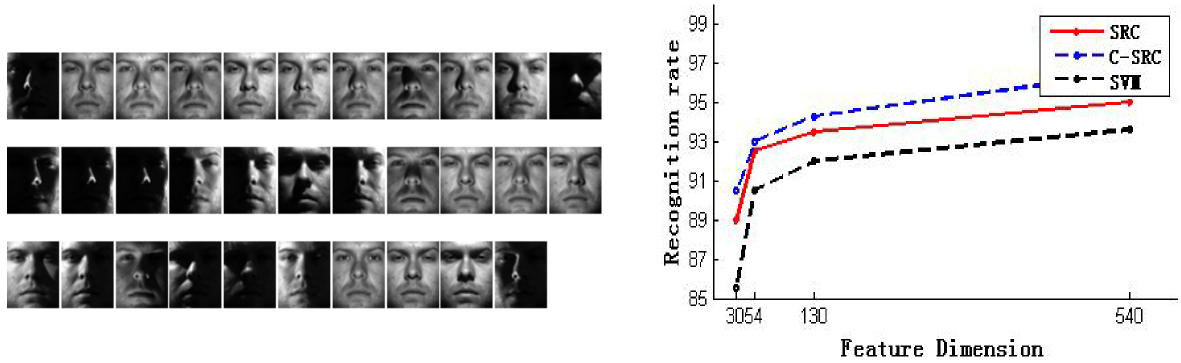
4.1 Experiment on Extended Yale B Database

The Extended Yale B database includes the frontal-face images of 38 individuals, each individual has 64 images with varying illumination, and the original size of each image is 192×168. We randomly select half of all images as the training samples and the rest for testing. Like SRC algorithm, we also calculate the recognition rates with the feature space dimensions 30, 56, 120, and 504, which correspond to downsampling ratios of 1/32, 1/24, 1/16, and 1/8 respectively. Fig. 3 shows the recognition rates of three different classifiers under four feature space dimensions.

In Fig. 3(a) Show the test sample of one person. (b) (c) Show the recoginition rate of three classification methods.The experimental results show that the maximum recognition rates for SRC, SVM, and C-SRC are 95%, 93.6%, and 96.5% respectively under 504D feature spaces. The performance of the proposed algorithm is better than SRC and SVM.

4.2 Experiment on AR Database

The AR database consists of over 4000 frontal images from 126 volunteers. These images include more angle variations, illumination change, expressions and occlusion, and each individual consists of 26 pictures, In the experiment, we choose 50 subjects with 165×120 size. For each subject, we select 14 images as the training sets and the rest for testing. Like the classical SRC algorithm, we also compute the recognition rates with the feature space dimensions 30, 56, 120, and 504, which correspond to downsampling ratios of 1/24, 1/18, 1/12, and 1/6 respectively. Fig. 4 shows the recognition rates of three different classifiers with the change of dimension.

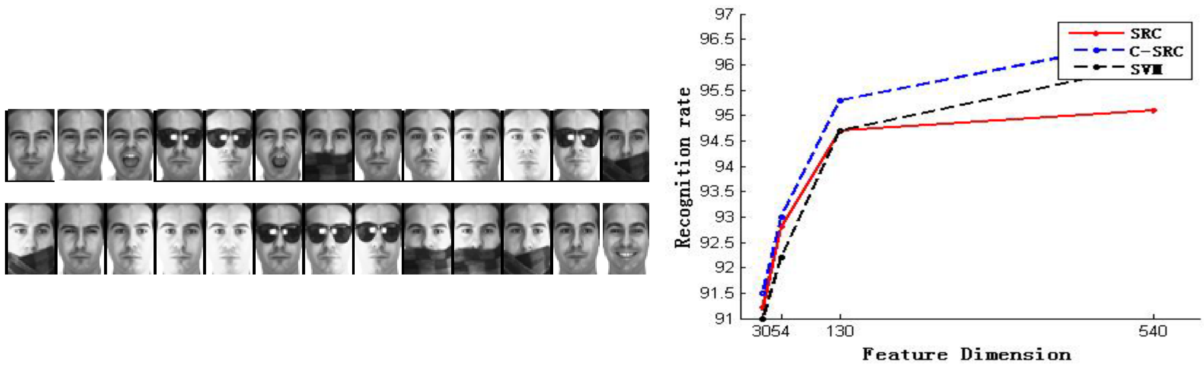


(a) (b)

Classifiers	Recognition rates under different dimension			
(Dimension)	30	56	120	504
SVM	0.875	0.905	0.92	0.936
SRC	0.89	0.925	0.935	0.95
C-SRC	0.905	0.93	0.943	0.965

(c)

Fig. 3. Recognition rates on Extended Yale B database



(a) (b)

Classifiers	Recognition rates under different dimension			
(Dimension)	30	56	120	504
SVM	0.91	0.922	0.947	0.96
SRC	0.912	0.928	0.947	0.951
C-SRC	0.913	0.93	0.953	0.963

(c)

Fig. 4. Recognition rates on AR database

In Fig. 4, (a) show the test sample of one person. Fig. 4(b) and Fig. 4(c) show the recoginition rate of three classification methods.The experimental results illustrate that the maximum recognition rates for SRC, SVM, and C-SRC are 96%, 95.1%, and 96.3% respectively under 504D feature spaces. The performance of the proposed algorithm is better than SRC and SVM. However, if the sample dimension is too low, the recognition effect of our algorithm is not as good as SRC.

5 Conclusions

Based on the simulation results on two available databases, we draw some conclusions. Through fusing the coefficients and residual, the strategy enhances the positive relativity between the testing sample and object class, and weaken the negative correlation between them at the same time. Furthermore, the larger positive coefficient of any non-object class can be weaken by these negative coefficients of the same class, and the cube can ensure to maintain the positive and negative properties of sparse terms. In addition, the algorithm can overcome the error classification of SRC with varying illuminations and expresses effectively, and improve the robustness of sparse classification.

Face recognition is a challenging technology, its purpose is to establish a system which can eliminate the disturbance of uncontrollable conditions and accurate automatic identification, the system can be applied in the field of national security, financial education and life. However, the identification process will be affected by uncontrollable conditions. In the face of the changeable unknown environment, the technology still has some problems to be overcome: The image samples used in this experiment are all from the common face database, and are the aligned images. For unaligned images, system recognition will be limited. When there is a large degree of rotation, the classification effect is worse, which should be further studied and improved in the future identification process. The face sparse representation algorithm is based on a large sample database, which is also verified in different training samples in this paper. When the training sample is small, the treatment result is not satisfactory. How to solve this problem is also a difficult problem. When the image contains multiple noises, the resulting sparse coefficient does not have good sparsity. For example, the image is covered by a large area and the classification effect is not good. Therefore, an efficient calculation method should be found to optimize the image. When the detection of face to face in the libralry with special similar characteristics, for example the twins, the classification of the effect is not obvious, dealing with specific details to feature extraction, should be improved.

References

[1] M.A. Turk, A.P. Pentland, Face recognition using eigenfaces, in: Proc. 1991 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 1991.

[2] J. L. Tu, X. M. Liu, P. Tu, On optimizing subspaces for face recognition, in: Proc. the 12th International Conference on Computer Vision (ICCV’ 09), 2009.

[3] J.W. Lu, Y.P. Tan, Cost-sensitive subspace learning for face recognition, in: Proc. the 2010 Computer Vision and Pattern Recognition (CVPR’ 10), 2010.

[4] A. Rao, S. Noushath, Subspace methods for face recognition, Computer Science Review 4(1)(2010) 1-17.

[5] Y. Wang, Y. Wu, Face recognition using intrinsicfaces, Pattern Recognition 43(10)(2010) 3580-3590.

[6] P.N. Belhumeur, J.P. Hespanha, D.J. Kriegman, Eigenfaces vs. fisherfaces: recognition using class specific linear projection, IEEE Transactions on Pattern Analysis and Machine Intelligence 19(7)(1997) 711-720.

[7] X. He, P. Niyogi, Locality preserving projections, in: Proc. Advances in Neural Information Processing Systems (NIPS), 2003

[8] Y. Fu, T.S. Huang, Graph embedded analysis for head pose estimation, in: Proc. IEEE International Conference on Automatic Face and Gesture Recognition, 2006.

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

[10] J. Yang, A.F. Frangi, J.-Y. Yang, D. Zhang, Z. Jin, KPCA plus LDA: a complete kernel fisher discriminant framework for feature extraction and recognition, IEEE Transactions on Pattern Analysis and Machine Intelligence 27(2)(2005) 230-244.

[11] H. Cevikalp, M. Neamtu, M. Wilkes, Discriminative common vector method with kernels, IEEE Transactions on Neural Networks 17(6)(2006) 1550-1565.

[12] Z. Zhang, H. Zha, Principal manifolds and nonlinear dimensionality reduction via tangent space alignment, SIAM Journal on Scientific Computing 26(1)(2004) 313-338.

[13] Z.Y. Zhang, J. Wang, H.Y. Zha, Adaptive manifold learning, IEEE Transactions on Pattern Analysis and Machine Intelligence 34(2)(2012) 253-265.

[14] H.-T. Chen, H.-W. Chang, T.-L. Liu, Local discriminant embedding and its variants, in: Proc. IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05), 2005.

[15] J. Wright, A. Yang, A. Ganesh, S. Sastry, Y. Ma, Robust Face Recognition via Sparse Representation, IEEE Transactions on Pattern Analysis and Machine Intelligence 31(2)(2009) 210-227.

[16] I. Ramirez, G. Sapiro, Universal sparse modeling, Technical Report, arXiv:1003.294v1[cs.IT], University of Minnesota, 2010.

[17] S.H. Ji, Y. Xue, L. Carin, Bayesian compressive sensing, IEEE Transactions on Signal Processing 56(6)(2008) 2346-2356.

[18] Y. Liu, F. Wu, Z. Zhang, Y. Zhuang, S. Yan, Sparse representation using nonnegative curds and whey, in: Proc. IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2010.

[19] S.H. Gao, I.W.H. Tsang, L.T. Chia, P. Zhao, Local features are not lonely-laplacian sparse coding for image classification, in: Proc: IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2010.

[20] M. Yang, L. Zhang, Gabor feature based spare representation for face recognition with gabor occlusion dictionary, in: Proc. 11th European Conference on Computer Vision (ECCV 2010), 2010.

[21] J.C. Yang, K. Yu, T. Huang, Supervised translation-invariant sparse coding, in: Proc. IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2010.

[22] 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.

[23] L.M. Zhang, S. Chen, L. Qiao, Graph optimization for dimensionality reduction with sparsity constraints, Pattern Recognition 45(3)(2012) 1205-1210.

[24] L.S. Qiao, S.C. Chen, X.Y. Tan, Sparsity preserving projections with applications to face recognition, Pattern Recognition 43(1)(2010) 331-341.

[25] J. Wright, Y. Ma, Dense error correction via l1-minimization, IEEE Transactions on Information Theory 56(7)(2010) 3540-3560.

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