Journal of Computers Vol. 28, No. 2, 2017, pp. 247-259
doi:10.3966/199115592017042802020

The Fusion of Gabor Feature and Sparse Representation for Face Recognition

Yujuan Hao, Liquan Zhang, and De Zhang

College of Electrical and Information Engineering, Beijing University of Civil Engineering and Architecture, Beijing100044, China
1243476593@qq.com

Received 21 September 2016; Revised 21 December 2016; Accepted 21 January 2017

Abstract. In the field of face recognition, the application of sparse representation is very successful, even if the images are taked under different illumination or including facial expression, it still has 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 a wrong face recognition. In sparse representation, the l_1-norm was used to define the fidelity of sparse coding. And the face feature extraction based on sparse representation is too simple, and the sparse coefficient is not sparse. In this paper, we propose a simple and effective face recognition algorithm, in which the classification algorithm sparse representation and Gabor feature are fused effectively. The useful information of feature can be fully reflected in the sparse representation. 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 classificcation. We conduct several experiments on publicly available database to verify the efficacy of the proposed approach and corroborate our claims.

Keywords: collaborative representation, Gabor feature, sparse coefficients, sparse representation, the least square L1 norm

1 Introduction

The face recognition has the advantage of convenience and safety, easy to accept, not easy to counterfeit, which are widely used in the identification, video retrieval, safety monitoring and other fields. It is a hot research topic in the field of pattern recognition and artificial intelligence. In recent years, with development of compressed sensing theory and the L1 norm optimization technology, sparse representation has received the attention of many scholars at home and abroad.

In sparse representation, a signal can be represented as the most sparse linear combination of a given dictionary atom. Therefore, sparse representation has been widely used in image denoising and recovery, sparse representation has been widely used in pattern recognition and face recognition. In 2009, Wright et al. applied sparse representation to the field of face recognition, and achieved good recognition results in the face image under the condition of noise pollution or partial occlusion [1]. The principle of sparse representation is to use the dictionary to process the signal processing [2], The signal represents a linear combination of a small number of atoms. Many scholars have done researches and improvement based the sparse representation algorithm.

Yang et al. discuss the rationality of L1 norm solution of SRC and the validity of the classification method [3]. Zhang, Yang and Feng believe that collaborative representation is more important than the L1 norm in the effective face recognition, and proposed a collaborative representation classification method (CRC), which is more efficient in computing [6]. However, CRC does not provide a method to remove the image noise, so there is still a lack of recognition. Based on kernel principal component analysis (KPCA) and Fisher linear discriminant analysis (LDA), a complete kernel Fisher discriminant analysis (CKFD) algorithm [4]. Cevikalp propose kernel discriminant common vector method. The local tangent space alignment algorithm (LTSA) is proposed by Zhang and Zha [9], Zhang, Wang and Zha propose an adaptive manifold learning based on LTSA[10]. Lai [11] put up a sparse 2D local discriminant projection (S2DLDP) algorithm. Research raise discriminant analysis of robust feature extraction of principal component analysis based on non parametric maximum entropy principle [7-8]. Lai et al. proposed GSRC (based SRC Gabor-feature) method, which uses Gabor features significantly reduce the size of the block dictionary and improve the rate of face recognition [12].

Fu and Huang [16] propose a neighborhood preserving embedding algorithm, the proposed algorithms not only solve the weakness of traditional linear methods such as PCA, which are difficult to maintain the nonlinear flow of the original data, but also solve the shortcomings of the nonlinear method which are hard to obtain the low dimensional projection of the new sample points. Chen et al. [17] present a local discriminant embedding (LDE) method for popular learning and pattern classification. Yan et al. [18] propose a graph embedding framework to develop a new dimension reduction algorithm. Cai et al. [19] propose a subspace method of spatial smoothing for face recognition and orthogonal Laplacianfaces face recognition [21]. Fan, Xu and Zhang [20] raise an improved LDA framework (LLDA) that can effectively capture the local structure of the samples.

Based on the current level of technology development, L1 norm sparsity determines SRC method whether it can be successfully applied to the problem of face recognition. And for this issue, Zhang and Yang [13] analysis and experiment the basic principle of SRC, and propose that SRC is more important than L1 norm sparsity in the classification methods based SRC, Zhang proposed a classification method based on sparse representation and rule of least squares SRC-RLS [14-15]. Compared with the CRC method [6], (SRC-RLS) the results of the recognition is not only much better, but also can reduce the complexity of calculating. However, the SRC-RLS method directly uses the gray characteristics of the image recognition, so when the situation changes of light, posture and facial expressions, the accuracy of this recognition method will not be very high.

In the second part, we introduced Gabor feature extraction principle and our improved method. We lists the data of the method, in the third part, we introduce the sparse representation theory and its shortcomings, and point out that the sparse coefficients contain negative values which affect the result of the experiment, and introduces the collaborative representation. Finally, we validate the effectiveness of our method on a database which can improve the robustness of face recognition.

This paper proposes a new face recognition algorithm based on a improved Gabor feature and sparse representation (GSRC), and the least square method is used to replace the L1 solution. After the feature extraction of face recognition image, we fully constitute a complete dictionary of facial features, and then fuse the sparse representation algorithm, which will improve the recognition rate.

2 Gabor Feature Extraction

2.1 Gabor Filter Principle

Gabor function is obtained through the Fourier transform, the spectrum feature of Fourier transform is only in the infinite time domain which is difficult to fit for the local signal frequency analysis . A part of the time domain signal is multiplied with the Gauss window function and the Fourier transform is performed, we can get Gabor transformation. It was first proposed by Gabor Dennis, in 1946. The Gabor transform will be defined as:

$$G_f(\omega,b)=\int_{-\infty}^{\infty}e^{-i\omega t}g_a(t-b)f(t)dt \tag{1}$$

The function of Gabor filter in the image processing is to achieve the input signal wavelet transform [17], so as to obtain the image of the local texture characteristics:

$$\begin{gathered}I_{mlpq}=\iint f(x,y)\varphi_{ml}(x-p\Delta x,y-q\Delta y)dxdy\\(m=0,\cdots,M-1,\ l=0,\cdots,L-1)\\\varphi_{ml}(x,y)=a^{-m}\varphi(\tilde{x},\tilde{y})\end{gathered} \tag{2}$$

M represents directions, L represents the scales; p represents the x-coordinate of the image pixels, q represents the ordinate of the image pixels; Δx and Δy are spatial sampling interval, usually set $\Delta x = \Delta y = 1$, $\varphi_{ml}(x, y)$ is the parent wave:

$$\tilde{x} = a^{-m}(x\cos\theta + y\sin\theta), \tilde{y} = a^{-m}(-x\sin\theta + y\cos\theta) \quad \textbf{(3)}$$

Direction change θ is defined as: $\begin{array}{l} \theta = l\Delta\theta \\ \Delta\theta = 2\pi / L \end{array}$

θ can choose different angles with the same spacing in 2π, thus making the filter in different directions have the same size of energy.

$$g(x, y) = \left(\frac{1}{2\pi\delta_x\delta_y}\right)\exp\left[-\frac{1}{2}\left(\frac{x^2}{\delta_x^2} + \frac{y^2}{\delta_y^2}\right) + 2\pi jW_x\right] \quad \textbf{(4)}$$

W represents frequency bandwidth. When $W = 0.5$, the filter can match the visual system. The frequency characteristic of Gabor wavelet can be realized in the space of u axis and v axis:

$$\begin{aligned} \psi(u, v) &= \exp[-2\pi^2(\delta_x^2 u^2 + \delta_y^2 v^2)] * \delta(u - W) \\ &= \exp\left(-2\pi^2\left(\delta_x^2\left(u - W\right)^2 + \delta_y^2 v^2\right)\right) \\ &= \exp\left(-\frac{1}{2}\left(\frac{(u - W)^2}{\delta_u^2} + \frac{v^2}{\delta_v^2}\right)\right) \end{aligned} \quad \textbf{(5)}$$

Frequency variance $\delta_u = \dfrac{1}{2\pi\delta_x}, \delta_v = \dfrac{1}{2\pi\delta_y}$.

The However, the output of Gabor filter in high frequency state is relatively small, and it is necessary to change the basis function calculation method.

$$\hat{\varphi}_{ml}(x, y) = \varphi(\tilde{x}, \tilde{y}) \quad \textbf{(6)}$$

$$\hat{I}_{mlpq} = \iint f(x, y)\hat{\varphi}(x - p\Delta x, y - q\Delta y)dxdy \quad \textbf{(7)}$$

After filtering for each pixel, the resulting is that:

$$F_{mlpq} = \left|\hat{I}_{mlpq}\right| \quad m = 0, \cdots, M - 1 \quad l = 0, \cdots, L - 1 \quad \textbf{(8)}$$

For people's facial features, there is also a unique characteristics of unity. Uniformity is the consistent pattern of the distribution of the eyes, nose and mouth, eyes located on the top of the face, both sides of the nose, nose is almost in the middle of the face, the bottom is the mouth. This rule allows us to clearly distinguish between people and other things. But in order to separate people and people, it should make full use of the unique characteristics of these key features, including the texture distribution of facial organs, geometric distance and the correlation between the various organs, Only grasp these small differences, we can distinguish the different people from the crowd, which is also the main theoretical basis for face recognition. First, we need to determine the general location of each organ in the face distribution according to the unity. We use the "three vertical and five horizontal" law, face detection in the vertical direction is divided into three parts: from the forehead to the eyebrows is the first part, the eyebrows to the nose is for the second part, below the tip of the nose act as the third part. In the horizontal direction, the face is divided into five equal parts, each is equal to the width of a eyes, so that you can determine two eyes will fall in the second and fourth. Such a rule can be roughly positioned to the various organs of the human face.

Fig. 1 shows the location of the eyes, nose and mouth.

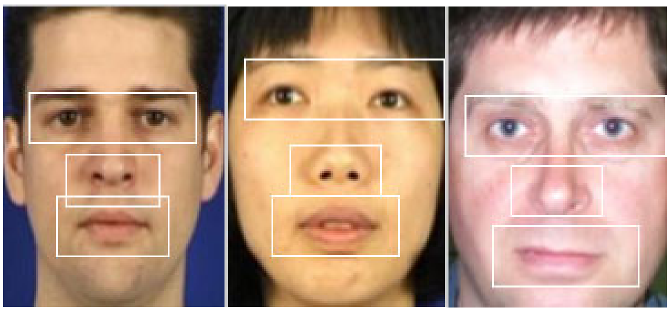


Fig. 1. Facial organ location

We chose 40 filters to filter the image which contains 8 directions and 5 scales . The characteristics described by each filter is different. In order to be more intuitive to identify these differences, we intercepted a few filtered images were analyzed. the end; spacing before 6 pt, after 6 pt.

Fig.2 shows the characteristics of different features after filtering.

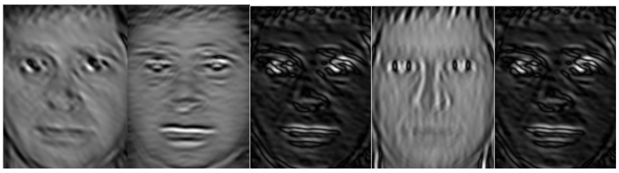


Fig. 2. Filtered image

In the face images, we need extract the key feature points, we choose the distance between each feature points as the feature vector of the face.

In Fig. 3, we show the process of the feature extraction. Eye feature points: after analysis of the color components, in the Cb image, the eye section shows a high pixel value, in Cr image, the eye section shows a low pixel value. Considering the character and select the appropriate threshold which can separate the eyes into a black area (the eye) and two white areas (eye ball on either side of the white part). Select exterior point of the white part (corner of the eye), then we get the center point of the eyeball in the black area. Finally, we can get 6 feature points of the eyes. Feature point extraction: mouth is located in the face of the bottom, relative to the eyes, it is easy to get the feature points of the mouth, Because the size of the pixel values in the Cr component has a great relationship with red, The red lips make it more in the pixel in the Cr component. Thus, the mouth region can be segmented from the image, then the sobel filter can be used to extract its contour, and then it can extract the characteristic points of the extreme point on both sides of the mouth. Nose feature points: the nose lies middle lower part between the eyes and the mouth, at the top of the mouth. According to this inherent attribute, we can roughly determined in the image of the nose area. For the extraction of the feature points of the nose, it also can be expressed by the Sobel filter, and then extract the points on both sides of the nose root as a feature point. We choose the root of the nose rather than the nose to do feature points, because different people have different textures in the root of the nose.

Fig. 3. Feature map of geometric distance

2.2. Gabor Feature Extraction Method

Based on the Gabor function, principle and design theory of filter, we have made some changes in the selection of Gabor kernel function, so as to be conducive to the implementation of the program. The two-dimensional Gabor kernel function we choose are shown in [18-19].

$$\begin{cases} W(x,y,\theta,\lambda,\phi,\sigma) = \exp\left(-\dfrac{x'^2+y'^2}{2\sigma^2}\right)\cos\left(2\pi\dfrac{x'}{\lambda}+\phi\right) \\ x' = x\cos(\theta) + y\sin(\theta) \\ y' = -x\sin(\theta) + y\cos(\theta) \end{cases} \tag{9}$$

(x, y) representative a *pixel coor*dinates in the diagram, $\theta,\lambda,\phi,\sigma$ The parameters of $\theta,\lambda,\phi,\sigma$ are shown in Table 1.

Table 1. Data of $\theta,\lambda,\phi,\sigma$

Parameters	Symbol	Set value
direction	θ	$\left\{0,\frac{\pi}{8},\frac{2\pi}{8},\frac{3\pi}{8},\frac{4\pi}{8},\frac{5\pi}{8},\frac{6\pi}{8},\frac{7\pi}{8}\right\}$
wavelength	λ	$\left\{4,4\sqrt{2},8,8\sqrt{2},16\right\}$
phase	ϕ	$\left\{0,\frac{\pi}{2}\right\}$
Gauss radius	σ	$\sigma=\lambda$

Based on the Gabor function, principle and design theory of filter, we have made some changes in the selection of Gabor kernel function, so as to be conducive to the implementation of the program. The two-dimensional Gabor kernel function we choose are shown in [18-19]. Direction θ : Represent the direction of parallel stripes in the Gabor function. Wavelength λ : In image processing, it is said that the unit is a pixel, there is a certain range of needs less than the size of the image 1/5, the minimum is 2. Phase φ : In the range of -180 degrees to 180 degrees, the center-off function is obtained at 90 degrees, the Center-On function is obtained at 0 degrees. Gauss radius σ : Represent the standard deviation of Gauss factor in the Gauss's function used in Gabor transform, $\sigma = \lambda$.

A set of filters consists of 5 different spatial frequencies and 8 different directions, which make up 40 different filters, as shown in Fig. 4. The figure shows the characteristics of different features after filtering.

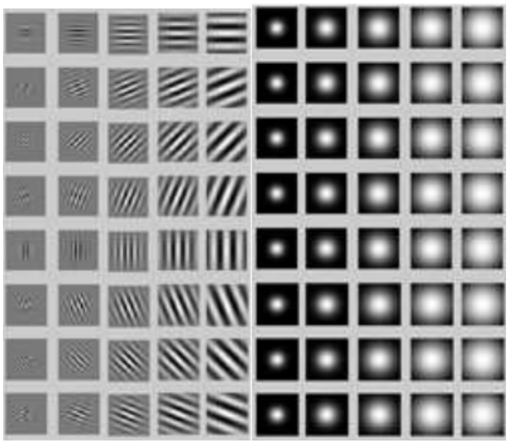


Fig. 4. Filtered image

When we use this set of filters to deal with a face image, the number of processed images is equal to the number of filters. If these filters are applied to each pixel in the image, the filtered vector dimension is very large (proportional to the size of the image). So, this will lead to a huge amount of computing and storage costs, which will consume a lot of time to classify and identify. In order to solve this problem and make the algorithm has some robustness, our method is only applied to the filter to extract the ten benchmark points, to obtain the filter response. If the number of filters is M, then the vector dimension of the filter response is 10 * M. The number of filters corresponding to different values of θ is shown in Table 2.

Table 2. Numbers of filter and the values of θ

Numbers of filter	Values of different directions (θ)
5	$\{0\}$
10	$\left\{0,\frac{4\pi}{8}\right\}$
20	$\left\{0,\frac{2\pi}{8},\frac{4\pi}{8},\frac{6\pi}{8}\right\}$
30	$\left\{0,\frac{\pi}{8},\frac{2\pi}{8},\frac{4\pi}{8},\frac{5\pi}{8},\frac{6\pi}{8}\right\}$
40	$\left\{0,\frac{\pi}{8},\frac{2\pi}{8},\frac{3\pi}{8},\frac{4\pi}{8},\frac{5\pi}{8},\frac{6\pi}{8},\frac{7\pi}{8}\right\}$

Finally, we choose to characterize the facial feature vectors including the Gabor filter after the response of the 10 x M values and 8 geometric distance parameters, in this paper, we use such a set of feature vectors to perform a high recognition rate, and shorten the recognition time.

3 Sparse Representation Robust Recognition

3.1 Sparse Representation Theory

In the face recognition technology based on sparse representation, each image size is $w \times h$, which can be regarded as a point of m dimensional space, $m = w \times h$. We define n_i pictures., n_i training samples of i subject form a matrix, $A_i = [v_{i,1}, v_{i,2}, \cdots, v_{i,n}] \in R^{m \times n_i}$. The matrix A is defined for the whole training sample set, it is regarded as a series of k classes of training samples. All n_i are added together: $n = n_1 + n_2 + \cdots + n_k$, we regard n columns as vector A: $A = [A_1, A_2, \cdots, A_k] = [v_{1,1}, v_{1,2}, \cdots, v_{k,n_k}]$, the test sample y can be expressed as a linear combination of the training sample set.:

$$x_0 = [0, \cdots, a_{i,1}, a_{i,2}, \cdots, a_{i,n_i}, 0, \cdots, 0]^T \in R^n \quad \textbf{(10)}$$

x_0 is coefficient vector, Non - zero items of x_0 should be corresponding to the class i of the training set. By sparse representation and compressive sensing theory, if the solution of x_0 is sparse enough, then the problem of l^0 norm can be replaced by the minimum l^1 norm.

In summary, the process of the classical sparse representation recognition algorithm can be described as follows:

Algorithm 1 Classifier based on sparse representation (SRC):

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

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

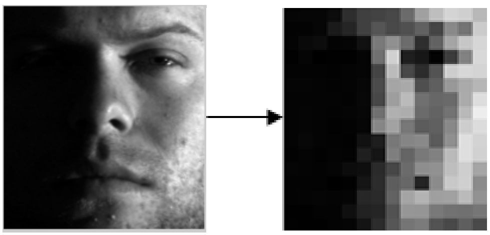
3. Solve the l^1 minimization problem: $\hat{x}_1 = x \arg\min \|x\|_1$ subject to *to* $Ax = y$ or alternatively, solve: $\hat{x}_1 = x \arg\min \|x\|_1$ subject to $\|Ax - y\|_2 \le \varepsilon$.

4. Compute the residuals: $r_i(y) = \|y - A\delta_i(x_i)\|_2$, $\delta_i(\hat{x}) \in R^n$ is a new vector, its non 0 element is the element in x, and it is related to the target i, and other elements are 0. $i = 1, 2, \cdots, k$.

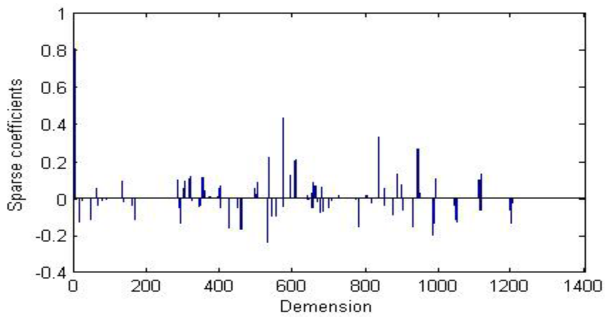
5. Output: $identify(y) = \arg\min_i r_i(y)$.

When the noise of an image is large, a small amount of negative correlation coefficient (absolute value larger) will also appear in the target class, other classes will also contain a greater positive values, but also have multiple sets of negative values. This is because sparse coefficients are not sparse, when SRC is applied into non-ideal circumstances such as illumination variation, occlusion and posture. When the image local space appears specular reflection, projection or occlusion, SRC processes it as the large scale sparse error. In addition, pose variations also result in a large number of nonlinear transformations, which can reduce the recognition performance.

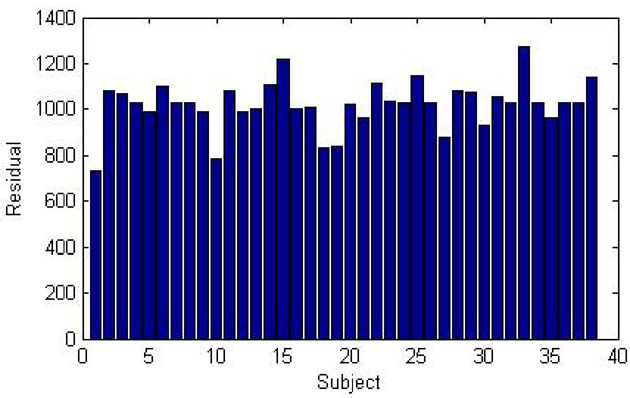
We choose the ORL face database to illustrate the shortcomings of sparse representation. We select an image of subject 8, which have glasses occlusion from the test images to detect, whose mouth posture change, the recognition result is not correct caused by linear conversion errors, the results as shown in Fig. 5 (a). From the Fig. 5 (b) and Fig. (c) show that the maximum coefficient does not belongs to the correct subject by the classical SRC method. The error of training images detected although sparse coefficient is the largest, but it also has a negative correlation coefficient absolute value maximum, other values are mostly negative. So the effect is not obvious.



(a)



(b)



(c)

Fig. 5. Classical sparse representation. (a) Left: Test sample belongs to subject 1. Right: Recognition with 12×8 downsampled images as features. (b) Sparse coefficients calculated by SRC. (c) 38 humanoid residuals by SRC, the ratio between the two smallest residuals is 1: 1.1

We choose to use Yale B Extended face database to illustrate how our algorithm is to improve the effectiveness of the judgment of the light. The database contains 38 people face images, each person has 64 photos under different illumination conditions, we randomly selected 32 samples as the training samples, and the rest as the test samples. Under the illumination, the improved residual is more distinct than the classical algorithm. We select subject 1 in the shadow image generated by the light to prove the simplicity of the method. In Fig. 5 (c) shows that the residuals of subject 1, subject 9, subject 18, subject 19 are similar, so there is a strong interference. According to the distribution of coefficients in Fig. 5 (b), subject 1 have small negative values and the maximum value of the entire sparse coefficient, subject 9 contain more positive coefficients, subject 18and subject 19 contain larger positive values in the overall coefficients, and also contain more negative values. Once again we merge the sparse coefficients into the residual, The results of these residuals can reflect the size of the sparse coefficient, and the positive and negative information, we can find the maximum residual from the positive and ignore the negative values. In face recognition, changes in illumination have a greater impact on the human faces, varying degrees of light or dark shaded area caused by light intensity or angle will reduce the recognition rate of SRC algorithm, the sparsity of coefficients is pathetic calculated by sparse representation, residuals $r_i(y)$ obtained by the sparse coefficient are similar, then the classification is not obvious. This is because the residual calculation formula is $r_i(y) = \|y - A\delta_i(x_i)\|_2$, The value of $r_i(y)$ only changes with the change of $\delta_i(x_i)$, that is to say, what really determines the residuals is the coefficient vector. But because the values of the coefficients are small, the residuals calculated by l_2-norm can not express the negative correlation information of the sparse coefficients.

Feature obtained by the down sampling is not obvious, it also can not express the face feature. And the sparse coefficients of the traditional sparse said method are negative, which has great interference in the category judgment, if the coefficient is positive, this problem can be optimized.

3.2 Face Recognition Algorithm Based on Collaborative Representation

A number of literatures have highlighted the importance of sparse for classification, and have not studied the role of various training samples in the collaborative representation of test samples. Improve the accuracy of face recognition, A collaborative representation based classification (CRC) method is proposed in the literature [6]. In a practical face recognition system, in order to ensure the high recognition rate, feature dimension usually cannot be set too low. Therefore, there is no need for us to l1 norm. Taking into account the dictionary X may be under determined, we use 2L norm regularization. In order to use X to express the test sample y, we propose the following collaborative representation of the regularized least squares method:

$$\hat{a} = \arg\min_{a} \left\{ \|y - xa\|_2^2 + \lambda \|a\|_2^2 \right\} \tag{11}$$

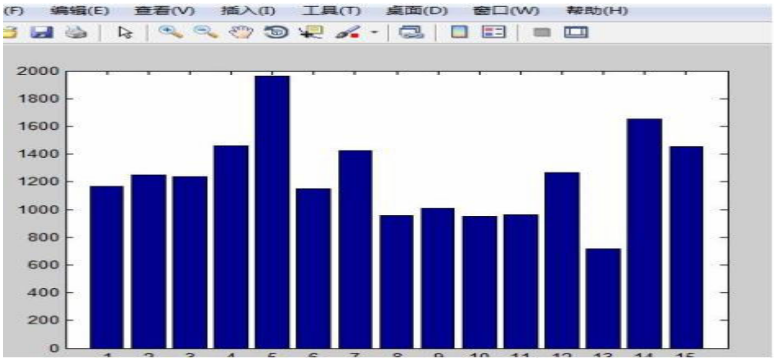
λ is regularization parameter, The function of l^2-norm regularization $\|a\|_2$ is twofold. First of all, it makes the least squares solution stable, especially when X is less regular; secondly, it introduces a certain number of sparse to $\hat{a}$. However, this degree of sparsity is weaker than the l^1-norm. we can get this:

$$\hat{a} = \left(x^T x + \lambda I\right)^{-1} x^T y \tag{12}$$

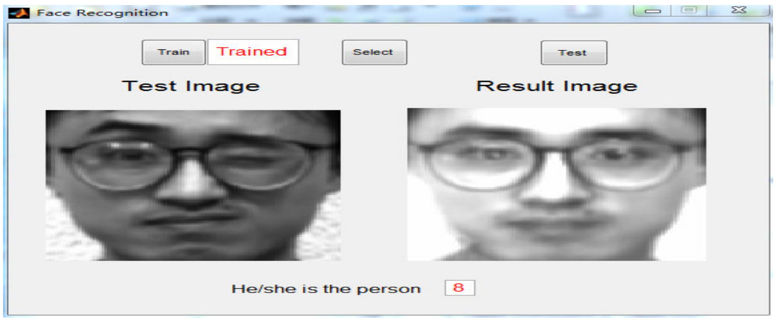
Order $P = \left(x^T x + \lambda I\right)^{-1} x^T$, Obviously, P is independent relative to y, We can calculate P as a projection matrix. Once the test sample y arrival, we will simply pass the P to thev Py, which makes the calculation speed of the collaborative representation is very fast.

Depending on the classification of $\hat{a}$ in the CRC and the classification of the $\hat{a}$ depending on the SRC are similar. In addition to using the specified class of $\|y - X_j \hat{a}_j\|_2$ to classify the reconstructed residuals where $\hat{a}_j$ is corresponding to the class of coding coefficient vector, l^2-norm sparse term can also bring some discriminant information. Therefore, we propose to use both of them in classification, so that the classification accuracy can be improved by the use of $\|y - X_j\hat{a}_j\|_2$.

(a) The recognition result



(b) The residuals of 15 people



(c) The correct result of our method

Fig. 6. The recognition result of our methord

We choose the ORL face database to illustrate our method, test image belongs to subject 8, the result of SRC shows it belongs to subject 13.Our merhod is correct.

3.3 Combination of Improved Gabor Features and Sparse Representation

When the images exist illumination, expression and pose variation, the recognition rate traditional sparse representation of will be decreased. In order to further improve the algorithm accuracy and robustness, We will combine feature extraction method which characterize the facial feature vectors including the Gabor filter in response to the 10 x M value and 8 geometric distance parameters of the feature extraction method with sparse representation algorithm, And the least square method is used to solve the norm, and the condition of the sparse coefficient is positive.

According to the proposed extension of the Gabor feature definition, Extracting all the Gabor features of the training sample set image, the corresponding Gabor feature set is extracted from the eyes, nose and lips, and set up the corresponding feature set. $X_i = [x_{i,1}, x_{i,2}, \ldots, x_{i,n_i}]$, The Gabor feature set for all classes of training samples is $X = [X_1, X_2, \ldots, X_K]$, All the training samples are set to represent the test samples., That is to simplify the SRC algorithm based on norm of sparse constraints to the rule of least squares method,

$$(\hat{p}) = \arg\min_{p} \left\{ \|y - Xp\|_2^2 + \lambda \|p\|_2^2 \right\} \tag{13}$$

p is coefficient vector, λ is expressed rule parameter, Regularization term $\lambda\|p\|_2^2$ has a double effect. First of all, it makes the least squares solution stable, especially when X is less regular; secondly, it introduces a certain number of sparse to $\hat{p}$. P is independent relative to y, We can calculate P as a projection matrix. Once the test sample y arrival, we will simply pass the P to the Py, which makes the calculation speed of the collaborative representation is very fast.

Improved algorithm flow:

1. Extract the Gabor features of the three parts of the training sample set and the test sample, then respectively get X and Y;
2. The test sample feature vector y is based on the training sample set X collaborative representation coding $\hat{p} = Py$, The projection matrix is $P = \left(x^T x + \lambda I\right)^{-1} x^T$;
3. calculate regular residual $r_i = \|y - X_i \hat{p}_i\|_2 / \|\hat{p}_i\|_2$, $\hat{p}_i$ is the encoding coefficient vectors for class i;
4. Output: $identity(y) = \arg\min_i \{r_i\}$

4 Simulation and Results

In order to verify the effectiveness of the algorithm, we conduct a simulation experiment, we select two standard face recognition databases, that is, the performance comparison test of the proposed algorithm is carried out on Extended Yale B database and AR database. In this paper, we use the support vector machine classification algorithm (SVM) and the classical SRC algorithm to compare with the improved algorithm.

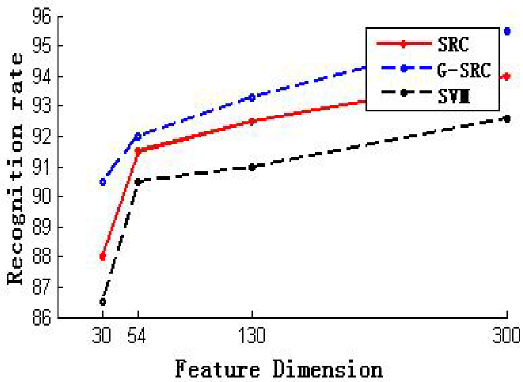
4.1 Experiment on Extended Yale B

The Extended Yale B face database contains 38 individual frontal face image, each in different illumination conditions have 64 pictures, the size of each photo is 192×168. We randomly selected 32 samples as training samples, and the rest as the test samples. Like the classical sparse representation, we select the feature space dimension is 30,54,130,300, Fig. 7 Recognition rate of three kinds of classifiers in four kinds of feature dimension.

4.2 Experiments on the AR Database

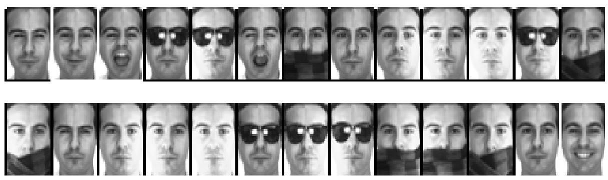
AR face database contains a total of 4000 images collected from 126 volunteers, including the changes in light, angle, facial expression and occlusion. We select 50 classes, each class of 26 samples, the image size is 165×120. We choose 14 images per class as a training sample, and the others are used as test samples. The feature space dimension is 30,54,130,300. Fig. 8 The experimental results obtained by the three algorithms with the change of dimension.

(a) Test sample images

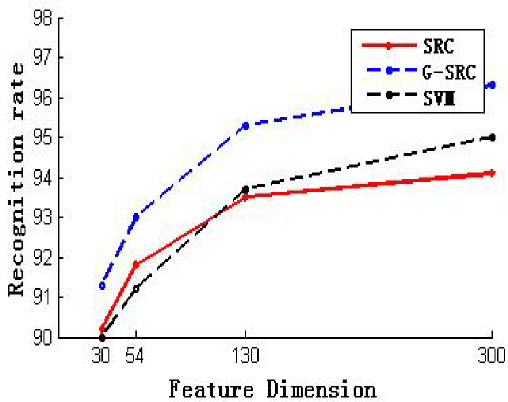


(b) Recognition rate of three kinds of classification methods

Fig. 7. Recognition rate on extended Yale B database



(a) Sample of someone



(b) The recognition rate of three kinds of classification methods

Fig. 8. Recognition rate on AR database

Acknowledgments

In the research of image recognition, feature extraction can be described as an essential step. The advantage of feature extraction determines the accuracy of the recognition. We propose an improved Gabor feature extraction algorithm to obtain the accurate position of the eyes and mouth and extract features; Using Sobel edge detection technology determines the position of the nose and extract the feature points.

We calibrate the facial feature points, then apply Gabor filter at each characteristic point respectively, The purpose of this process is to extract the local details of the facial organs.

The Gabor feature extracted by this method is not only low in dimension, but also has good robustness. Then the least square method is used to solve the solution, which ensures the non-negativity of the coefficients. The experimental results show that the new method is robust to illumination, occlusion and other changes.

References

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

[2] J.Z. Huang, X.L. Huang, D. Metaxas, Simultaneous image transformation and sparse representation recovery, in: Proc. the 26th IEEE Conference on Computer Vision and Image Recognition, 2008.

[3] J. Yang, L. Zhang, Y. Xu, J.-Y. Yang, Beyond sparsity: the role of L1-optimizer in pattern classification, Pattern Recognition 45(3)(2012) 1104-1118.

[4] 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.

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

[6] L. Zhang, M. Yang, X.C. Feng, Sparse representation or collaborative representation: which helps face recognition?" in: Proc. IEEE International Conference on Computer Vision, 2011.

[7] R. He, W.-S. Zheng, B.-G. Hu, Maximum correntropy criterion for robust face recognition, IEEE Transactions on Pattern Analysis and Machine Intelligence 33(8)(2011) 1561-1576.

[8] R. He, B.-G. Hu, X. T. Yuan, Robust discriminant analysis based on non-parametric maximum entropy, in: Proc. Asian Conference on Machine Learning, 2009.

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

[10] Z.Y. Zhang, J. Wang, H.Y. Zha, Adaptive manifold learning, IEEE Transaction on Pattern Analysis and Machine Intelligence 32(2)(2012) 253-265.

[11] Z. Lai, M. Wan, Z. Jin, J. Yang, Sparse two-dimensional local discriminant projections for feature extraction, Neurocomputting 74(4)(2011) 629-637.

[12] Z. Lai, M. Wan, Z. Jin, J. Yang, Sparse two-dimensional local discriminant projections for feature extraction, Neurocomputting 74(4)(2011) 629-637.

[13] N. Zhang, J.K. Yang, Nearest neighbor based local sparse representation classifier, in: Proc. the 2010 Chinese Conference on Pattern Recognition, 2010.

[14] R. Tibshirani, Regression shrinkage and selection via the lasso: a retrospective, Journal of the Royal Statistical Society:

Series B 73(3)(2011) 273-282.

[15] J. Zhang, R. Jin, Y.M. Yang, A.G. Hauptmann, Modified logistic regression: an approximation to SVM and its applications in large-scale text categorization, in: Proc. the 20th International Conference on Machine Learning, 2003.

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

[17] H.-T. Chen, H.-W. Chang, T.-L. Liu, Local discriminant embedding and its variants, in: Proc. IEEE Conference on Computer Vision and Pattern Recognition, 2005.

[18] S. Yan, D. Xu, B. Zhang, H. Zhang, Q. Yang, S. Lin, Graph embedding and extension: a general framework for dimensionality reduction, IEEE Transaction on Pattern Analysis and Machine Intelligence 29(1)(2007) 40-51.

[19] D. Cai, X. He, Y. Hu, J. Han, T. Huang, Learning a spatially smooth subspace for face recognition, in: Proc. IEEE Conference on Computer Vision and Pattern Recognition, 2007.

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

[21] D. Cai, X. He, Y. Hu, J. Han, and T. Huang, "Learning a spatially smooth subspace for face recognition," in Proc. IEEE Conference on Computer Vision and Pattern Recognition, 2007.