Journal of Information Hiding and Multimedia Signal Processing ©2014 ISSN 2073-4212
Ubiquitous International Volume **5**, Number **3**, July **2014**

Face Recognition Based on Multi-Wavelet and Sparse Representation

Lijian Zhou and Yukai Xu

College of Communication and Electronic Engineering
Qingdao Technological University
Qingdao, 266033, P. R. China
zhoulijian@qtech.edu.cn

Zhe-Ming Lu

School of Aeronautics and Astronautics
Zhejiang University
Hangzhou, 310027, P.R. China
zheminglu@zju.edu.cn

Tingyuan Nie

College of Communication and Electronic Engineering
Qingdao Technological University
Qingdao, 266033, P. R. China.
tynie@qtech.edu.cn

Received November, 2013; revised November, 2013

Abstract. *The feature dimension and redundancy can reduce the face recognition speed and rate, the shading and light changing can heavily affect the face recognition effect. So the first key to face recognition is how to effectively extract face features because the face image contains a lot of redundant information. Multi-wavelet transform has symmetry, orthogonality, compact support and high vanishing moments simultaneously, which can present the face features better than scalar wavelet in each band. By analyzing all the multi-wavelet frequency band components, we can see that the low frequency component of the face image can provide the main feature of a face, and the high frequency components often contain some noises caused by the external interference. Therefore, the low frequency component is often adopted to form the face features. For solving the second problem, we choose the sparse representation recognition method , which has strong robustness to shading and light changing problem. So this paper proposes a face recognition method, which combines multi-wavelet with sparse representation recognition. First, we extract the face features by multi-wavelet, then establish a fully redundant dictionary, and use the sparse representation recognition algorithm for face recognition at last. Based on the YALE, the AR and the FERET face databases, the experimental results show that our method can effectively recognize faces, reduce the dimension of features and has good robustness.*
Keywords: Face recognition, Multi-wavelet, Sparse representation.

1. **Introduction.** Compared with other biometrics technologies such as fingerprint recognition, iris recognition and voice recognition, face recognition has wider application prospects in finance, security, law and other fields, because it is natural and harmless, thus it gets more and more attentions.

Sparse representation[1] can describe signals by a linear combination of bases or very small amounts of elements in the dictionary, in this way, the inherent nature of the image can be fully used. Yang[2] et al. utilized the sparse representation (SR) method directly for face recognition, but they directly adopted the downsampling method to reduce the dimension of face database. This method does not effectively extract facial features, so the recognition rate is not very high. Zeng [3] et al. introduced the wavelet transform into the sparse representation, and adopted the Gabor wavelet to extract face features. Combined with the SRC (Sparse Representation-based Classifier) algorithm, they presented the GSRC face recognition algorithm. Although this method can get good recognition effect, the computational cost is very high. The reason is that the Gabor wavelet often has five scales and eight directions, resulting in forty Gabor face feature sub-images after each input face image is filtered. On the basis of the Gabor wavelet, Li[4] et al. presented a new SR based face recognition method, which partitions the energy sub-bands of the Gabor wavelet filtered image. This method can achieve the purpose of strengthening the local features, however it faces the problem of the large amount of calculation. So the feature extraction is an important factor to solve the high computational complexity problem. By analyzing the feature extraction method, we found that multi-wavelet[5] has not only good localization features in time and frequency domains, but also symmetry, orthogonality, compact support and high vanishing moments , which can present the face features better than scalar wavelet in each band.

Therefore, in order to better solve the high computational complexity in face recognition and take the nature of the human face image into account simultaneously, this paper proposes a new face recognition method based on multi-wavelet and sparse representation. First, the training face images must be pre-filtered. Next, the multi-wavelet transform is performed on these images. Then, the low frequency components of the face images are extracted to get the multi-wavelet face features. At last, a fully redundant dictionary is established using the SRC algorithm. During the recognition process, the test face images are pre-filtered by the pre-filter and transformed by the multi-wavelet to extract their low frequency components. And then these low frequency components are projected onto the fully redundant dictionary. At last, the residual is calculated to recognize the face image. Experimental results show that our method can not only improve the recognition rate, but also reduce the feature dimension and the computational complexity.

2. Facial Feature Extraction Based on Multi-wavelet.

2.1. **Multi-wavelet.** As we know, the wavelet transform can overcome the deficiency of the Fourier transform, it has a good localization characteristic in time and frequency domains, thus it gets wider attentions. Multi-wavelet not only inherits the characteristics of the wavelet, but also has symmetry, orthogonality, compact support and high vanishing moments. The orthogonality property can keep the energy of a signal, while the symmetry property is suitable for the human vision system and makes the signal easy to be handled at the border. Furthermore, the compactly supported wavelet filter is a kind of finite impulse response filter, which makes the fast wavelet transform convergent [6]. Usually, a scalar wavelet has a scale function $\phi(t)$ and a wavelet function $\psi(t)$, but multi-wavelet has many scale functions and many wavelet functions at the same time,which are called multi-scale function and multi-wavelet function respectively, usually represented by multidimensional vector functions. Let $\mathbf{\Phi}(t) = [\phi_1(t), \phi_2(t), \cdots, \phi_r(t)]^T \ (r \in N)$ represent the r-dimensional multi-scale function in the multi-resolution analysis space $\{V_k\}_{k \in Z}$. The corresponding multi-wavelet function is denoted by $\mathbf{\Psi}(t) = [\psi_1(t), \psi_2(t), \cdots, \psi_r(t)]^T \ (r \in N)$.Thus, $\mathbf{\Phi}(t)$ and $\mathbf{\Psi}(t)$ constitute the orthogonal basis in the orthogonal complementary subspace $\{W_k\}_{k\in Z}$. That is to say, W_j represents the orthogonal complementary space of V_j in the subspace V_{j+1} . $\mathbf{\Phi}(t)$ and $\mathbf{\Psi}(t)$ satisfy the following two scale equations

$$\mathbf{\Phi}(t) = \sqrt{2}\sum_{k\in Z} \mathbf{H}_k \mathbf{\Phi}(2t-k) \tag{1}$$

$$\mathbf{\Psi}(t) = \sqrt{2}\sum_{k\in Z} \mathbf{G}_k \mathbf{\Phi}(2t-k) \tag{2}$$

where $\{\mathbf{H}_k\}$ and $\{\mathbf{G}_k\}$ are $r \times r$ matrix filters. Then their dual orthogonality condition can be expressed as follows

$$\sum \mathbf{H}_k \mathbf{H}^*_{k-ml} = \sum \mathbf{G}_k \mathbf{G}^*_{k-ml} = \delta_{0l}\mathbf{I}_r \tag{3}$$

$$\sum \mathbf{H}_k \mathbf{G}^*_{k-ml} = \sum \mathbf{G}_k \mathbf{H}^*_{k-ml} = \mathbf{0}_r \tag{4}$$

where * represents the complex conjugate transpose of a matrix, and $\mathbf{I}_r$ is the $r \times r$ unit matrix and $\mathbf{0}_r$ is the $r \times r$ null matrix.

This paper takes the GHM multi-wavelet [7] constructed by Geronimo, Hardin and Massopust and the CL multi-wavelet [8] given by Chui and Lian.

2.2. **Facial Feature Extraction.** Multi-wavelet transform is only applicable to vector signals, therefore, pre filtering for the rows and columns of the image is necessary, then take the rows and columns to form vector signals according to certain rules, at last, multi-wavelet transform is processed. There are two kinds of pre-filters, i.e., repeat row pre-filter (RR) and approximation order pre-filter (AP).

Multi-wavelet transform is the convolution of an image and the multi-wavelet function, namely

$$\mathbf{M}(i,j) = \mathbf{\Psi}(i,j) \otimes \mathbf{I}(i,j) \tag{5}$$

where $\mathbf{\Psi}(i,j)$ is the wavelet kernel function, and $\mathbf{I}(i,j)$ is the input image.

When a gray-level image is transformed by multi-wavelet, we can get four sub-bands, LL1, HL1, LH1 and HH1, as shown in Fig. 1. The sub-band LL1 is the coarse version of the original image, which keeps the low frequency component of the original image. Sub-band LH1 keeps the vertical edge information of the original image, sub-band HL1 keeps the horizontal edge information of the original image and sub-band HH1 keeps the details of the diagonal direction of the original image. Sub-band LL1 contains almost all of the image information, while the other three sub-bands contain the edge information of the image. For face recognition, LH1, HL1 and HH1 belong to the redundant information, therefore, extracting them will increase the complexity of the algorithm. Thus, this paper takes the low frequency sub-band to form face features, which can not only represent the face features well and get rid of the noises caused by the external interfering, but also reduce the dimension of the face features. Taking the pre-filter RR as an example, the feature extraction result is shown in Fig. 2. From Fig. 2, we can see that most of the energy in the face image transformed by the multi-wavelet is in the low frequency component. Therefore, extracting the low frequency components not only gets the main information of the face image, but also reduces the dimension of the face features and the computational complexity.

3. **Sparse Representation Recognition.**

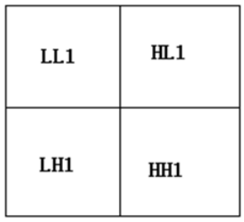

FIGURE 1. One-level multi-wavelet decomposition result.

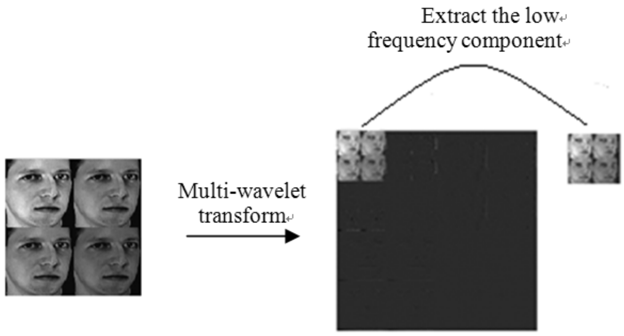

FIGURE 2. Two-level multi-wavelet transform example, where the low frequency component is extracted as face features.

3.1. **Sparse Sensing Representation.** Sparse representation recognition not only has a robustness to the changes in shade and illumination, but also can map the high-dimensional signal to the low-dimensional space according to the inherent nature of the image. Wright [9] et al. introduced the sparse representation method for face recognition in detail as follows.

Given enough training samples in the ith class, denoted by$\mathbf{A}_i \doteq [\boldsymbol{v}_{i,1}, \boldsymbol{v}_{i,1}, \cdots \boldsymbol{v}_{i,n_i}] \in \mathbb{R}^{m\times n_i}$. In the context of face recognition, we will view a $w \times h$ gray-scale image as the column vector of $\boldsymbol{v} \in \mathbb{R}^{m\times n_i} \,(m = w \times h)$. Thus, each column of $\mathbf{A}_i$ is a training face image from the ith class. Thus, any new test sample $\boldsymbol{y} \in \mathbb{R}^m$ from the ith class can be approximately expressed as the linear combination of the training samples in the ith class.

$$\boldsymbol{y} = \alpha_{i,1}\boldsymbol{v}_{i,1} + \alpha_{i,2}\boldsymbol{v}_{i,2} + \cdots + \alpha_{i,n_i}\boldsymbol{v}_{i,n_i} \tag{6}$$

where $\alpha_{i,j} \in \mathbb{R}, j = 1, 2, \cdots, n_i$ are scalar coefficients.

Since the class label of the test sample is initially unknown, we define a new matrix $\mathbf{A}$ for the entire training set as the concatenation of the n training samples of all k classes, i.e., $\mathbf{A} \doteq [\boldsymbol{A}_1, \boldsymbol{A}_2, \cdots, \boldsymbol{A}_k] = [\boldsymbol{v}_{1,1}, \boldsymbol{v}_{1,2}, \cdots, \boldsymbol{v}_{1,n_1}, \boldsymbol{v}_{2,1}, \boldsymbol{v}_{2,2}, \cdots, \boldsymbol{v}_{2,n_2}, \cdots, \boldsymbol{v}_{k,n_k}]$, Then, the linear representation of $\boldsymbol{y}$ can be rewritten in terms of all training samples as

$$\boldsymbol{y} = \mathbf{A}\boldsymbol{x}_0 \;\; \boldsymbol{y} \in \mathbb{R}^m \tag{7}$$

where $\boldsymbol{x}_0 = [0, \cdots, 0, \alpha_{i,1}, \alpha_{i,2} \cdots \alpha_{i,n_i}, 0, \cdots, 0]^T \in \mathbb{R}^n$ is a coefficient vector whose entries are all zero except for those associated with the ith class.

For determining the coefficients of a test sample $\boldsymbol{y}$ over the whole training set, we need to solve the linear equation $\boldsymbol{y} = \mathbf{A}\boldsymbol{x}$. Obviously, if $m > n$, this system equation $\boldsymbol{y} = \mathbf{A}\boldsymbol{x}$ is overdetermined, and the correct and unique $\boldsymbol{x}_0$ can usually be found. However, in robust face recognition, the equation $\boldsymbol{y} = \mathbf{A}\boldsymbol{x}$ is typically underdetermined, and thus its solution is not unique. Conventionally, this problem can be formulated by choosing the minimum l_2-norm solution as follows.

$$\left(l^2\right) : \hat{\boldsymbol{x}}_{(2)} = arg\ min\ \ \|\boldsymbol{x}\|_2\ \ subject\ to\ \mathbf{A}\boldsymbol{x} = \boldsymbol{y} \tag{8}$$

Furthermore, this optimization problem can be easily solved via calculating the pseudoinverse of $\mathbf{A}$, the solution $\hat{\boldsymbol{x}}_{(2)}$ is not especially informative for recognizing the test sample $\boldsymbol{y}$. For example, a valid test sample $\boldsymbol{y}$ can be sufficiently represented using only the training samples from the same class. This representation is naturally sparse if the number of training samples is sufficiently large. The more sparse the recovered $\boldsymbol{x}_0$ is, the easier it will be to accurately determine the identity of the test sample $\boldsymbol{y}$.

Therefore, we can seek the sparsest solution to $\boldsymbol{y} = \mathbf{A}\boldsymbol{x}$ by solving the following optimization problem:

$$\left(l^0\right) : \hat{\boldsymbol{x}}_{(0)} = arg\ min\ \|\boldsymbol{x}\|_0\ \ subject\ to\ \mathbf{A}\boldsymbol{x} = \boldsymbol{y} \tag{9}$$

where $\|\bullet\|_0$ denotes the l_0-norm, which counts the number of nonzero entries in a vector. In fact, the condition of solving $\boldsymbol{y} = \mathbf{A}\boldsymbol{x}$ having the unique sparsest solution has been investigated. However, the problem of finding the sparsest solution of an underdetermined linear equation is an NP-hard problem and it is even difficult to approximate such solution. Recent development in sparse representation reveals that if the solution $\boldsymbol{x}_0$ is sparse enough, the solution of the l_0-minimization problem is equivalent to the solution to the following l_1 -minimization problem [10].

$$\left(l^1\right) : \hat{\boldsymbol{x}}_{(1)} = arg\ min\ \|\boldsymbol{x}\|_1\ \ subject\ to\ \mathbf{A}\boldsymbol{x} = \boldsymbol{y} \tag{10}$$

This problem can be solved in polynomial time by standard linear programming methods.

3.2. **SRC Algorithm.** If we have k classes training samples expressed as $\mathbf{A} = [\boldsymbol{A}_1, \boldsymbol{A}_2, \cdots, \boldsymbol{A}_k] \in \mathbb{R}^{m\times n}$,and a test sample $\boldsymbol{y} \in \mathbb{R}^m$. In general, the optional tolerance constant $\varepsilon > 0$ is given in advance because of the small dense noise in $\boldsymbol{y}$, or to balance the coding error of $\boldsymbol{y}$ and the sparsity of α . We take $\varepsilon = 0.001$ in this paper. The SRC algorithm can be illustrated as follows

Step 1. Normalize the columns of $\mathbf{A}$ to have unit l_2-norm, where $\mathbf{A}$ is the original testing data.

Step 2. Solve the l_1-minimization problem to code $\boldsymbol{y}$ over $\mathbf{A}$ via $\hat{\boldsymbol{x}} = argmin\,\|\boldsymbol{x}\|_1$ $subject\ to\ \|\boldsymbol{y} - \mathbf{A}\boldsymbol{x}\|_2 \le \varepsilon$.

Step 3.Computethe residuals $\boldsymbol{e}_i(\boldsymbol{y}) = \|\boldsymbol{y} - \mathbf{A}\hat{\boldsymbol{x}}_i\|_2$,where $\hat{\boldsymbol{x}}_i$ is the coding coefficient vector associated with the class i.

Step 4. Output the identity of $\boldsymbol{y}$ as $identity(\boldsymbol{y}) = argmin_i\,\{\boldsymbol{e}_i\}$

4. **Face Recognition Based on Multi-wavelet and Sparse Representation.** The flow chart of our face recognition algorithm based on multi-wavelet and sparse representation is shown in Fig. 3.

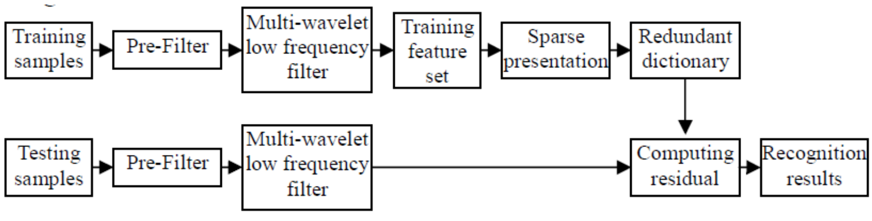

FIGURE 3. The flow chart of face recognition scheme based on multi-wavelet and sparse representation.

Assume that the size of each original face image $\mathbf{I}(i,j)$ is $m \times m$, after pre-filtering, each face image becomes $\tilde{\mathbf{I}}(i,j)$. Let the low frequency multi-wavelet filter be $\boldsymbol{\Psi}_{LL}(i,j)$, and the feature image after low frequency filtering be $\mathbf{M}(i,j)$. Assume that there are

N samples in the training set, $\mathbf{X} = [\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_N]$, $\boldsymbol{x}_i$ is a column vector, representing the features of the ith face. Thus, the proposed algorithm can be described as follows.

Step 1. Each training image $\boldsymbol{x}_i$ is pre-filtered to get $\tilde{\boldsymbol{x}}_i$, then, the feature set of the training set becomes $\tilde{\mathbf{X}} = [\tilde{\boldsymbol{x}}_1, \tilde{\boldsymbol{x}}_2, \cdots, \tilde{\boldsymbol{x}}_N]$.

Step 2. For each pre-filtered image $\tilde{\boldsymbol{x}}_i$, according to Eq. (5), it is filtered by the low frequency multi-wavelet filter as follows

$$\hat{\boldsymbol{x}}_i = \boldsymbol{\Psi}_{LL} \otimes \tilde{\boldsymbol{x}}_i \tag{11}$$

Thus, we can get the corresponding low frequency components of each image to form the feature database $\hat{\mathbf{X}} = [\hat{\boldsymbol{x}}_1, \hat{\boldsymbol{x}}_2, \cdots, \hat{\boldsymbol{x}}_N]$.

Step 3. Based on the feature database $\hat{\mathbf{X}} = [\hat{\boldsymbol{x}}_1, \hat{\boldsymbol{x}}_2, \cdots, \hat{\boldsymbol{x}}_N]$, we establish a fully redundant dictionary D using sparse representation.

Step 4. Accordingly, the input test sample is also per-filtered and then filtered by the low frequency multi-wavelet filter, and then projected onto the fully redundant dictionary. At last, we can get the coefficients of its sparse representation by solving the minimum l_1 norm problem.

Step 5. Compute the residuals $\boldsymbol{e}_i(\boldsymbol{y}) = \|\boldsymbol{y} - \mathbf{A}\hat{\boldsymbol{x}}_i\|_2$ and identify the category of the test sample.

5. **Experimental Results.** To test the superiority of the proposed scheme over the original SRC-based face recognition algorithm, this paper adopts the YALE, AR and FERET face databases to perform the comparison experiments. The YALE face database totally has 165 face images from 15 people, each having 11 face images, and the size of each image is 320×243. This database considers all kinds of conditions, such as center/left/right lightening, with/without glasses, with surprise/sad feeling, with sleepy, blinking or other facial expressions. The AR face database is established by A.M. Martinez and R. Benavent, it has 126 people, including 70 male and 56 female, each having 26 face images. Thus, the number of images in the AR database are 3276, each having the size of 576×768. Similarly, this database also considers various light conditions, with/without the sunglasses or scarf, and different facial expressions. We choose 1400 cropped images to do experiments from the FERET face database, which has 200 people, each having 7 face images with the size of 80×80. For convenience, our experiments adopt the face images of size 64×64 after preprocessing.

5.1. **Experiment on YALE Face Database.** In this experiment, we randomly select 1 to 10 images from each person respectively in the YALE database as the training set, and take the rest as the test set. Then, we perform the SRC, GHM_RR+SRC, GHM_AP+SRC, CL_RR+SRC and CL_AP+SRC face recognition methods respectively on the same set. To better evaluate the experimental results, we perform 5 runs for each experiment, and then calculate the mean and variance of the results. The comparison results are shown in Table 1.

From Table 1, we can see that the recognition rate of the face features which are pre-filtered and transformed by multi-wavelet is higher than the recognition rate by the SRC algorithm. The highest rate of our scheme can reach 97.33%, and it has a lower variance. In addition,, the proposed method has a higher robustness for different lightening directions, obstacle conditions and facial expressions.

5.2. **Experiment on AR Face Database.** In this experiment, we randomly select 1 to 9 images from each person respectively in the AR database as the training set, and take the rest as the test set. Then, we perform the SRC, GHM_RR+SRC and CL_RR+SRC face recognition methods respectively on the same set. To better evaluate the experimental

Table 1. The average recognition rate (ARR) and variance (VR) of YALE face database(%)

Number of samples of each person	SRC		GHM_RR+SRC		GHM_AP+SRC		CL_RR+SRC		CL_AP+SRC	
	ARR	VR	ARR	VR	ARR	VR	ARR	VR	ARR	VR
1	47.47	4.84	48.40	4.94	45.73	2.73	48.27	4.80	**48.53**	5.19
2	72.59	2.77	**74.67**	3.12	**74.67**	2.89	74.52	3.25	73.19	3.37
3	84.00	2.07	84.50	1.92	**85.33**	1.26	83.83	1.83	84.50	2.54
4	85.52	3.26	86.67	4.62	86.67	1.90	**87.23**	2.64	85.71	2.43
5	**89.33**	1.86	88.89	1.76	89.11	1.99	**89.33**	2.17	88.22	0.99
6	91.20	3.21	92.00	2.98	93.07	2.89	**93.60**	1.46	91.73	3.18
7	91.67	3.91	92.33	3.25	91.67	2.64	**93.00**	2.98	92.00	2.74
8	89.78	1.99	**90.67**	1.86	89.33	3.30	**90.67**	1.86	**90.67**	2.90
9	96.00	2.79	95.33	2.98	96.00	2.79	**96.67**	2.36	**96.67**	3.80
10	94.67	5.58	96.00	5.46	96.33	4.71	**97.33**	3.65	96.00	3.65

results, we perform 5 runs for each experiment, and then calculate the mean and variance of the results. The comparison results are shown in Table 2.

From Table 2, we can see that the recognition rate of the face features which are pre-filtered and transformed by multi-wavelet is higher than the recognition rate by the SRC algorithm. The highest rate of our scheme can reach 99.17%, and it has a lower variance. In addition, the proposed method has a higher robustness for different lightening directions, obstacle conditions and facial expressions.

Table 2. The average recognition rate (ARR) and variance (VR) of AR face database(%)

Number of samples of each person	SRC		GHM_RR+SRC		CL_RR+SRC	
	ARR	VR	ARR	VR	ARR	VR
1	38.70	4.58	38.89	3.41	**38.98**	3.20
2	71.25	4.85	71.88	3.38	**73.54**	5.46
3	81.19	1.09	**81.79**	2.58	81.55	2.63
4	90.17	1.34	91.17	0.78	**92.17**	0.68
5	93.17	1.61	**93.83**	0.76	93.68	0.29
6	95.42	2.53	96.04	1.91	**96.25**	1.91
7	97.22	1.73	**97.50**	1.67	**97.50**	1.67
8	97.50	2.34	**98.20**	1.90	98.08	1.43
9	97.50	1.98	**99.17**	1.44	98.33	1.44

5.3. **Experiment on FERET Face Database.** In this experiment, we randomly select 1 to 6 images from each person respectively in the FERET database as the training set, and take the rest as the test set. We perform the SRC, GHM_RR+SRC, GHM_AP+SRC, CL_RR+SRC and CL_AP+SRC face recognition methods respectively on the same set. We perform 5 runs for each experiment too, and then calculate the mean and variance of the results. The comparison results are shown in Table 3.

From Table 3, we can see that the recognition rate of the face features which are pre-filtered and transformed by multi-wavelet is higher than the recognition rate by the SRC algorithm. Especially, the CL_AP +SRC method is the best one at the same condition.

The highest rate of our scheme can reach 78.20%, and it has a lower variance. It proves that the proposed method has a higher robustness for more complex conditions.

Table 3. The average recognition rate (ARR) and variance (VR) of FERET face database(%)

Number of samples of each person	SRC		GHM_RR+SRC		GHM_AP+SRC		CL_RR+SRC		CL_AP+SRC	
	ARR	VR	ARR	VR	ARR	VR	ARR	VR	ARR	VR
1	27.90	0.48	28.97	0.52	29.50	0.94	30.17	0.82	**31.87**	0.65
2	42.24	2.01	42.76	1.87	46.44	1.25	45.80	1.18	**51.20**	1.41
3	48.65	1.63	49.90	1.34	57.25	1.88	55.85	2.13	**61.35**	1.67
4	54.53	2.18	56.80	2.79	65.40	2.72	62.40	2.88	**69.00**	2.47
5	53.40	4.17	58.30	4.40	69.20	3.37	66.10	3.58	**72.30**	1.68
6	59.20	3.27	66.60	1.95	75.60	4.34	74.00	1.22	**78.20**	2.77

6. **Conclusions.** In this paper, we propose a face recognition method based on multi-wavelet and sparse representation. First, we extract the low frequency components of the face features by multi-wavelet which can reduce the dimension of the face database. Then, we adopt the sparse representation classifier to recognize faces. The proposed face recognition algorithm not only brings together the localization characteristics in time and frequency domains, orthogonality, symmetry, compact support and high order disappear moment of the multi-wavelet simultaneously, but also possesses the properties of SR which is insensitive to light and shade. Compared with the traditional SRC-based face recognition method, our method not only has higher recognition rate and better robustness, but also reduces the feature dimension.

Acknowledgment. This work was partially supported by the Zhejiang Provincial Natural Science Foundation of China under grant R1110006. The authors also gratefully acknowledge the helpful comments and suggestions of the reviewers, which have improved the presentation.

REFERENCES

[1] D. L. Donoho, Compressed sensing, *IEEE Trans. Information Theory*, vol. 52, no. 4, pp. 1289-1306, 2006.
[2] R. G. Yang, M. W. Ren, and J. Y. Yang, Sparse representation based face recognition algorithm, *Journal of Computer Science*, vol. 37, no. 9, pp. 267-269, 2010.
[3] X. Y. Zeng, J. Y. Gan, Y. K. Zhai, Face recognition based on fast sparse representation of gabor dictionary and l0 norm, *Journal of Signal Processing*, vol. 29, no.2, pp. 256-261, 2013.
[4] Z. X. Li, Research of Sparse Representation Face Recognition Algorithm Based Image Feature, Yanshan University, pp.39-60, Qinhuangdao, China, 2012.
[5] G. Strang, and V. Strela. Short wavelets and matrix dilation equation, *IEEE Trans. Signal Processing*, vol. 43, no. 1, pp. 108-115, 1995.
[6] D. F. Wu, Multi-wavelet Transform and Its Application in Digital Image Processing, Jilin University, pp. 7-12, Changchun, China, 2005.
[7] J. M. Geronimo, D. P. Hardin, and P. R. Massopust, Fractal functions and wavelet expansions based on several scaling functions, *Journal of Approximation Theory*, vol. 78, no. 3, pp. 373-401, 1994.
[8] C. K. Chui, and J. A. Lian, A study of orthonormal multi-wavelets, *Journal of Applied Numerical Mathematics*, vol. 20, no. 3, pp. 273-298, 1996.
[9] J. Wright, A. Y. Yang, A. Ganesh, and S. S. Sastry, Robust face recognition via sparse representation, *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 31, no. 2, pp. 210-227, 2009.

[10] D. L. Donoho, and M. Elad, Optimal sparse representation in general dictionaries via l^1 Minimization, *Proceedings of the National Academy of Sciences of the United States of America*, vol. 100, no. 5, pp. 2197-2202, 2003.