Journal of Information Hiding and Multimedia Signal Processing ©2015 ISSN 2073-4212
Ubiquitous International Volume **6**, Number **3**, May **2015**

A Step-wise Refinement Algorithm for Face Recognition Based on Blocking Wavelet Transforms

Yi-Bing Wang[1], Tian-Yun Zhou[2] and Bang-Jun Hu[2,*]

[1]Center of Computer Teaching
Anhui University
No.111 Jiulong Road, Hefei, Anhui, 236061, China
wyb@ahu.edu.cn

[2]Key Laboratory of Biotechnology
Anhui Institute of Biology
No.9 Haiguan Road, Hefei, Anhui, 230088, China
jerryzhou@ankebio.com
*hubb1977@163.com

Received July, 2014; revised February, 2015

Abstract. *In this paper, a new human face recognition algorithm is proposed. Blocking wavelet transforms are used for local features extraction, and serial integration of classifiers is adopted for final verification. The proposed technique consists of three stages. 1) A simple and efficient method is presented to extract global features, so as to obtain the mean face image. 2) In order to fully extract local features and overcome the problem of the small sample size, the concept of blocking wavelet transforms is presented. First of all, the image is divided into several non-overlapping blocks, and wavelet coefficients are obtained by using wavelet transform for each sub-block. Then, an improved dimensionality reduction technique, called Bidirectional two Dimensional Principal Component Analysis (B2DPCA) is used to reduce dimensionality. 3) To further improve the recognition performance, the global and local facial features are combined in a serial manner. Global features are used for coarse classification, and the global and local features are integrated for fine classification. Comprehensive experiments on three data sets (ORL, AR and FERET) demonstrate the effectiveness of our scheme.*
Keywords: Blocking wavelet transforms, Stepwise refinement, Classifier integration, A coarse-to-fine strategy, Small sample size

1. **Introduction.** Over the past two decades, numerous face recognition researches and studies have been carried out in the field of Computer Vision (CV) [1]. Owing to the remarkable development in this area, face recognition systems have been applied to a wider range of commercial and law enforcement, such as drivers license, voter registration card, passports, database security, safety management, virtual reality and law enforcement. A small child can identify a human face without effort, but it is difficult to the computer. Therefore, the major aim is to design such systems which can perfectly and completely identify a human face like a small child. Although massive intricacies, the human visual system can efficiently discriminate the human face and eventually finish the face identification. But the small changes in human face such as aging, illumination, cluttered background, viewpoint, and occlusions have seriously affected the performance of face recognition.

Generally, face recognition methods are composed of a feature extractor like Principal Component Analysis (PCA), Wavelet decomposer to reduce the size of input and a classifier like Neural Networks, Support Vector Machines, and Nearest Distance Classifiers to find the features which are most likely to be found. In the field of face recognition, the accuracy and robustness of the algorithm largely depend on the facial feature description [2]. In general, the facial features include global and local features. Global features mainly describe the whole properties of the face, such as color, shape, used for general matching; local features, used for confirming accuracy, mainly describe the details on human face, such as scars, dimples. [3].

In recent years, a large number of global and local features extraction methods have been proposed [4-10], including some of the classical global features extraction methods, such as PCA, Linear Discriminate Analysis (LDA), Locality Preserving Projections (LPP), etc. Particularly, PCA and LDA have become the benchmark for face recognition. However, the insensitivity and robustness to the variations of local features such as illumination, occlusion are lacked when global features are being extracted. In order to extract more local features and overcome the sensitivity to the local variations, a few methods based on local features extraction were proposed. Wang *et al.* [11] proposed three-patch Local Binary Pattern and four-patch Local Binary Pattern features which encode the similarities among neighboring patches around the center pixels so as to obtain the information which can complement the LBP features. Mayank Agarwal *et al.* [12] proposed a system in which they put forward a method for face verification based on information theory approach to coding and decoding the face image. This method consists of two steps. Firstly, the facial features are extracted with PCA, and then the face recognition is realized by the feed forward back propagation Neural Network. In [13] Raman Bhati *et al.* used Eigen faces to represent the feature vectors. This approach generates more proficient results in all aspects including mean square error, recognition accuracy and training time.

However there are problems regarding variations in different postures on an individual's image, so Pushpaja V. Saudagare *et al.* [14] proposed a method which involves inhibiting the consequences of changing illumination and pose. They proposed a 2-level wavelet transform system that is used to divide the facade image into seven sub-image bands. Then from these sub-image bands, eigenfaces features are extracted and these features are further classified. In [15], Ki Chung *et al.* referred to a new face recognition method based on PCA and Gabor filter responses. In the first step, Gabor filtering is applied on predefined fiducial points to get robust facial features from the original face image. The second step involves transformation of the facial features into Eigen space by PCA, leading to the optimal classification of individual facial representations.

Some of the above methods are widely used for face recognition. It is worth mentioning that wavelet transform method and the sub-image method are of great importance. The sub-image method is to divide the original image into several non-overlapping blocks, from which features are extracted. Many researchers pay more attention to the sub-image method, which is simple and intuitive, and also can effectively solve the problem of small sample size, improving the robustness in variations such as illumination, occlusion etc. The wavelet transform has achieved good results in performance evaluation on many face databases, so it is considered to be one of the most effective ways of face representing. Further, blocking wavelet transforms can not only give prominence to the local information, but also realize dimensionality reduction rapidly. Sun *et al.* [16] proposed a face recognition algorithm in which face image is divided into several blocks, and then for each sub-block PCA is used to extract features, which is called blocking PCA. Similarly, Xie *et al.* [17] proposed using wavelet transform for each sub-block to get high frequency and low frequency wavelet coefficients in different sub-block first, and then use them for feature extraction.

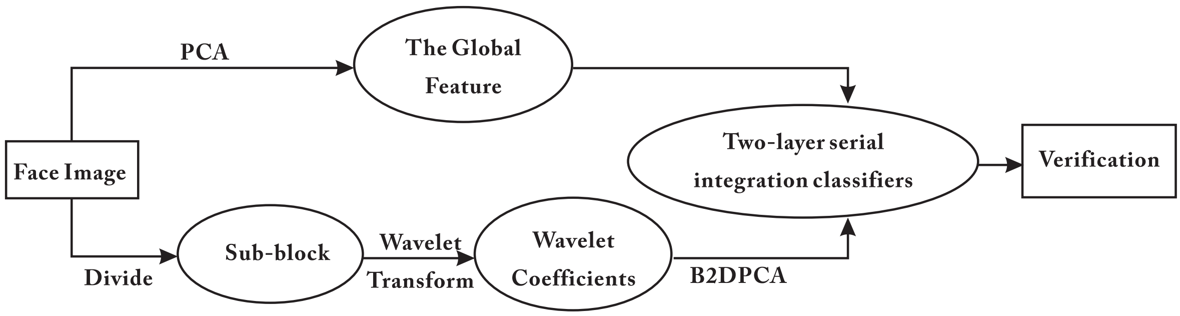


Figure 1. The system diagram

In addition to the feature extraction, classifier also has a great influence on the performance of face recognition algorithm. In theory, almost all of the classifiers can be used to classify face images. However, currently, classifiers based on statistical learning, is the first choice for face recognition. The statistical learning theory has made great achievement in generalization ability of learning algorithm, but practically it is not the most effective approach in face recognition, so comes another key issue in the study of face recognition that is to design classifier with high accuracy and improve its generalization ability.

K-Nearest Neighbor (KNN) is a simple and effective classification method. Wang *et al.* [18] proposed a method which contains two parts. Firstly, the image is divided into multiple non-overlapping blocks. After using Kernel principle component analysis to extract features, KNN classifier is used to finish classification and recognition. In [19] Ergun Gumus *et al.* proposed a method based on wavelet transform and Support Vector Machine (SVM). They extract features by using wavelet decomposition and Eigenfaces method which is based on PCA, and after feature vectors are generated, SVM completes the final classification. KNN has high efficiency without supervision, while SVM has very good effect on classifying in few categories. Therefore, some researchers combine these two methods for classification.

Based on the above analysis, considering both feature extraction and classifier design, we propose a step-wise refinement algorithm for face recognition based on blocking wavelet transforms. In this method, global features are extracted by PCA, and local features are extracted by the blocking method after wavelet decomposition of each sub-block. After that, a two-layer classifier with a global classifier and multiple local classifiers is built, and then the global and local features are combined in a serial manner. The experimental results show that the proposed method can not only significantly improve the recognition accuracy, but also can greatly reduce the time needed for the recognition.

The remaining part of this paper is divided into three sections. Section 2 discusses the proposed face recognition system, including how to extract global features, how to use blocking wavelet transforms, and how to build an efficient classifier, etc. Section 3 reports experimental results obtained on ORL, AR and FERET databases and compares them with the state-of-the-art face recognition methods. Finally, conclusions are drawn in Section 4.

2. SYSTEM DESIGN.

In order to improve the performance of face recognition system, this paper proposes the face recognition system diagram shown in Figure 1. First, use PCA to extract global features of the face image, then divide the face image into several non-overlapping blocks, and use wavelet transform for each sub-block to obtain the wavelet coefficients. After that, local features are extracted by using bidirectional two dimensional principal component analysis (B2DPCA) to reduce dimensionality rapidly. Finally, the exacted features are eventually classified and recognized by two-layer serial integration classifiers with KNN-SVM.

2.1. **Global features extraction by Principal Component Analysis.** There are many kinds of methods to extract global features such as PCA, LDA and LPP etc. PCA is a statistical analysis technology on the basis of Karthunen-Loeve transform. The purpose of PCA is to reduce the dimensionality of the face data from R dimensional space to M dimensional space, where $R >> M$. The main information is preserved after dimensionality reduction, so that classification for face image becomes easier. By using PCA, the global information of face, such as organs location, color, etc. is very clear, but the details are blurred. According to the above, in the proposed method, PCA is used to extract global information of face image.

2.2. **Image division and local features extraction.** The most common ways to divide the face image are organ division method and block division method. The former is to make division according to the facial organs like eyes, nose, mouth etc, which has some limitations in actual operation because of the accuracy of facial feature detection or some manual intervention needed, while in the latter method, the face image is divided into several blocks with the help of specific coordinate, so that much better identification effect can be achieved. Therefore in this paper, it is block division method that we adopt.

Wavelet transforms are multi-resolution image decomposition tools that provide a variety of channels representing the image feature by different frequency subbands at multi-scale. It has the well-known advantage of the capabilities of multi-scale analysis, which is beneficial to the extraction of key facial features in coarse-to-fine order. In this paper, blocking wavelet transforms are used to extract local features, and to get the detail information of face image. The flow diagram of the proposed blocking wavelet transforms is shown in Figure 2.

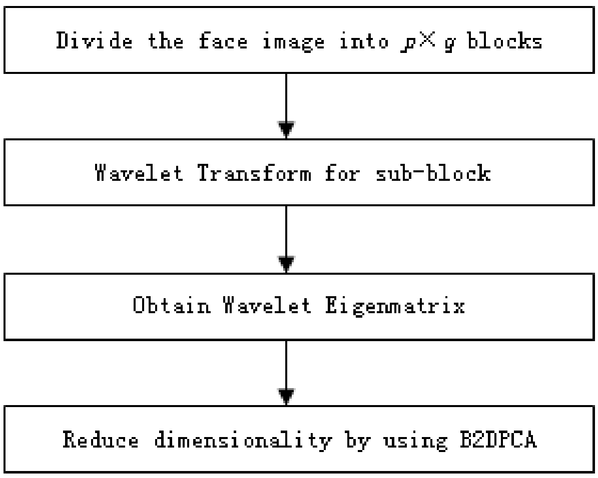


FIGURE 2. The flow diagram of blocking wavelet transforms

This section firstly discusses how to choose a proper wavelet basis and decomposition levels, then concludes the rule of image blocking, and finally conducts the detailed steps of blocking wavelet transforms.

A. Wavelet basis choosing and decomposition levels confirmation

Experiments have been conducted for many times to compare the identification performance via different wavelet basis on the ORL face database from the aspects of the training time and recognition accuracy. In the experiments, different wavelet bases are used to obtain low frequency sub-image for recognition by 2-level wavelet decomposition. The results show that each wavelet series has wavelet basis with high recognition accuracy and short training time, and there is not much difference between them. The experimental results are shown in Table 1, from which we can see that the Rbio1 has the highest recognition accuracy, but with the longest training time. From the two different sizes of images, the Coif1 has the best performance in the comprehensive consideration of recognition accuracy and training time, so in this paper, Coif1 is chosen as the wavelet basis.

Table 1. Performance of different wavelet basis

Wavelet basis	Size is 28×23		Size is 112×92	
	Time(ms)	Mean accuracy(%)	Time(ms)	Mean accuracy(%)
Ribio1.3	1.59	94.33	5.07	96.56
Db3	1.37	94.67	4.35	95.63
Sym3	1.28	94.75	4.42	95.87
Coif1	1.46	95.03	4.46	96.09
Bior1.3	1.53	95.12	4.72	95.94

The level of decomposition relies on the size of the image. In the experiments, images with different sizes (112×92 and 28×23) are chosen from the ORL face database. The recognition accuracy of different size images by using different decomposition levels is shown in Table 2.

Table 2. Recognition accuracy by different decomposition levels

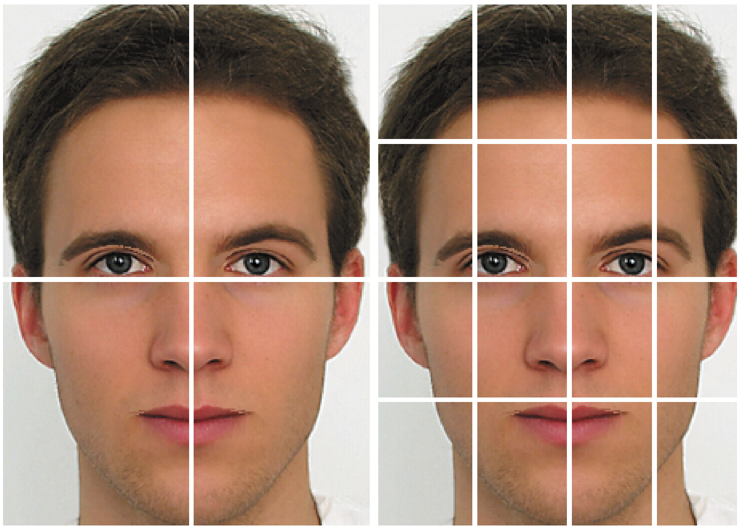
The level of Decomposition	Size is 28×23	Size is 112×92
1	94.46	93.87
2	95.02	96.11
3	91.49	96.57
4	81.63	94.83

Table 2 shows that, it can reach the best result when using 3-level wavelet decomposition to images with the size of 112×92, and 2-level wavelet decomposition to images with the size of 28×23. Therefore, the images with the size of 112×92 and 3-level wavelet decomposition are taken in the following experiments.

B. Division method for face image

If the face image is divided into 2×2 sub-block, shown in Figure 3 (a), four respective non-overlapping blocks, upper, lower, left, right, the forehead, eyebrows and eyes are in the same block. If it is divided into 4×4 sub-block, forehead is with more sub-blocks, cheeks and mouth are separated, as shown in Figure 3 (b). It is supposed that the recognition effect will not be significantly affected when illumination or occlusions exist in the face when the image is divided into more sub-blocks.

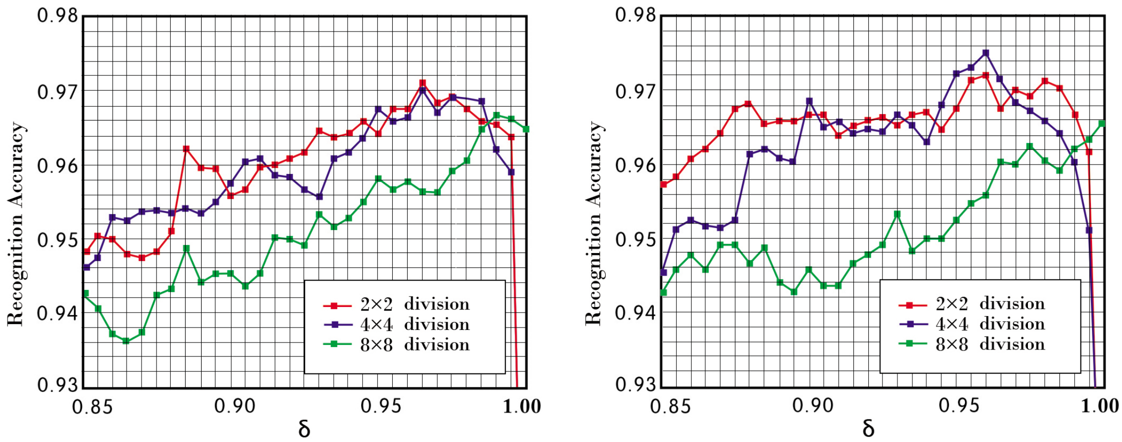
In order to find a suitable division method, experiments are performed by using 2-level and 3-level wavelet decomposition respectively for each sub-block. The recognition



(a) divided into 2×2 blocks (b) divided into 4×4 blocks

Figure 3. different division methods for face image

accuracy is shown in Figure 4. Figure 4 (a) and (b) respectively express the recognition accuracy by using 2-level and 3-level wavelet decomposition. In Figure 4, δ represents the energy ratio of previous d eigenvalue in PCA.



(a) by 2-level wavelet decomposition (b) by 3-level wavelet decomposition

Figure 4. Recognition accuracy by different wavelet decomposition levels

The Figure 4 shows that the recognition accuracy almost remains at more than 96% when the image is divided into 2×2 sub-block, while it declines but not increases when into 8×8 sub-block. Besides, it is obvious that the more sub-blocks are divided, the longer training time is needed. Therefore, in this paper, the face image is divided into 2×2 sub-block.

C. Dimensionality reduction by using B2DPCA

Let X denote a q dimensional unitary column vector. Linear transformation $Y = AX$ is applied to projecting a $p \times q$ image matrix A to X, resulting into a projected vector Y. The total scatter of the projected samples can be used to test the discriminatory power of X. If S_x represent the covariance matrix of the projected feature vectors, it can be defined as:

$$S_x = E[Y - E(Y)][Y - E(Y)]^T = E[(A - EA)X][(A - EA)X]^T \tag{1}$$

The total scatter is presented by the trace of S_x; $J(X) = tr(S_x)$, where the trace of S_x can be represented as:

$$tr(S_X) = X^T[E(A - EA)^T(A - EA)]X \tag{2}$$

The covariance matrix G_t of image can be defined:

$$G_t = E[(A - EA)^T(A - EA)] \tag{3}$$

It can be proved from the above that G_t is a non-negative $q \times q$ image covariance matrix. Provided there are N training samples, $\overline{A}$ represents the average image of N training samples, and the αth image sample is denoted by a $p \times q$ matrix $A_\alpha(\alpha = 1, 2, \ldots, N)$. G_t can be calculated as follows:

$$G_t = \frac{1}{N}\sum_{\alpha=1}^{N}(A_\alpha - \overline{A})^T(A_\alpha - \overline{A}) \tag{4}$$

$$J(X) = X^T G_t X \tag{5}$$

The unitary vector X_{opt} that makes the generalized total scatter criterion $J(X)$ maximized is called the optimal projection axes. Generally speaking, X_{opt} can refer to a list of N orthonormal eigenvectors $X_1, X_2, \ldots, X_N$ of G_t correspondent to N largest eigenvalues. Thus, every image A_α can achieve the reduction in dimensionality by being post-multiplied and pre-multiplied with optimal projection axes as $X_{opt}^T A_\alpha X_{opt}$.

One problem with 2DPCA is that its operability only goes along row direction, which affects the speed of classification. Based on an assumption that all of the training images are zero mean, Zhang *et al.* [20] proposed B2DPCA, in which image covariance matrix can be obtained by using outer product of row/column vectors. Two image covariance matrices G_{tRow} and G_{tCol} can be figured out by Eq. (4) firstly following the row vectors of A_α and $\overline{A}$, and repeating the same procedures for column vectors. The optimal projection axes of G_{tRow} and G_{tCol} are calculated and signed as X_{1opt} and Z_{1opt}. Since both G_t and G_{tRow} are calculated along rows, the optimal projection axes X_{opt} and X_{1opt} are similar. A_α is formed as $Z_{1opt}^T A_\alpha X_{1opt}$ via dimensionality reduction.

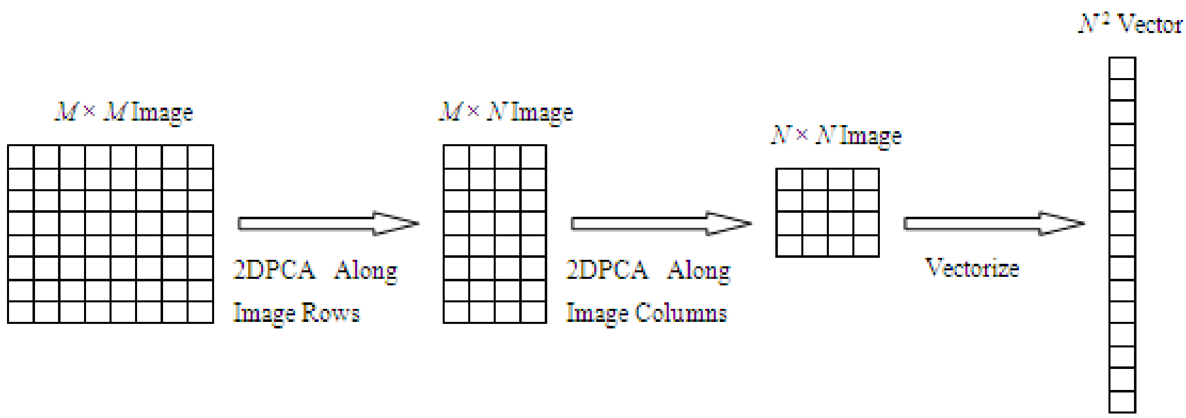


FIGURE 5. Block diagram of proposed B2DPCA

The proposed dimensionality reduction method runs respectively along row and column directions which are shown in Figure 5 so that the neighborhood relationship is better preserved and the generation of distinctive feature sets can be assured. The purpose of B2DPCA is to generate an image covariance matrix $G_{t\alpha}$ by optimal projection axes and further optimizes it. As long as $X_{opt\alpha}$ is worked out, the dimensionality of A_α is reduced along its columns, bringing about new image sets A_β by using Eq. (6). The new image sets are then taken as a fresh database to figure out the new image covariance matrix $G_{t\beta}$ and optimal projection axes $X_{opt\beta}$. Finally, A_β is pre-multiplied by $X^T_{opt\beta}$ via Eq. (7). Hence, instead of using the traditional 2DPCA, a two fold method is taken in B2DPCA to realize dimensionality reduction. The output matrix A_θ is obtained by multiplying A_β with $X^T_{opt\beta}$.

$$A_\beta = A_\alpha X_{opt\alpha} \tag{6}$$

$$A_\theta = X^T_{opt\beta} A_\beta \tag{7}$$

The steps are offered as below:
INPUT: Input sub-image $A_{M\times M}$
OUTPUT: $N \times N$ output matrix A_θ
(1) Use Eq. (4) to calculate non-negative image covariance matrix $G_{t\alpha}$
(2) Calculate the total scatter $J(X)$ by Eq. (5)
(3) $X_{opt\alpha} = X_1, X_2, \dots, X_N$, where X_i represents a orthogonal vector
(4) Use Eq. (6) to realize dimensionality reduction along columns
(5) Calculate $G_{t\beta}$, i.e., the image covariance matrix of A_β by $G_{t\beta} = \frac{1}{N}\sum_{\alpha=1}^{N}(A_\beta - \overline{A_\beta})^T(A_\beta - \overline{A_\beta})$
(6) According to step(2), calculate $X_{opt\beta}$
(7) Reduce dimensionality along rows by Eq. (7)

In conclusion, the local feature extraction steps by blocking wavelet transforms are as follows:
Step 1: The face image is divided into 2×2 sub-block;
Step 2: Choose Coif1 as the wavelet basis and use wavelet transforms for each sub-block. Because of the main energy concentrated in the low frequency range after wavelet decomposition, therefore, in this paper, only the low frequency components of each sub-block are extracted;
Step 3: Set the image size as $M\times N$, and divide it into $n\times n$ (in this paper, $n = 2$). Mark the size for $\frac{M}{n} \times \frac{N}{n}$ sub area as $I_{a,b}$, so the mean wavelet coefficients can be obtained by Eq. (8).

$$\overline{w}_{a,b} = \frac{1}{n^2}\sum_{I_{a,b}} I(x,y), 1 < a < n, 1 < b < n \tag{8}$$

Therefore, a new matrix W composed of $\overline{w}_{a,b}$ is the blocking wavelet eigenmatrix:

$$W = \begin{pmatrix} \overline{w}_{1,1} & \overline{w}_{1,2} & \dots & \overline{w}_{1,n} \\ \overline{w}_{2,1} & \overline{w}_{2,2} & \dots & \overline{w}_{2,n} \\ \vdots & \vdots & \dots & \vdots \\ \overline{w}_{n,1} & \overline{w}_{n,2} & \dots & \overline{w}_{n,n} \end{pmatrix} \tag{9}$$

Step 4: B2DPCA is used to reduce dimensionality of the wavelet coefficients of each sub-block, so as to obtain eigenmatrix B of each sub-block.

2.3. **Setting of classifier with a coarse-to-fine strategy.** This section firstly describes how to build the correspondent classifier by using the global eigenvectors and the wavelet eigenvectors of each sub-block, and then make serial integration of these classifiers with a coarse-to-fine strategy.

Global features mainly describe the whole property of human face, used for rough matching. Local features describe the details of human face, used for accurate confirmation. Therefore, in order to improve the classification speed and recognition accuracy, a two-layer classifier is built. The first layer with the global classifier is used for rough matching. In the second layer, the global classifier and N local component classifiers are integrated (called the whole classifier) for accurate classification. The existing literature by Giorgio Fumera *et al.* [21] proved that, the performances are almost the same between using simple summation and using weighted summation to integrate the component classifiers when there are many of component classifiers. Therefore, in this paper, the simple summation is used to integrate multiple component classifiers.

Because fewer features are used, the global classifier in the first layer is with fast speed and low accuracy, and because of more features used in the whole classifier composed of $N+1$ component classifiers, the second layer is with high accuracy and low speed. That is, the global classifier used in the first layer speeds up the classification, and the local classifier added in the second layer improves the recognition accuracy.

Firstly, with the method of the K-nearest neighbor (KNN), the global classifier used in the first layer filters data from the training samples by comparing the eigenvectors. Then with SVM, the whole classifier used in the second layer further recognizes the face image from the remaining samples. Thus, the whole classifier with the limitation of slow speed only needs to handle the small subset of the original face database, so it can obviously shorten the classification time. Figure 6 shows a schematic diagram of serial integration classifier, here FD_0 represents the original face database, and FD_1 represents the post-filtering data via global classifier in the first layer.

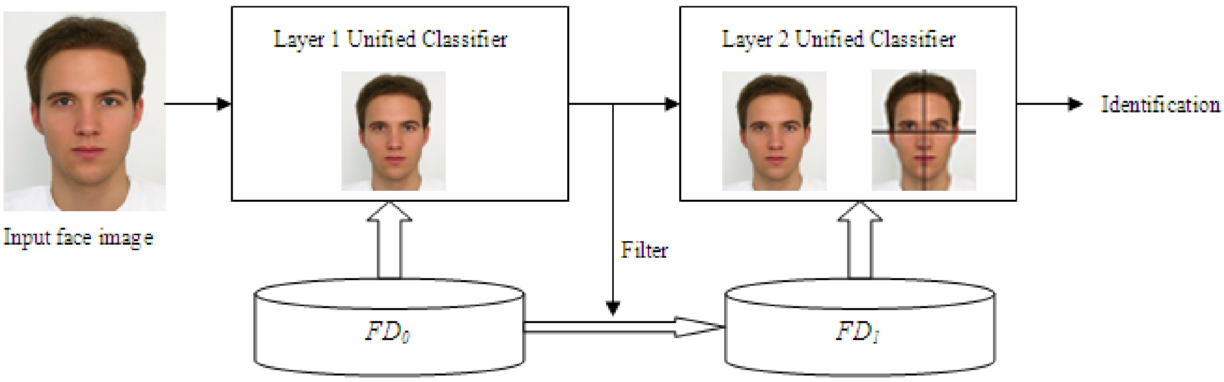


FIGURE 6. Serial integration of classifiers for face recognition

3. **EXPERIMENT PROCESS.** In this section, the method proposed in this paper are compared with some other advanced methods based on the two face databases: ORL and AR, and the performance of serial integration of classifiers on the FERET database is analyzed. The experiments are conducted on a PC with Intel Pentium D 3.10 GHz CPU and 8GB RAM.

3.1. **Face Databases.** The ORL face database contains 10 different images for each of the 40 distinctive subjects. Each subject is imaged at different time and different illumination. The size of each image is 112×92 with 256 gray-level.

The AR face database established by Computer Vision Center (Barcelona, Spain), has nearly 4000 images, including 126 persons (male 70, female 56). In this face database, each person's facial expressions are not the same, with almost all kinds of skin color. The reason for selecting AR face database is that it is one of the few databases that contains the occlusions. Figure 7 lists all the 26 images of one person in AR database.

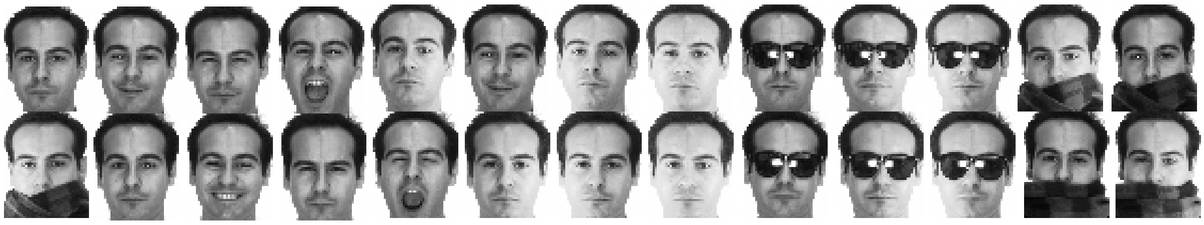


Figure 7. All the 26 images of one person in AR database

The FERET database sponsored by the Department of Defense is to develop a system with automatic face recognition capability to be employed for assistance in intelligence, security and law enforcement. It contains 1208 face images with different expression, direction and illumination, in which the number of samples is about 10000.

3.2. **Experiments on ORL database.** Comparison experiments are performed on ORL face database. Randomly 10 images for per person are selected, of which 5 images are as the training set and the other 5 images as testing set. Size of each image is cropped to 64×64 in the repeated experiments for 10 times by different decomposition levels and different division methods. The experimental results are shown in Table 3.

Table 3. Recognition accuracy by different decomposition levels and different division methods on ORL database

Division methods	Decomposition levels							
	1-level		2-level		3-level		4-level	
	Mean accuracy (%)	Time (ms)	Mean accuracy (%)	Time (ms)	Mean accuracy (%)	Time (ms)	Mean accuracy (%)	Time (ms)
1×1 (without division)	94.73	145.60	96.01	143.20	95.38	138.60	95.21	1450
2×2	95.49	21.60	96.01	19.50	96.59	23.23	95.67	22.20
4×4	96.38	6.80	96.65	8.60	95.77	9.60	95.31	9.80
8×8	96.03	4.90	96.09	5.70	94.96	5.80	95.47	5.60

The obtained accuracy from the first row in Table 3 (without division) are always lower than that of post-division, which can prove that much higher recognition accuracy and much faster speed can be obtained when using blocking wavelet transforms.

3.3. **Experiments on AR database.** To evaluate the performance of the proposed method on the occluded AR database, we choose 100 samples. For each sample, we respectively select K images ($K = 1, 2, 4, 6, 8$) without any occlusions as the training set, and two subsets with sunglasses (occlusion percentage is about 30%) and scarf occlusions (occlusion percentage is about 50%) as the testing set. In the same way, all the images are cropped to 64×64.

All the results of the recognition accuracy by the proposed method in this paper, compared to "Discriminative K-SVD for dictionary learning" [22], "A new face coding model, namely regularized robust coding (RRC)" [23], "Alignment-Free Approach" [24] and "Discriminant sparse neighborhood preserving embedding (DSNPE)" [25], are shown in Figure 8 for each different value of $K(K = 1, 2, 4, 6, 8)$. Figure 8(a) and Figure 8(b) show the recognition accuracy with sunglasses and scarf occlusions respectively.

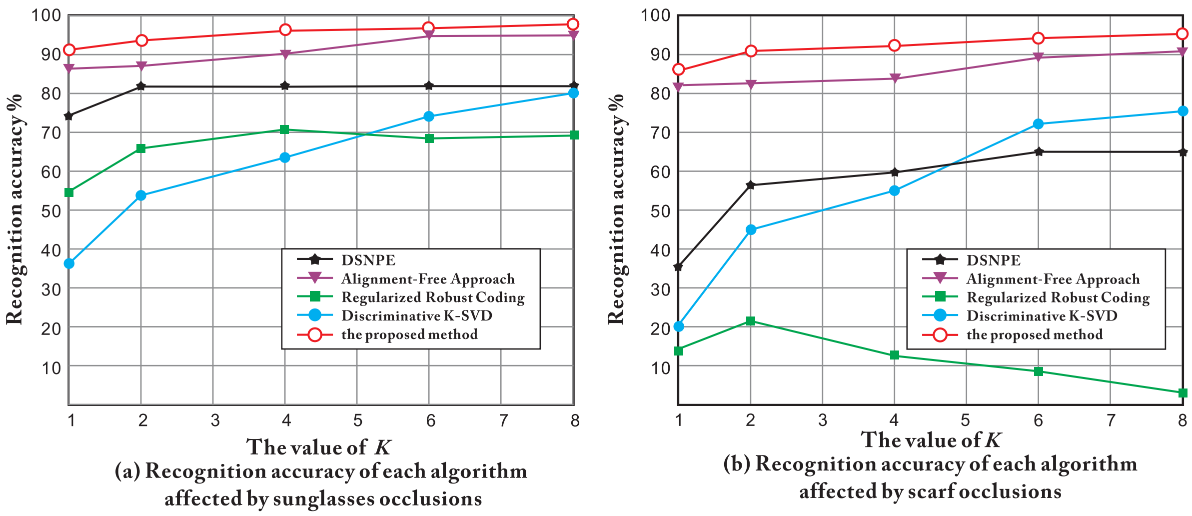


FIGURE 8. Recognition accuracy with different value of K on the AR database

From Figure 8, The implementation effect of both Alignment-Free Approach [24] and the proposed method are relatively better than that of others. When only one training image per person is available ($K = 1$), the best recognition accuracy still can be obtained for all different occlusions by the proposed method. From comprehensive analysis of the 2 figures, no matter what the value of K is, our method is better than others.

3.4. **Performance analysis for serial integration of classifiers.** In this experiment, the adopted training set is a subset of standard training set from FERET database, which contains 1002 frontal images of 429 persons, and the testing set includes four standard testing sets from the FERET database: fafb, fafc, dup1, dup2. Expression variations are mainly in the fafb set, illumination variations in the fafc set, and acquisition time variations in the dup1 set and dup2 set. In this experiment, the recognition accuracy is obtained in four testing sets by the global classifier, the local classifiers and the whole classifier.

Table 4 shows the recognition accuracy on each testing set for FERET database by the global classifier, local classifier and the whole classifier.

From Table 4, the whole classifier by serial integration can improve the performance on dup1 and dup2 testing sets. On fafb and fafc testing sets, the recognition accuracy by the local classifiers is close to 100%, that is to say, the whole classifier does not improve its performance. The proposed method achieve better results than that of any other compared algorithm on the four standard testing sets of FERET database.

TABLE 4. Performances comparison on FERET database

Methods	fafb	fafc	dup1	dup2
Mult-Region Histogram[26]	99.36	98.71	79.63	75.67
Combined b/g samples[27]	98.91	98.69	80.74	77.39
Alignment-Free Approach[24]	99.62	99.33	81.46	78.57
Our methods-the Global classifier	97.63	75.63	66.57	39.94
Our methods-the Local classifiers	99.54	99.01	91.83	84.29
Our methods-the Whole classifier	99.61	99.03	93.79	88.13

3.5. **Performance comparison.** The superiority of the proposed method is not be defined fully just by the high accuracy. Therefore, the running time of other algorithms, including the time of training in total and the time of testing one image, is recorded to be compared with that of the proposed method. The comparison results are shown in Table 5.

TABLE 5. Running time by other algorithms on ORL database

Methods	Time of training in total (s)	Time of testing one image (s)
Mult-Region Histogram[26]	167.983	3.733
Combined b/g samples[27]	184.039	4.873
Associate-Predict[28]	130.298	4.192
Discriminative K-SVD[22]	102.067	4.609
Alignment-Free Approach[24]	98.680	3.992
DSNPE[25]	99.187	3.685
Our method	98.697	2.891

Table 5 shows that, the time of training in total of the proposed method is only slightly longer than that of the Alignment-Free Approach [24], just only 0.017 seconds longer, but much shorter than those of several other algorithms. It only needs 2.891 seconds to test one image, obviously shorter than that of the each compared algorithm, which significantly proves the superiority of the proposed method.

4. **CONCLUSIONS.** In this paper, a stepwise refinement algorithm based on blocking wavelet transforms for face recognition is proposed. Firstly, global features of the face image are extracted by PCA, and the face image is divided into 2×2 sub-block. Secondly, local features are extracted by using wavelet transform for each sub-block and using B2DPCA to further reduce dimensionality. Finally, a two-layer classifier is built, which combines the global and local facial features in serial manner with a coarse-to-fine strategy. Many experiments on the three face data sets demonstrate that our scheme outperforms most of the other state-of-the-art methods.

In the future, the proposed method will be applied on other much larger databases to further improve the performance together with other advanced technology based on plenty of experiments.

Acknowledgment. The authors would like to thank the Associate Editor and the Reviewers for their valuable comments and suggestions.

REFERENCES

[1] M. Srinivasan and N. Ravichandran, A 2D discrete wavelet transform based 7-state hidden Markov model for efficient face recognition, *The 4th International Conference on Intelligent Systems Modelling & Simulation*, pp. 199-203, 2013.
[2] Y. Su, S. G. Shan, X. Chen and W. Gao, Interation of Global and Local Feature for Face Recognition, *Journal of Software*, vol. 21, no. 8, pp. 1849-1856, 2010.
[3] Y. Zhu and S. Can, Sub-Image Method Based on Features Sampling and Feature Fusion for Face Recognition, *Journal of Software*, vol. 23, no. 12, pp.3210-3220, 2012.
[4] X. Tan and B Triggs, Enhanced local texture feature sets for face recognition under difficult lighting conditions, *IEEE Trans. on Image Processing*, vol. 19, no. 6, pp. 1635-1650, 2010.
[5] W. Gu, C. Xiang, YV. Venkatesh, D. Huang and H. Lin, Facial expression recognition using radial encoding of local Gabor features and classifier synthesis, *Pattern Recognition*, vol. 45, no. 1, pp. 80-91, 2012.
[6] C. Geng and X. Jiang, Fully automatic face recognition framework based on local and global features, *Machine vision and applications*, vol. 24, no. 3, pp.537-549, 2013.
[7] M. Yang, L. Zhang, SCK Shiu and D. Zhang, Gabor feature based robust representation and classification for face recognition with Gabor occlusion dictionary, *Pattern Recognition*, vol. 46, no. 7, pp. 1865-1878, 2013.
[8] X. Zhu and D.Ramanan, Face detection, pose estimation, and landmark localization in the wild, *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 2879-2886, 2012.
[9] K. Bonnen, B.F. Klare and A.K. Jain, Component-based representation in automated face recognition, *IEEE Trans. Information Forensics and Security*, vol. 8, no. 1, pp. 239-253, 2013.
[10] C.F. Shan, Learning local binary patterns for gender classification on real-world face images, *Pattern Recognition Letter*, vol. 33, no. 4, pp. 431-437, 2012.
[11] X. Wang, Y. Zhang and X. Mu, The Face Recognition Algorithm Based on Improved LBP, *Opto-Electronic Engineering*, vol. 39, no. 7, pp. 109-114, 2012.
[12] M. Agarwal, H. Agawal and N. Jain, Face Recognition Using Principal Component Analysis, Eigenface and Neural Network, *IEEE International Conference on Signal Acquisition and Processing*, pp. 310-314, 2010.
[13] R. Bhati, S. Jain and N. Maltare, A Comparative Analysis of Different Neural Networks for Face Recognition Using Principal Component Analysis, Wavelets and Efficient Variable Learning Rate, *IEEE International Conference on Computer and Communication Technology*, pp. 526-531, 2010
[14] S.A. Daramola and O.S. Odeghe, Efficient Face Recognition System using Artificial Neural Network, *International Journal of Computer Applications*, vol. 41, no. 3, pp. 12-15, 2012.
[15] P. Agarwal and N. Prakash, An Efficient Back Propagation Neural Network Based Face Recognition System Using Haar Wavelet Transform and PCA, *Journal of Computer Science and Mobile*, vol. 2 no. 5, pp. 386-395, 2013.
[16] X. Sun and B. Liu, Face Recognition Based on Block-PCA, *Computer Engineering And Applications*, vol. 41, no. 17, pp. 79-82, 2005.

[17] Y. Xie, F. Chen, S. Zhang and J. Yang, Feature Extraction And Recognition Of Human Face Based On Intersected Wavelet Transform And SVD Threshold Compression, *Computer Applications And Software*, vol. 25, no. 1, pp. 30-34, 2008.

[18] Y.B. Wang and B.J. Hu. Robust face recognition based on low frequency DCT coefficients retransforming optimized by CLAHE, *Computer Engineering and Applications*, vol. 50, no. 9, pp. 135-140, 2014.

[19] E. Gumus and N. Kilic, A Sertbas, Evaluation of face recognition techniques using PCA, wavelets and SVM, *Expert System with Applications*, vol. 37, no. 9, pp. 6404-6408, 2010.

[20] D. Zhang and Z. Zhou, (2D)2PCA: two-directional two-dimensional PCA for efficient face representation and recognition, *Elservier Neurocomputing*, vol. 69, no. 1, pp. 224-231, 2005.

[21] G. Fumera and F. Roli, A theoretical and experimental analysis of linear combiners for multiple classifier systems, *IEEE Trans. On Pattern Analysis and Machine Intelligence*, vol. 27, no. 6, pp. 942-956, 2005.

[22] Q. Zhang, B. Li, Discriminative K-SVD for Dictionary Learning in Face Recognition, *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 2691-2698, 2010.

[23] M. Yang, D. Zhang,J Yang and D. Zhang, Regularized Robust Sparse Coding for Face Recognition, *IEEE Conference on Computer Vision and Pattern Recognition*, pp.625-632, 2011.

[24] S. Liao, A.K. Jain and S. Li, Partial Face Recognition: Alignment-Free Approach, *Pattern Analysis and Machine Intelligence*, vol. 35, no. 6, pp. 1193-1205, 2013.

[25] J. Gui, Z. Sun and W. Jia, Discriminant sparse neighborhood preserving embedding for face recognition, *Pattern Recognition*, vol. 45, no. 8, pp. 2884-2893, 2012.

[26] C. Sanderson and B. C. Lovell, *Multi-region probabilistic histograms for robust and scalable identity inference*, LNCS 5558, Springer, pp. 199-208, 2009.

[27] L. Wolf, T. Hassner and Y. Taigman, *Similarity scores based on background samples*, LNCS 5995, Springer, pp. 88-97, 2010.

[28] Q. Yin, X. Tang, and J. Sun, An associate-predict model for face recognition, *The 6th IEEE Conference Computer Vision and Pattern Recognition*, pp. 497-504, 2011.