Journal of Information Hiding and Multimedia Signal Processing ©2013 ISSN 2073-4212
Ubiquitous International Volume **4**, Number **2**, April **2013**

A Scale Invariant Feature Transform Based Method

Yu-Yao Wang
Harbin Institute of Technology Shenzhen Graduate School
shinewyy@gmail.com

Zheng-Ming Li
Harbin Institute of Technology Shenzhen Graduate School
lizhengming2004@126.com

Long Wang
Harbin Institute of Technology Shenzhen Graduate School
patrick8863@gmail.com

Min Wang
Harbin Institute of Technology Shenzhen Graduate School
leafcano@gmail.com

Received September 2012; revised April 2013

Abstract. *In this paper, we propose a novel method for face recognition. Its basic idea is to use a coarse to fine strategy based on Scale Invariant Feature Transform (SIFT) feature. To recognize a test sample, our method contains the three main steps: The first step identifies a certain number of candidates from training samples, depending on the Euclidean distance between the test sample and all training samples. The second step counts the numbers of well matched pairs of SIFT features between each candidate chosen in the first step and the test sample, then chooses several samples from candidates with the greater number of well matched pairs. The third step calculates the similarity between the test sample and each class that includes training samples chosen in second step, and chooses a class with the highest similarity to be the recognition result. By the first two steps, our method succeeds in greatly reducing the computational complexity, and avoiding the interference of those samples that may cause error recognition to a certain extent. The third step enhances the robustness of our method. Extensive experiments on different public face databases confirm that our method obtains high recognition accuracy and has a good robustness.*
Keywords: pattern recognition, face recognition, SIFT features.

1. **Introduction.** In the present society, recognition technology on biometrics, such as palmprint, iris, and face, has received much attention [1-6]. As one of the most representative biometrics techniques, face recognition has undergone considerable development in recent years [7-15].

Until now, many methods of face recognition have been proposed. For example, Eigenface [16], Independent component analysis (ICA) [17] and Fisher Discriminate Analysis [18] are the classic holistic face recognition methods which are simple and effective [19]. Although the holistic feature can measure the entire characteristic of an image, it cannot avoid losing some details within an image. Many recent researches show that local features are more effective to describe the detailed and stable information of an image.

Among these local features, Scale Invariant Feature Transform (SIFT) proposed by D. Lowe [20] became popular in face recognition. One interesting thing about the SIFT is its capability to capture the main gray level features of an objects view by means of local patterns extracted from a scale-space decomposition of an image [21]. Luo et al. [22] showed the good representation ability of SIFT features through combining the person-specific SIFT features and a simple matching strategy. So far, researchers have made a lot of researches on SIFT, including some improvements [23-26]. Mikolajczyk and Schmid [27] identified the SIFT as the most resistant to common image deformations. Furthermore, they applied SIFT to face recognition and compared it with PCA, 2DPCA, and so on [28], which proved SIFT to be a powerful matching tool. B. Dai et al. [29] kept all initial SIFT keypoints as features and detected the keypoints described by a partial descriptor at a large scale. This method performed well on ORL and AR face databases by using such kind of features. Ajmal Mian [30] also used SIFT feature for video-based face recognition, and achieved high recognition and verification accuracy.

Although these SIFT based methods have a good performance, they pay too much attention on local features, leading the overall information of an image ignored. However, the linear representation methods [31-39], which used the whole image for representation, showed the powerful ability in classification. Xu et al. [40] proposed a method called Two-Phase Test Sample Sparse Representation (TPTSR). This method calculated a coefficient matrix using all training samples to represent the test sample, and chose a certain number of neighbors to classify the test sample. This method was effective on many face databases; however, it cost a large amount of time to calculate the coefficient matrix in practice.

Inspired by [30] and [40], we propose our method based on SIFT for face recognition. In our method, we make use of Euclidean distance to choose a certain number of candidates, and apply SIFT features extracted from those candidates to classify the test sample. We decrease the size of data at a low computational cost by choosing some candidates, and combine the holistic feature (Euclidean distance) with the local feature (SIFT) for face recognition.

The rest of this paper is organized as follows. Section 2 describes the extraction procedure of SIFT features. Section 3 describes the details of our method; Section 4 analyzes our method in detail. Section 5 shows our experiments on different face databases and the comparison with some other methods; Section 6 gives a brief conclusion.

2. **FEATURES EXTRACTION.** The SIFT is usually used to extract local features of an image, and it has several advantages: (1) SIFT features keep invariant on rotation, scale scaling, and illumination change of images; (2) SIFT features maintain a certain degree of stability on perspective changes, affine transformation, and robustness to the noise; (3) SIFT features are uniqueness and informative [41].

The feature extraction procedure of SIFT can be described as follows.

STEP1. Construct the DOG scale-space:

$$D(x, y, \sigma) = (G(x, y, k\sigma) - G(x, y, \sigma)) * I(x, y) = L(x, y, k\sigma) - L(x, y, \sigma) \tag{1}$$

where $G(x, y, \sigma) = \frac{1}{2\pi\sigma^2} e^{-(x^2+y^2)/2\sigma^2}$, $I(x, y)$ is the original image, and $L(x, y, \sigma) = G(x, y, \sigma) * I(x, y)$.

STEP2. Get the keypoints: We get the keypoints at the scale space extreme in the difference of Gaussian function convolved with the image.

STEP3. Assign an orientation and gradient modulus to each keypoint:

$$m(x, y) = \sqrt{(L(x+1, y) - L(x-1, y))^2 + (L(x, y+1) - L(x, y-1))^2} \tag{2}$$

$$\theta(x,y) = \tan^{-1}\left(\frac{L(x,y+1) - L(x,y-1)}{L(x+1,y) - L(x-1,y)}\right) \tag{3}$$

STEP4. Construct the descriptor of SIFT features: We sample within an 8*8 neighborhood window centered on the keypoints, and divide the neighborhood into four 4*4 child windows as shown in Fig.1. Then we calculate the gradient orientation histogram [41] with eight bins in each child window and get a 128-dimensional vector called descriptor.

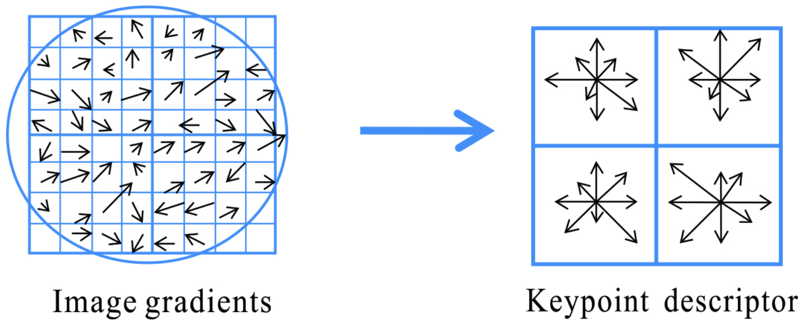

FIGURE 1. Get descriptors from the gradient orientation histogram

We can extract different numbers of descriptors from different images. Take images of one subject in FERET face database as an example, the descriptors of SIFT features extracted in the images can be shown as a series of points in Fig.2.

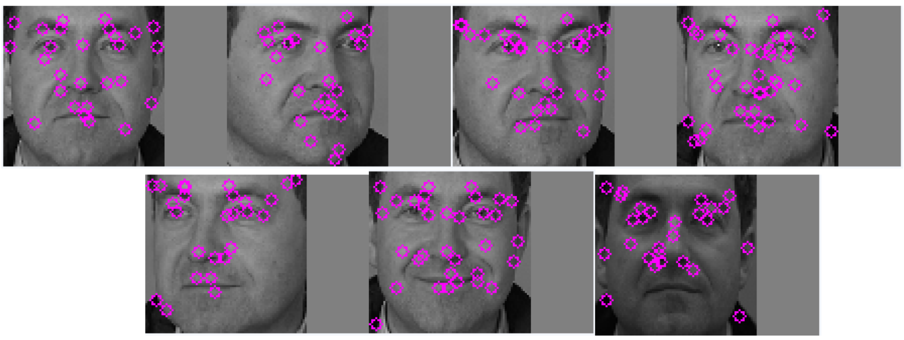

FIGURE 2. A show for the descriptors of SIFT extracted

3. DETAILS OF OUR METHOD.

Our method mainly consists of the following three steps:

STEP1. Choose N "candidates" based on the Euclidean distance. We define $X = \{x_{11}, \cdots, x_{ij}, \cdots\}$ as a test sample, $Y = \{y_{11}, \cdots y_{ij}, \cdots\}$ as a training sample, where x_{ij} and y_{ij} are the pixels at position (i.j) of the test and training samples respectively. Let $B = \{Y_1, Y_2, \ldots\}$ denote the training set. Let D(X,Y_k) be the Euclidean distance between a test sample X and a training sample Y_k.

$$D(X, Y_k)\sqrt{\sum_{i=1,j=1}^{i=H,j=W} (x_{ij} - y_{ij})^2} \tag{4}$$

H and W are the height and width of a test sample respectively. The training samples should have the same size as the test sample.

We sort the whole training samples in the ascending order of Euclidean distance, and then choose N training samples with smaller distance as "candidates" set C.

$$C = \{(Y_{1\dots}Y_N)|D(X,Y_1) \leq D(X,Y_2) \leq \ldots \leq D(X,Y_N) \leq D(X,Y_{N+1})\} \tag{5}$$

Where $N \leq n$, n is the number of training samples in training set B.

STEP2. Choose P new "candidates" based on SIFT features. In this step, we choose P new "candidates" from C based on the number of well matched pairs of SIFT features. First of all, we define the criterion of well matched pair of SIFT features. We build a KD-tree [42] using the descriptors of SIFT features in a training sample. And then, for each descriptor a in the test sample, we employ Best-Bin-First search algorithm to find out k nearest nodes $b_1, b_2, \ldots b_k$ in the KD-tree (usually k= 2), which are sorted in descending order. Let d_1,d_2 respectively be the distances between a and b_1, a and b_2. We then calculate the ratio of d_1, d_2:

$$ratio = {}^{d_1}\!/_{d_2} \tag{6}$$

If $ratio < Threshold$ (defined manually), we define a and b_1 are a well matched pair of SIFT features. Fig.3 shows the effect of $Threshold$ on the recognition accuracy. When $Threshold$ is below a certain value, the recognition accuracy increases rapidly and reaches the highest while $Threshold$ is 0.5. Thus, we fix it as 0.5 in our method.

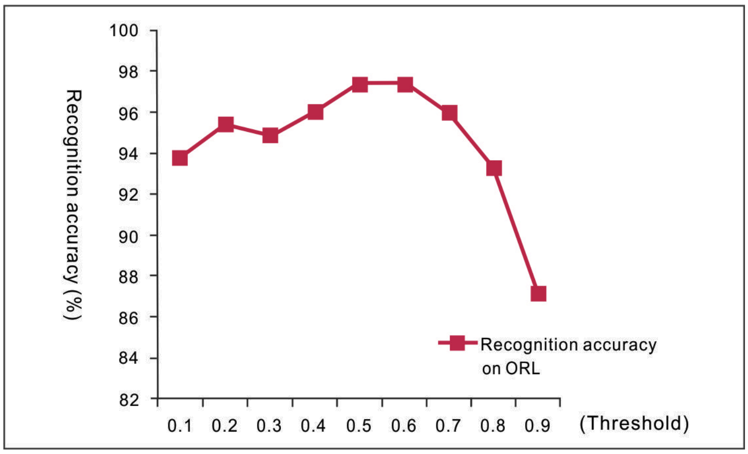

FIGURE 3. Recognition accuracy with respect to varying threshold

We take some images of a subject from FERET face database for example to show the well matched pairs of two images, as shown in Fig.4.

In this step, we count the number of well matched pairs of SIFT features between the test sample and each "candidate" in C, and sort the "candidate" samples in descending order. Then we choose P samples from C with the greater number of well matched pairs as the new "candidates" set $C'=\{C'_1,C'_2,\cdots,C'_P\}$.

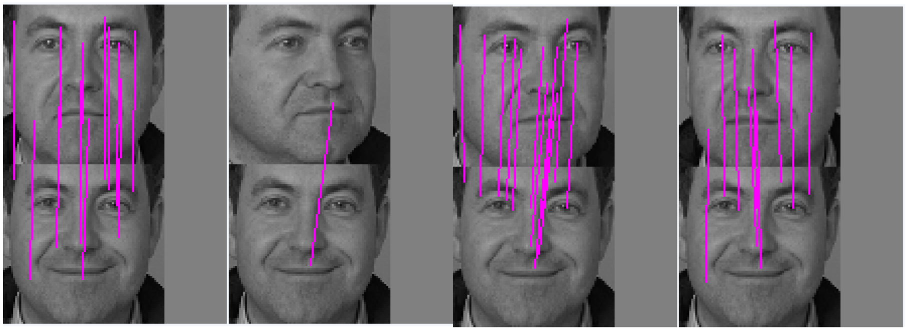

FIGURE 4. An example of matched pairs of SIFT features

STEP3. Calculate the average similarity between classes. Let M_i be a class with the same class label as new "candidates" $c'_i \in C'$, e_ij the angle of SIFT feature descriptors between a test sample and the j-th training sample in M_i, and $\bar{e}_i$ be the average angle between the test sample and each class M_i. We calculate e_ij by using Eq. (7), and $\bar{e}_i$ by using Eq. (8).

$$e_{ij} = \cos^{-1}\left(\frac{f_{test} \cdot f_{ij}}{\|f_{test}\| \cdot \|f_{ij}\|}\right) \tag{7}$$

$$\bar{e}_i = \frac{\sum\limits_{j=1}^{L} e_{ij}}{L} \tag{8}$$

Where f_{test} and f_{ij} are the descriptors of the test sample and the j-th training sample in M_i respectively, and L is the number of training samples in M_i.

In our method, we describe the similarity of a test sample and the class M_i by using $\bar{e}_i$ and choose the class with the minimum angle as the recognition result.

To get high recognition accuracy, we adjust some parameters involved in our method by using a small validation set. The parameters mainly include the numbers N,P of "candidates" chosen in the first two steps and the $Threshold$ used to count the number of well matched pair of SIFT.

4. ANALYSIS OF OUR METHOD.

The procedure of our method in section 3 can be shown in Fig.5.

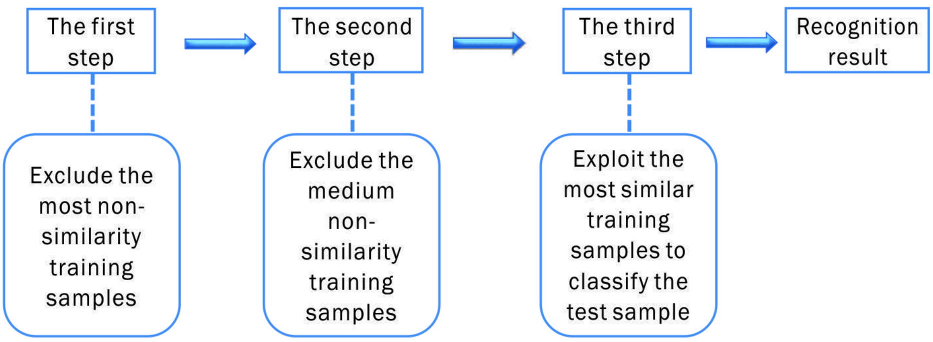

FIGURE 5. The general procedure of our method

In our method, the first step is on the assumption that for a test sample, theres more possible to have smaller distance with the samples from the same class than the ones from different classes. Through an experiment on AR and FERET face databases, we find that more than 93% of the "candidates" chosen in the first step have the same class label as the test sample. This indicates that our assumption is ground-truth. Take a test sample X_1 with the class label of "1" for example; Fig.6 shows the Euclidean distances of X_1 and the samples in the same class are less than that of X_1 and the ones in the different classes. So this step excludes those samples which are much different from the test sample, and decreases their interferences on the recognition.

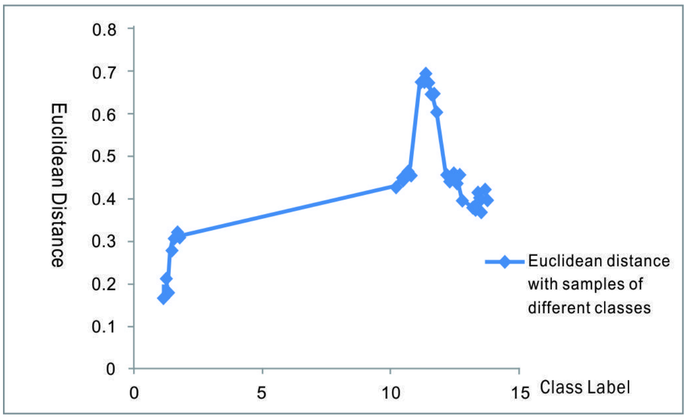

FIGURE 6. The horizontal axis (x) represents the class label of different samples; the vertical axis (y) represents the Euclidean distance of X_1 with samples. The distance of X_1 with samples from the same class is smaller than those from different classes.

The second step aims to modify the first step. We choose a certain number of samples from "candidates" chosen in the first step based on the number of matched pairs of SIFT features to modify the rank of "candidates". Whats more, by the first two steps, we combine the Euclidean distance with SIFT features. This makes us pay more attentions on the holistic as well as the detailed information of an image.

In the third step, we make our decisions depending on the new "candidates" set C'. However, just using the information of $c_i{}' \in C'$ is not abundant. In fact, this situation maybe happens: a sample is similar to the test sample, but its not from the same class as the test sample. Obviously, this situation leads error recognition. To avoid this situation as much as possible, we make use of all training samples in the same class as $c_i{}'$, which contains more information of different poses, gestures, et al. to calculate the mean similarity between each class with the test sample, and pick the class with the highest similarity as recognition result.

Taking a test sample (shown in Fig.7) in AR face database for example, we show the corresponding results of three steps in Fig.8, Fig.9, and Fig.10 respectively. By the first step, we choose $N = 15$ "candidates". Fig.8 shows the "candidates" in the ascending order of Euclidean distance, and the nearest sample is not from the same class as the test sample. By the second step, we choose $P = 5$ new "candidates". Fig.9 shows them in the descending order of the number of matched pairs of SIFT features. We can find the ranking of the sample from the same class as the test sample increased. At last, Fig.10 shows the recognition result after the third step. From these figures, we can find that not only the interference samples are less after the first two steps, but also the rank of the correct sample increases after the second step.

FIGURE 7. The test sample

FIGURE 8. The "candidates" chosen in the first step

FIGURE 9. The new "candidates" chosen in the second step

FIGURE 10. The recognition result after the third step

In contrast, if we recognize a test sample by calculating the number of matched pairs of SIFT features directly without choosing "candidates", the rate of error recognition will be higher. As shown Fig.11, an error recognition result is given for the test sample. This shows the effectiveness of choosing "candidates":

FIGURE 11. The recognition result without choosing "candidates"

In summary, by the first two steps, we reduce the size of data significantly, and abandon non-similarity samples as many as possible, which helps us pay more attentions on those more similar samples; in the third step, we characterize the similarity well using a weighted sum. Thus, our method obtains high recognition accuracy.

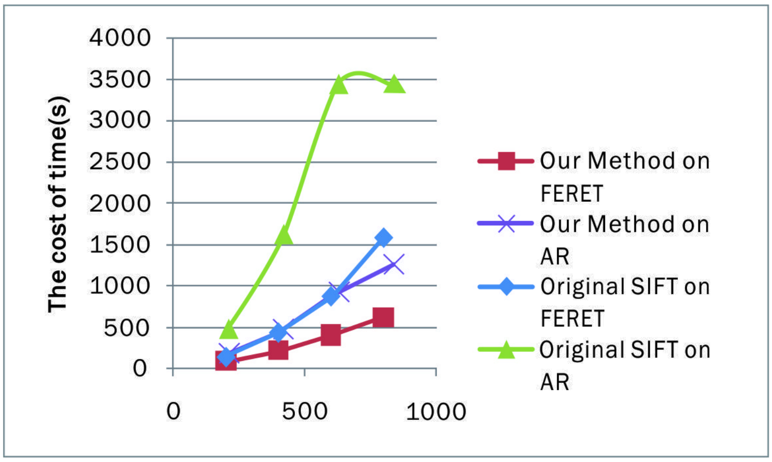

FIGURE 12. Comparison of our method with the original SIFT method that classifies on the entire training set. The horizontal axis (x) represents the number of training samples; the vertical axis (y) represents the average cost of time on the condition that N=15 and P=15. It shows that our method has low computational complexity.

FIGURE 13. Partial results of our method on FERET database, the first and third lines are test samples, and images below them are their corresponding recognition result. It shows that our method gets the correct recognition result in the most cases.

5. **EXPERIMENT AND COMPARISON.** We conduct a number of experiments on AR [43], FERET [44], and ORL [45] face databases to test our method, and make comparisons with some other methods, including the original SIFT, TPTSR [41], Principal Component Analysis (PCA), Sparse Representation Classification (SRC) [46] and Linear Regression Classification (LRC) [47].

5.1. **Parameters of Our Method.** To verify the influence of the number N of "candidates" chosen in the first step and the number P of "candidates" chosen in the second step on our method, we conduct a series of experiments with different N and P on a small validation set from AR, FERET, and ORL. Fig.14 and Fig.15 show the result.

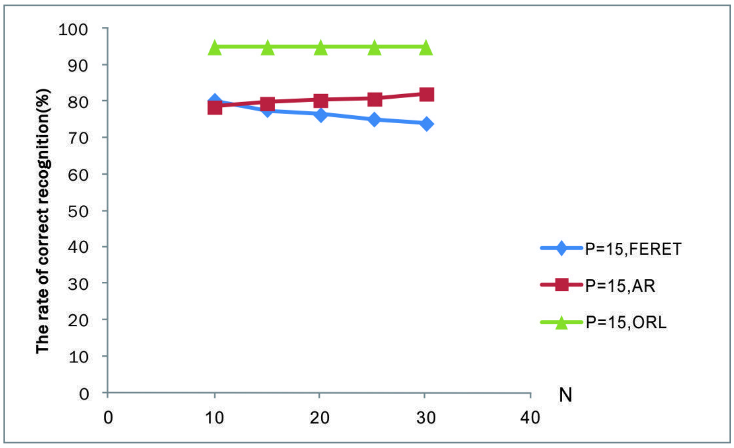

FIGURE 14. The horizontal axis (x) represents the value range of N; the vertical axis (y) represents the recognition accuracy on different databases on the condition of P=15.

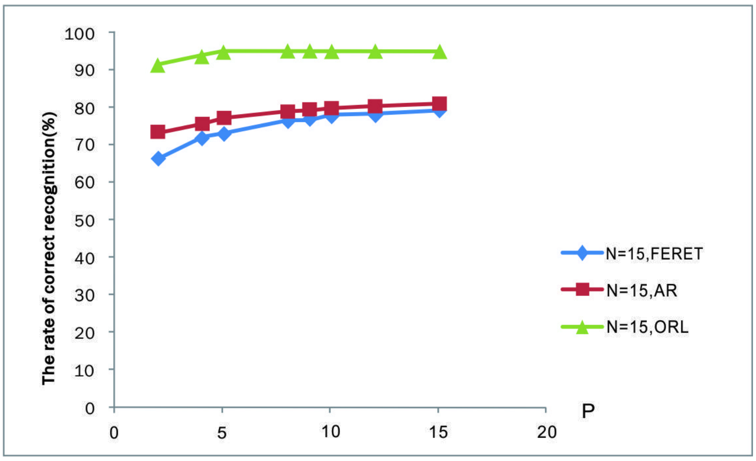

FIGURE 15. The horizontal axis (x) represents the value range of P; the vertical axis (y) represents the recognition accuracy on different databases on the condition of N=15.

Fig.14 shows too large N may decrease the recognition rates. The probable reason is that more "candidates" contain more non-similarity samples, which increases the probability of error recognition. Fig.15 shows larger P increases the recognition accuracy to a certain extent, but it cannot be bigger than N. Moreover, the bigger N and P will be of higher computational complexity. Finally we find the fittest values of N=15 and P=15, on which the following experiments are conducted.

5.2. **Face Databases Used In Experiments.**

5.2.1. *FERET Face Database.* In this database, we use 1400 images whose resolution is 80 * 80, from 200 subjects with each subject providing 7 images. We choose the former 4 images of each subject as training samples, and the remained 3 images are regarded as the corresponding test samples. Finally, we get a training set including 800 images with 200 subjects, and a test set including 600 images with 200 subjects. Fig.16 shows some samples in FERET database, and TABLE 1 shows the recognition accuracy of our method on different number of subjects from FERET database.

FIGURE 16. Some face images from FERET database

TABLE 1. The recognition accuracy of our method on FERET

Number of Subject	Number of Test Set	Number of Training Set	OUR METHOD	PCA	SRC	LRC	SIFT FEATURE ONLY	TPTSR
50	150	200	86%	46.3%	74.7%	87.3%	76.7%	64.5%
100	300	400	79.7%	46.3%	65.7%	82.3%	70%	60.1%
150	450	600	79.1%	40.7%	64.7%	80.7%	68.4%	55.8%
200	600	800	75.7%	53.1%	61%	---	64.3%	50.2%

5.2.2. *ORL Face Database.* In this database, we use 1400 images whose resolution is 46*56, from 40 subjects with each subject providing 10 images. We choose the former 5 images of each subject as training samples, and the remained 5 images are the test samples. Finally, we get a training set including 200 images with 40 subjects, and a test set including 200 images with 40 subjects. Fig.17 shows some samples in ORL database, and TABLE 2 shows the recognition accuracy of our method on different number of subjects from ORL database.

FIGURE 17. Some face images from FERET database

Table 2. The recognition accuracy of our method on AR

Number of Subject	Number of Test Set	Number of Training Set	OUR METHOD	PCA	SRC	LRC	SIFT FEATURE ONLY	TPTSR
10	50	50	90%	82%	100%	100%	90%	93%
20	100	100	93%	83%	93%	94%	93%	93.5%
30	150	150	92%	82.5%	90.7%	87.3%	92%	94%
40	200	200	94%	84%	88%	88%	94%	94%

Figure 18. Some face images from AR database

Table 3. The recognition accuracy of our method on AR

Number of Subject	Number of Test Set	Number of Training Set	OUR METHOD	PCA	SRC	LRC	SIFT FEATURE ONLY	TPTSR
30	210	210	91.4%	53.7%	77.2%	75.2%	98.1%	79.2%
60	420	420	87.6%	53%	83.6%	77.9%	95.9%	77.9%
90	630	630	87.9%	53%	81.9%	74.9%	93.6%	77.6%
120	840	840	80.7%	52.9%	76.8%	69.4%	89.6%	75.7%

5.2.3. *AR Face Database.* In this database, we use 1680 images whose resolution is 80 * 100, from 120 subjects with each subject providing 14 images. We choose the former 7 images of each subject as training samples, and the remained 7 images are regarded as the corresponding test samples. Finally, we get a training set including 840 images with 120 subjects, and a test set including 840 images with 120 subjects. Fig.18 shows some samples in AR database, and TABLE 3 shows the recognition accuracy of our method on different number of subjects from AR database.

5.3. **Comparison with other methods.** To test the performance of our method, we conduct a series of experiments on different face databases to compare our method with other methods. The experiment results are shown as follows.

Using the coarse to fine strategy and the information of classes rather than only single sample, is more effective and can achieve a better performance. To prove it, we design an experiment to compare our method with the original SIFT which counts the number of matched pairs of SIFT features between each test sample and all training samples, and chooses the training sample with the largest number of matched pair as the recognition result. Fig.19 and Fig.20 show the result of comparison on different databases.

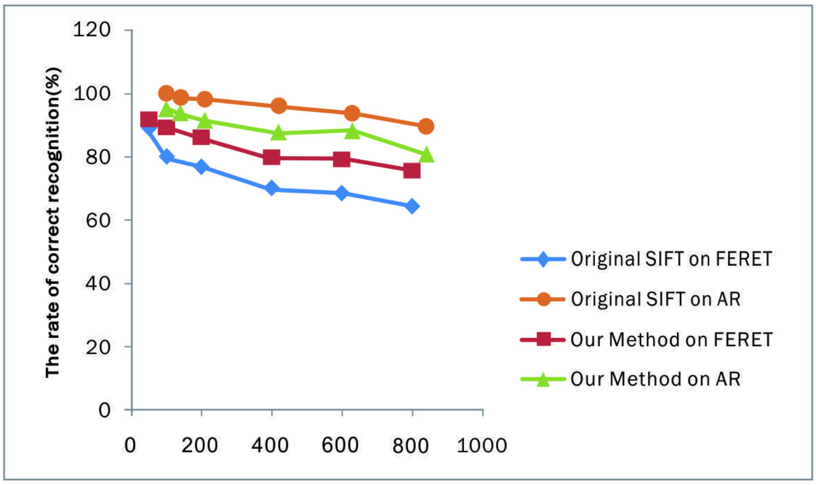

FIGURE 19. Comparison our method and the original SIFT on FERET and AR databases; the horizontal axis (x) represents the number of training set; the vertical axis (y) represents the recognition accuracy.

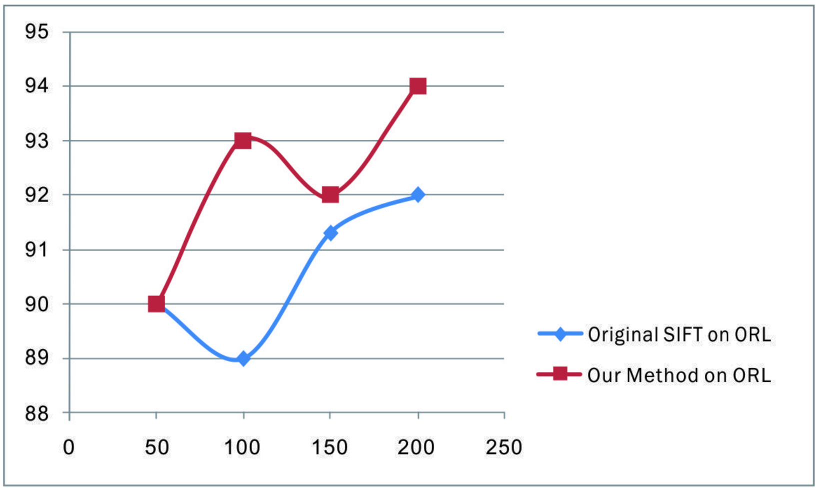

FIGURE 20. Comparison of our method and the original SIFT on ORL; the horizontal axis (x) represents the number of training set; the vertical axis (y) represents recognition accuracy. Our method has higher recognition accuracy.

We compare our method with PCA that uses 50, 100, 150, 200 transform axes on ORL database, 140, 210, 420, 630 transform axes on AR database, 100, 200, 400, 600 transform axes on FERET database, respectively. Fig.21 and Fig.22 show the result of the comparisons.

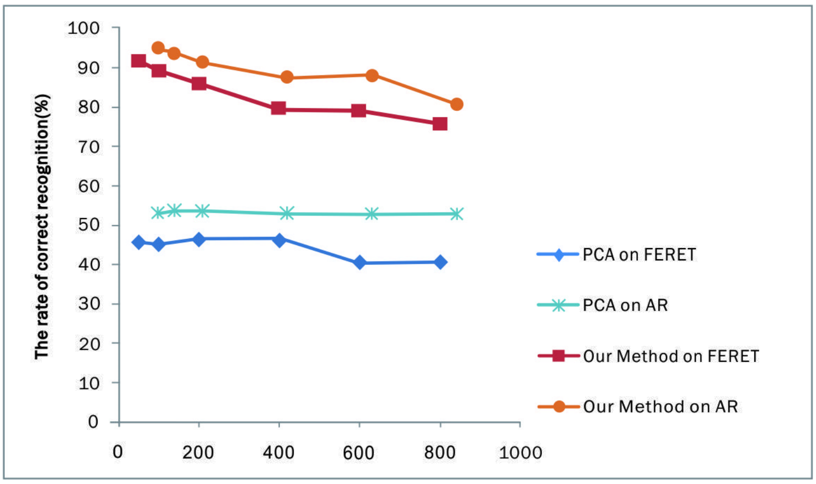

Figure 21. Comparison of our method and PCA on FERET and AR; the horizontal axis (x) represents the number of training set for our method/transform axes for PCA; the vertical axis (y) represents the recognition accuracy.

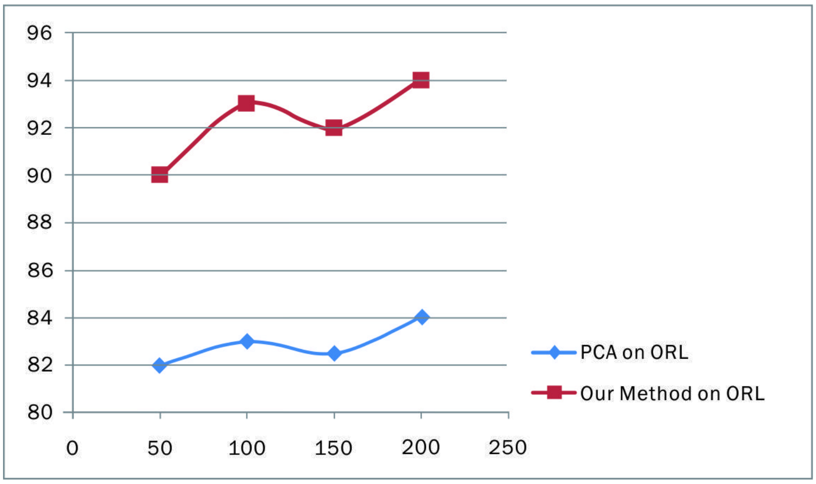

Figure 22. Comparison of our method and PCA on ORL; the horizontal axis (x) represents the number of training set for our method/transform axes for PCA; the vertical axis (y) represents the recognition accuracy. From Fig.21 and Fig.22, Its easy to find that our method performs much better than PCA.

To prove the good performance of SIFT features used in our method, we compare our method with TPTSR [38], which makes use of sparse representation based on choosing neighbors of the test sample from all training samples. Fig.23 shows the result of the comparison on FERET and AR databases.

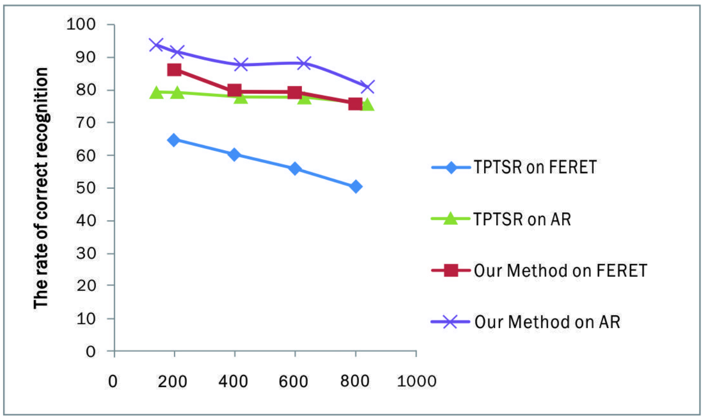

FIGURE 23. Comparison of our method and TPTSR on FERET, AR, and ORL; the horizontal axis (x) represents the number of training set; the vertical axis (y) represents the recognition accuracy. We can find that our method outperforms over TPTSR, especially on FERET and AR databases.

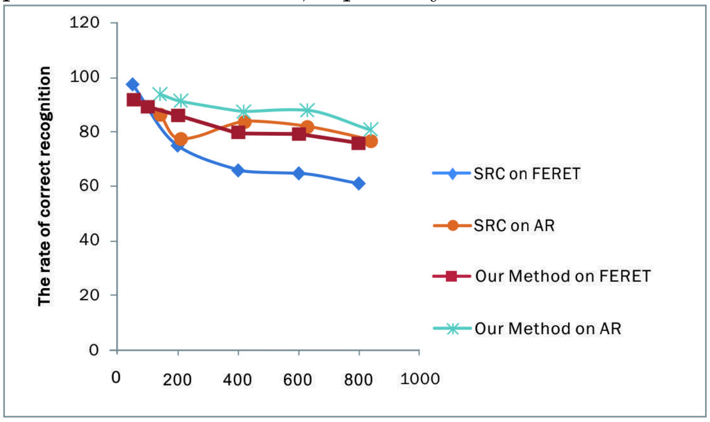

FIGURE 24. Comparison of our method and SRC on FERET, AR, and ORL; the horizontal axis (x) represents the number of training set; the vertical axis (y) represents the recognition accuracy.

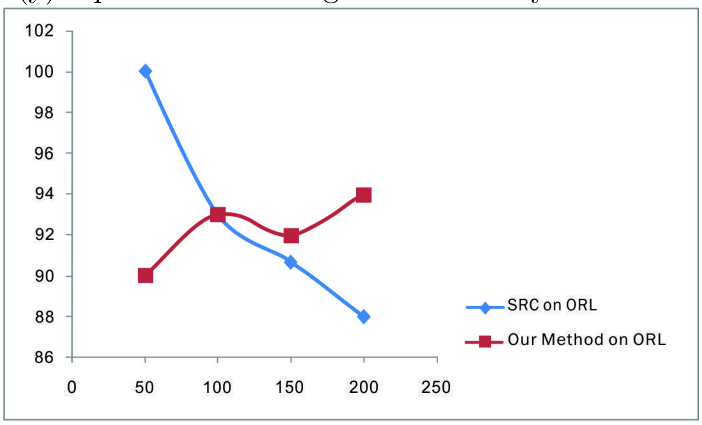

FIGURE 25. Comparison of our method and SRC on ORL; the horizontal axis (x) represents the number of training set; the vertical axis (y) represents the percentage of correct recognition. From Fig 24 and Fig 25, we can find that our method outperforms over SRC, especially on FERET and AR.

Fig.24 and Fig.25 show the comparison of our method with SRC. We can easily find that our method has a better performance than SRC.

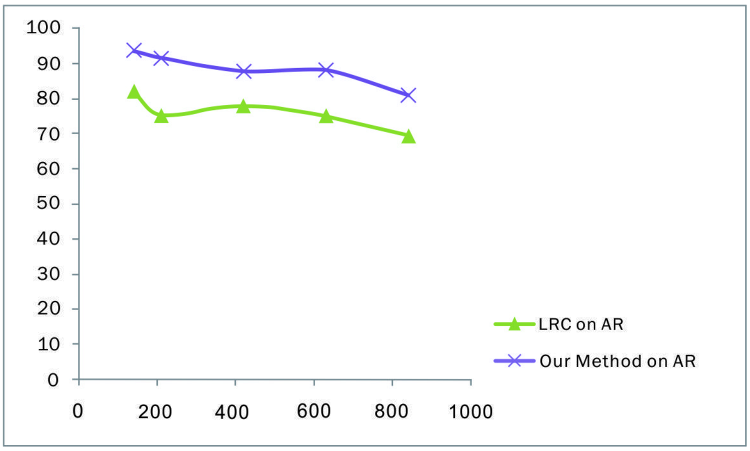

FIGURE 26. Comparison of our method and LRC on AR database; the horizontal axis (x) represents the number of training samples; the vertical axis (y) represents recognition accuracy.

6. **CONCLUSION.** This paper proposes a SIFT features based method for face recognition, and adopts the coarse to fine strategy by choosing "candidates" to reduce the size of data. The proposed method combines a local feature with a holistic classification method. To optimize the data set and compute effectively, we choose "candidates" twice based on Euclidean distance and SIFT features respectively, which reduce the computational complexity and enhance the effectiveness for face recognition at the same time. Besides, after picking out new "candidate" samples, we use the whole samples in the same class as the new "candidates" rather than only the new "candidates". A number of experiments show that our method performs well on several face databases. In the future, we will try to use our method in practical application and continue digging out more possible ways to apply SIFT features, and then improve its performance on face recognition.

REFERENCES

[1] C. L. Lin, T. C. Chuang, and K. C. Fan, Palmprint verification using hierarchical decomposition, *Journal of Pattern Recognition*, vol. 38, no. 12, pp. 2639-2652, 2005.

[2] Y. Xu, D. Zhang, and J. Y. Yang, A feature extraction method for use with bimodal biometrics, *Journal of Pattern Recognition*, Vol. 43, no. 3, pp. 1106-1115, 2010.

[3] D. Zhang, F. Song, Y. Xu, and Z. Liang, *Advanced pattern recognition technologies with applications to biometrics*, IGI, Hershey, USA, 2008.

[4] Y. Xu, A. Zhong, J Yang, and David Zhangn, Lpp solution schemes for use with face recognition, *Journal of Pattern Recognition*, vol. 43, no. 12, pp. 4165-4176, 2010.

[5] Y. Xu, Q. Zhu, and D. Zhang, Combine crossing matching scores with conventional matching scores for bimodal biometrics and face and palmprint recognition experiments, *Journal of Neurocomputing*, vol. 74, no. 18, pp. 3946-3952, 2011.

[6] Y. Xu, W. Zuo, and Z. Fan, Supervised sparse representation method with a heuristic strategy and face recognition experiments, *Journal of Neurocomputing*, vol. 79, no. 1, pp. 125-131, 2012.

[7] X. Li, A. Li, and X. Bai, 3D face detection and face recognition: state of the art and trends, *Proc. of International Society for Optical Engineering*, pp. 78201E-1-78201E-7, 2010.

[8] M. A. Akhloufi, A. Bendada, and J. C. Batsale, Multispectral face recognition using non linear dimensionality reduction, *Proc. of International Society for Optical Engineering*, pp. 73410J-1-73410J-10, 2009.

[9] M. K. Hsu, C. Hsu, T. N. Lee, and H. Szu, Face recognition from a moving platform via sparse representation, *Proc. of SPIE*, vol. 8401, pp. 840106-1-840106-6, 2012.

[10] H. F. Wang, Y. Han, and Z. X. Zhang, Applying local gabor ternary pattern for video-based illumination variable face recognition, *Proc. of International Conference on Machine Vision*, pp. 83490W-1-83490W-6, 2011.

[11] D. Colbry, F. Oki, and G. Stockman, 3D face identification: experiments towards a large gallery, *Proc. of SPIE*, vol. 6944, pp. 694403-1-694403-9, 2008.

[12] W. Feng, Q. He, Y. Yan, G. Jin, and M. Wu, Real-time human face recognition system with high parallelism, *Proc. of SPIE*, vol. 3817, pp. 108-115, 1999.

[13] D. Y. Huang, C. H. Chen, W. C. Hu, S. C. Yi, and Y. F. Lin, Feature-based vehicle flow analysis and measurement for a real-time traffic surveillance system, *Journal of Information Hiding and Multimedia Signal Processing*, vol. 3, no. 3, pp. 279-294, 2012.

[14] C. Y. Chang, C. W. Chang, and C. Y. Hsieh, Applications of block linear discriminant analysis for face recognition, *Journal of Information Hiding and Multimedia Signal Processing*, vol. 2, no. 3, pp. 259-269, 2011.

[15] D. Y. Huang, C. J. Lin, and W. C. Hu, Learning-based face detection by adaptive switching of skin color models and AdaBoost under Varying Illumination, *Journal of Information Hiding and Multimedia Signal Processing*, vol. 2, no. 3, pp. 204-216, 2011.

[16] M. A. Turk, and A. P. Pentland, Face recognition using eigenfaces, *Journal of cognitive neuroscience*, vol. 3, no. 1, pp. 586-591, 1991.

[17] L. Zhang, Q. Gao, and D. Zhang, Block independent component analysis for face recognition, *Journal of The 14th International Conference on Image Analysis and Processing*, vol. 27, no. 9, pp. 217-222, 2007.

[18] J. Wang, Q. Li, J. You, and Q. Zhao, Fast kernel fisher discriminant analysis via approximating the kernel principal component analysis, *Journal of Neurocomputing*, vol. 74, no. 17, pp. 3313V3322, 2011.

[19] S. Arca, P. Campadelli, and R. Lanzarotti, A face recognition system based on local feature analysis, *Proc. of The 4th International Conference on Audio and Video-Based Biometric Person Authentication*, pp. 182-189, 2003.

[20] D. G. Lowe, Distinctive image features from scale-invariant keypoints, *Journal of Computer Vision*, vol. 60, no. 2, pp. 91-110, 2004.

[21] M. Bicego, A. Lagorio, E. Grosso, and M. Tistarelli, On the use of sift features for face authentication, *Proc. of Conference on Computer Vision and Pattern Recognition Workshop*, 2006.

[22] J. Luo, Y. Ma, E. Takikawa, S. Lao, M. Kawade, and B. L. Lu, Person-specific sift features for face recognition, *Proc. of The IEEE International Conference on Acoustics, Speech and Signal Processing*, pp. 593-596, 2007.

[23] Y. Ke and R. Sukthankar, Pca-sift: a more distinctive representation for local image descriptors, *Proc. of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 506-513, 2004.

[24] C. Geng, Face recognition using sift features, *Proc. of The 16th IEEE International Conference on Image Processing*, pp. 3313-3316, 2009.

[25] A. Majumdar, and R. K. Ward, Discriminative sift features for face recognition, *Proc. of Canadian Conference on Electrical and Computer Engineering*, pp. 27-30, 2009.

[26] J. Križaj, V. Štruc and N. Pavešić, Adaptation of sift features for robust face recognition, *Proc. of The 7th international conference on Image Analysis and Recognition*, pp. 394-404, 2010.

[27] K. Mikolajczyk, and C. Schmid, A performance evaluation of local descriptors, *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 27, no. 10, pp. 1651-1630, 2005.

[28] B. Dai, D. Zhang, H. Liu, S. Sun, and K. Li, Evaluation of face recognition techniques, *Proc. of SPIE*, vol. 7489, pp. 74890M-1-74890M-7, 2009.

[29] C. Geng, and X. Jiang, Face recognition using sift features, *Proc. of The 16th IEEE International Conference on Image Processing*, pp. 3277-3280, 2009.

[30] A. Mian, Online learning from local features for video-based face recognition, *Journal of Pattern Recognition*, vol. 44, no. 5, pp. 1068-1075, 2011.

[31] Z. Lai, Z. Jin, J. Yang, and W. K. Wong, Sparse local discriminant projections for feature extraction, *Proc. of The 20th International Conference on Pattern Recognition*, pp. 926-929, 2010.

[32] J. Wright, Y. Ma, J. Mairal, G. Sapiro, T. Huang, and S. Yan, Sparse representation for computer vision and pattern recognition, *Proc. of IEEE*, vol. 98, no. 6, pp. 1031-1044, 2010.

[33] J. Wright, A. Y. Yang, A. Ganesh, S. S. Sastry, and Y. Ma, Robust face recognition via sparse representation, *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 31, no. 2, pp. 210-227, 2009.

[34] Y. Xu, Z. Fan, and Q. Zhu, Feature space-based human face image representation and recognition, *Journal of Optical Engineering*, vol. 51, no. 1, pp. 017205-1-017205-7, 2012.

[35] Y. Xu, A. Zhong, J. Yang, and D. Zhanget, Bimodal biometrics based on a representation and recognition approach, *Journal of Optical Engineering*, vol. 50, no. 3, pp. 037202-1-037202-7, 2011.

[36] Y. Shi, D. Q. Dai, C. C. Liu, and H. Yan, Sparse discriminant analysis for breast cancer biomarker identification and classification, *Journal of Progress in Natural Science*, vol. 19, no. 11, pp. 1635-1641, 2009.

[37] Y. Xu, and Q. Zhu, *A simple and fast representation-based face recognition method*, Springer, London, UK, 2012.

[38] E. Elhamifar, and R. Vidal, Sparse subspace clustering, *Proc. of IEEE International Conference on Computer Vision and Pattern Recognition*, pp. 2790-2797, 2009.

[39] L. Zhang, M. Yang, and X. Feng, Sparse representation or collaborative representation: which helps face recognition, *Proc. of IEEE International Conference on Computer Vision*, pp. 471-478, 2011.

[40] Y. Xu, D. Zhang, J. Yang, and J. Y. Yang, A two-phase test sample sparse representation method for use with face recognition, *IEEE Trans. Circuits and Systems for Video Technology*, vol. 21, no. 9, pp. 1255-1262, 2011.

[41] SIFT Feature, 2011, `http://www.cnblogs.com/cfantaisie/archive/2011/06/14/2080917.html`.

[42] Rob Hess, 2011, `http://blogs.oregonstate.edu/hess/`.

[43] A. Martinez, and R. Benavente, The ar face database, 1998, `http://www2.ece.ohio-state.edu/~aleix/ARdatabase.html`.

[44] P. J. Phillips, The facial recognition technology (feret) database, `http://www.itl.nist.gov/iad/humanid/feret/feret_master.html`.

[45] K. Vaishnavi, and G. P. R. Kumar, Face recognition using passion back propagation neural networks, *Journal of Computing Technology and Information Security*, vol. 1, no. 2, pp. 9-15, 2011.

[46] J. Wright, A. Y. Yang, A. Ganesh, S. S. Sastry, and Y. Ma, Robust face recognition via sparse representation, *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 31, no. 2, pp. 210-227, 2009.

[47] I. Naseem, R. Togneri, and M. Bennamoun, Linear regression for face recognition, *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 32, no. 11, pp. 2106-2112, 2010.