International Journal of Innovative Computing, Information and Control
Volume 8, Number 12, December 2012
 ISSN 1349-4198
pp. 8545–8556

COMBINING SPEEDED-UP ROBUST FEATURES WITH PRINCIPAL COMPONENT ANALYSIS IN FACE RECOGNITION SYSTEM

Shinfeng D. Lin*, Bo-Feng Liu and Jia-Hong Lin

Department of Computer Science and Information Engineering
National Dong Hwa University
No. 1, Sec. 2, Da Hsueh Rd., Shoufeng, Hualien 97401, Taiwan
*Corresponding author: david@mail.ndhu.edu.tw

Received October 2011; revised March 2012

Abstract. *Recently, the techniques of face recognition have been widely used in security application such as security monitoring, and access control. However, there are still some problems in face recognition system in which the light changes, expression changes, head movements and accessory occlusion are the main issues. In this article, a robust face recognition scheme is proposed. Speeded-Up Robust Features algorithm is used for extracting the feature vectors with scale invariance and pose invariance from face images. Then PCA is introduced for projecting the SURF feature vectors to the new feature space as PCA-SURF local descriptors. Finally, the K-means algorithm is applied to clustering feature points, and the local similarity and global similarity are then combined to classify the face images. Experimental results show that the performance of the proposed scheme is better than other methods, and PCA-SURF feature is more robust than original SURF and SIFT local descriptors against the accessory, expression, and pose variations.*
Keywords: Face recognition, Speeded-up robust features, Principal component analysis

1. Introduction.

In recent years, there has been a dramatic proliferation of research concerned with the face recognition system. For the access control, compared with the traditional personal guard or ID card, the face recognition is more effective and has better interaction between human and machine. However, there have been more difficult problems in the face recognition; they are the variance of illumination, pose, expression, accessory, and aging. Therefore, many algorithms focusing on how to conquer these problems have still received much attention.

The feature extraction of face image is an important procedure for face recognition. There are some traditional algorithms for face recognition such as PCA [1,2] and FLD [3] that are classical methods based on representing features of holistic image with the projection in subspace. The DWT [4] and DCT [5] have been utilized to extract feature information in various studies on face recognition. Recently, LBP [6] and Gabor filter [7] were proposed to increase the performance of recognition. Yan *et al.* [8] proposed a method to combine LBP and Gabor filter for improving the recognition rate. This method uses LBP and Gabor filter to extract the descriptor and combines the similarity score matrix of LBP and Gabor to obtain the improved performance in face recognition.

Scale Invariant Feature Transform (SIFT) [9] was proposed by Lowe in 2004. This algorithm has been successfully and widely applied to object recognition problems. SIFT algorithm contains three steps. The stable interest point detector in scale space is first applied and then SIFT precisely localizes the keypoints in scale space. Finally the rotation invariant local descriptors are constructed. There are some works based on the SIFT feature in face recognition such as SIFT_CLUSTER [10] and SIFT_GRID [11]. These works were proposed to investigate the performance of SIFT feature with different feature clustering strategy. Since the SIFT features in the matching stage have to consume a high computation time, some methods were proposed to reduce the dimension of features for speeding it up, such as PCA-SIFT [12] and GLOH [13]. However, the above-mentioned methods still cannot improve the complex problem of SIFT.

More recently, Bay *et al.* proposed a new descriptor which satisfies the speed requirement and is named Speeded-Up Robust Features (SURF) [14]. SURF is similar to the SIFT whereas both descriptors have scale and pose invariant characteristics; however, its performance and efficiency are better than SIFT. In SURF, the detectors are used for finding the interest points in an image; the feature vectors of interest point are then extracted as the descriptors. In contrast with SIFT that downsamples the image size and uses the difference of Gaussian (DoG) and Hessian detector to detect interest points, SURF applies the fast Hessian-matrix by up-scaling the filter size on the integral image to locating the interest points. For constructing the descriptor, the 64-dimension SURF descriptor is used to describe the feature of interest point's neighborhood. In the matching stage, the SURF descriptor only has 64 dimensions and the sign of Laplacian is used for building the fast indexing scheme, so the efficiency of SURF is much better than SIFT.

For the discrete SIFT and SURF features, the features between two images cannot be compared only based on their locations since the positions and number of features may locate in the different part of each image. The situation of a feature around the eye matching to a feature around the mouth corner will occur if a feature directly compares with the whole features of another image. Therefore, some methods were proposed to solve this problem. An overlapping regular grid matching strategy was proposed by Majumdar and Ward [15]. They apply three different matching methods, Maximal, Grid-Based, and Grid-Based Best matching, to matching features and comparing the performance in face recognition. In [16], K-means clustering algorithm was applied to constructing the subregions of images. Then the global similarity and local similarity were combined to classify face images. Both of these two methods have a good performance on face recognition.

Inspired by the methods mentioned above, a face recognition scheme which combines SURF, PCA, and K-means algorithm is proposed. The main contribution is its performance and efficiency are better than SIFT-based scheme with scale and pose invariance. The proposed scheme contains two parts. The first part is constructing the feature descriptors. The SURF descriptor is combined with the PCA algorithm to extract the features of an image. The second part is matching the features to classify the face images. We apply K-means algorithm to clustering image features and then calculate similarity between images to classify the face images. The performance under different conditions is evaluated to demonstrate the superiority of the proposed scheme.

The rest of the paper is organized as follows. Section 2 presents the proposed face recognition scheme. Experimental results are demonstrated in Section 3. Finally, Section 4 draws conclusions.

2. **The Proposed Face Recognition Scheme.** The general procedure of face recognition can be roughly divided to two parts: feature extraction, and feature classification. In the proposed scheme of face recognition, the algorithm of feature extraction bases on Speeded-Up Robust Features and PCA to establish the local descriptors. In the feature classification stage, the K-means algorithm is first applied to clustering the local descriptors, and then the local and global similarities are combined to classify the face images. The flowchart of the proposed scheme is illustrated in Figure 1 and three main stages of the proposed scheme (feature extraction, feature clustering, and the classification stage) are briefly described in Section 2.1, Section 2.2, and Section 2.3, respectively.

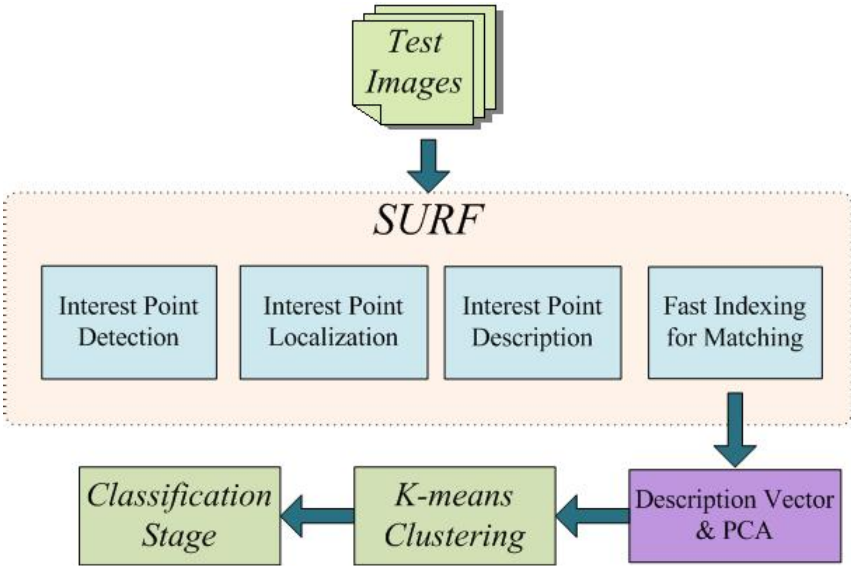


FIGURE 1. The complete flowchart of the proposed face recognition scheme

2.1. **The proposed PCA-based SURF descriptor.** Standard SURF procedure [14] can roughly be divided to three parts: interest point detection, interest point description, and interest point matching. The details of interest point detection and description are described as follows. After that, the PCA projection matrix is built and the descriptors are projected to a new space with lower dimension and further utilized to represent the feature of an image.

2.1.1. *Interest point detector and descriptor.* The detector of standard SURF is based on the approximate Hessian matrix. The determinant of the approximate Hessian matrix can represent the blob response at that location and scale of an image. For each point of an image in scale space, if the blob response at this location and scale is a local maximum, this point is denoted as the interest point. For constructing the SURF descriptors, a square region centered on the interest point is extracted and aligned to the dominate orientation. Then the region is spilt up equally into 4 smaller square sub-regions. The Haar wavelet responses within each sub-region are computed and summed up to describe the feature of the interest point.

In this scheme, the descriptor vector of each interest point has 128 elements. For face recognition, rotation-invariance is often not necessary, so the upright version of SURF is applied. For the upright version, the SURF descriptors use the reference implementation in [17].

2.1.2. *Local descriptor projection.* In the proposed scheme, the dimension of the feature vector is chosen to be 128 because the 64D vector is not good for finding the inline-points which is used in the classification stage. To improve computation efficiency, PCA is applied to modeling the identity of interest points and reducing the dimension of the feature. First, the eigenspace must be built with a training set. The training set is used for collecting the SURF feature vectors. Then, PCA is applied to the scatter matrix of these feature vectors and estimates the projection matrix.

The projection matrix is used for projecting the feature vectors to the new feature space as the PCA-SURF descriptors. The process of projection not only reserves the identity of the interest points but also discards unmodeled distortions. Besides, the computational complexity is greatly reduced because of the low dimensionality of new feature space. The flowchart of feature vectors' projection is shown as Figure 2.

2.2. **Feature clustering.** The features of the interest points cannot be compared only based on their locations because the positions and number of features are different in the images. The situation of a feature around the eye matching to a feature around the mouth corner will occur if this feature of an image is directly compared with the whole features of another image. Therefore, we divide the image into several sub-regions by K-means algorithm and compare the features of each sub-region separately.

K-means clustering algorithm based on the positions of PCA-SURF features is used. Figure 3 shows the flowchart of feature clustering. After the K-means algorithm is done,

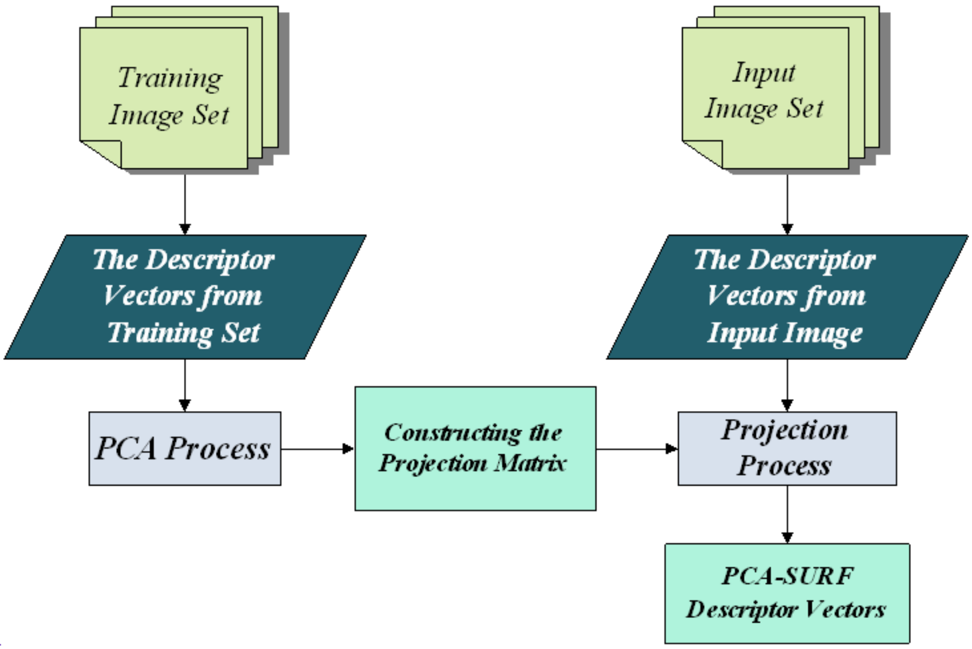


FIGURE 2. Flowchart of feature vectors' projection

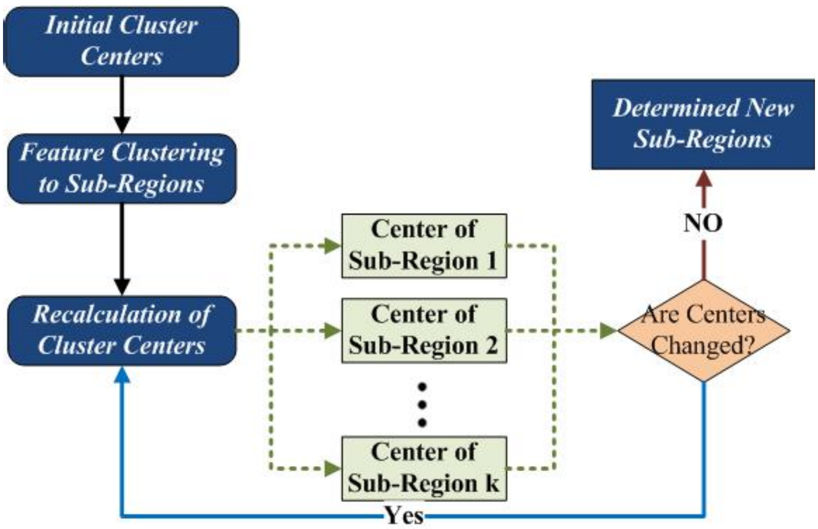
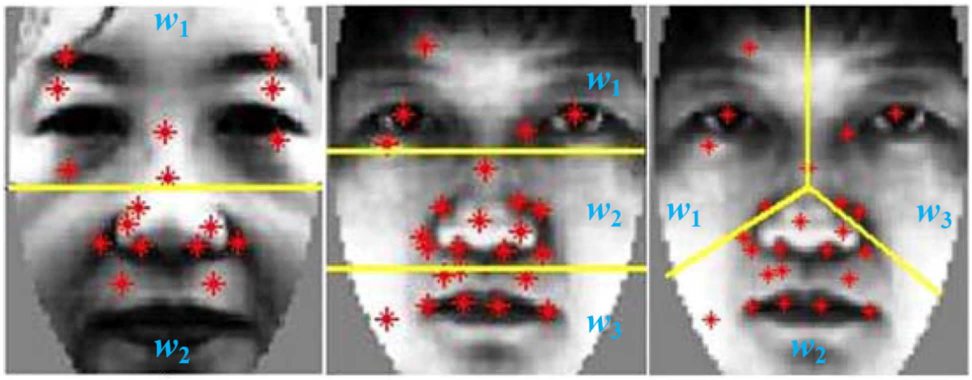


FIGURE 3. Flowchart of K-means algorithm



(a) (b) (c)

Figure 4. Face images of different clustering sub-regions. (a) The feature points are divided by the horizontal line. (b) The feature points are clustered to three equal parts. (c) The feature points are clustered and centered on left eye, right eye, and the center of mouth.

sub-regions of face images are determined. When a test image is inputted to the classification stage, feature points of the test image are classified to the sub-regions based on the cluster centers. In this article, three different approaches are proposed in the feature's clustering algorithm. The sample face images of different clustering approaches are illustrated in Figure 4.

2.3. **The classification stage.** The proposed classifying strategy is similar to [16] with a different global similarity and different situations of clustering. As the sub-regions are determined, local similarity and global similarity can be calculated and integrated to classify the face images. However, the complexity of computing similarity is very high. To improve this disadvantage, the fast indexing for matching method [14] is used for filtering the interest points with extreme difference. Figure 5 illustrates the flowchart of similarity computation, and the local similarity and global similarity are introduced in the following.

2.3.1. *Local similarity and global similarity.* We assume that the feature points of the input image I are located in k sub-regions and denoted as (1).

$$I = \left(f_1^1, \ldots, f_1^{m1}, f_2^1, \ldots, f_2^{m2}, \ldots, f_k^1, \ldots, f_k^{mk}\right) \tag{1}$$

where the f_k^j means the jth feature descriptor in the kth sub-region of image I.

For each feature in the sub-region of I_t, the first filter uses the fast index for matching method [14] to eliminate the extreme different features and then preserve the useful and similar features in the same sub-region of I_r and I_t. Then the local similarity S_L between a test image I_t and a reference image I_r is calculated as the following steps:

Step 1. Collect the interest points in each sub-region, and then compute the correlation between each pair of features in ith sub-regions of the test and reference image by (2).

$$d(f_{ti}^x, f_{ri}^y) = \frac{(f_{ti}^x, f_{ri}^y)}{\|f_{ti}^x\| \cdot \|f_{ri}^y\|} \tag{2}$$

Step 2. Choose the maximal similarity in ith sub-region as S_i by (3). The Steps 1 and 2 are illustrated in Figure 6.

$$S_i(I_t, I_r) = \max\left(d(f_{ti}^1, f_{ri}^1), \ldots, d(f_{ti}^1, f_{ri}^y), \ldots, d(f_{ti}^x, f_{ri}^y)\right) \tag{3}$$

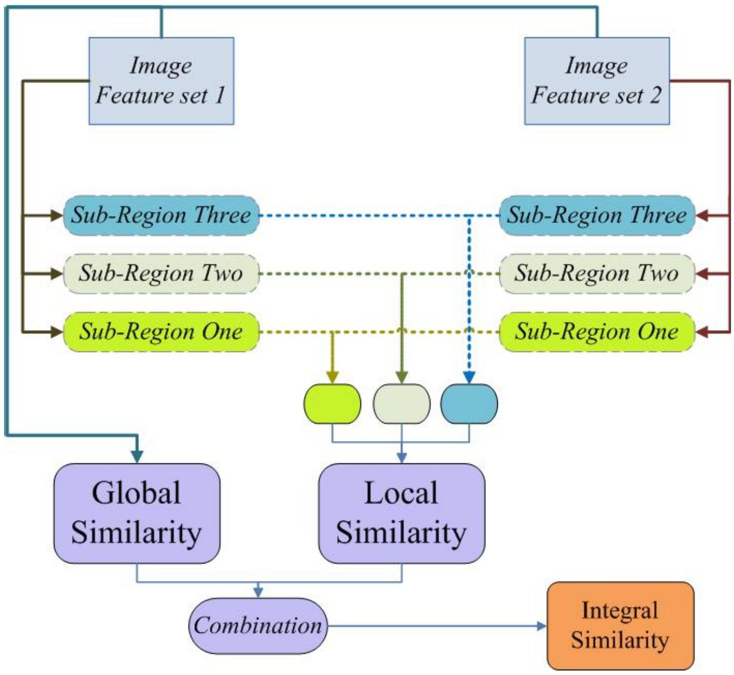


FIGURE 5. The flowchart of similarity computation

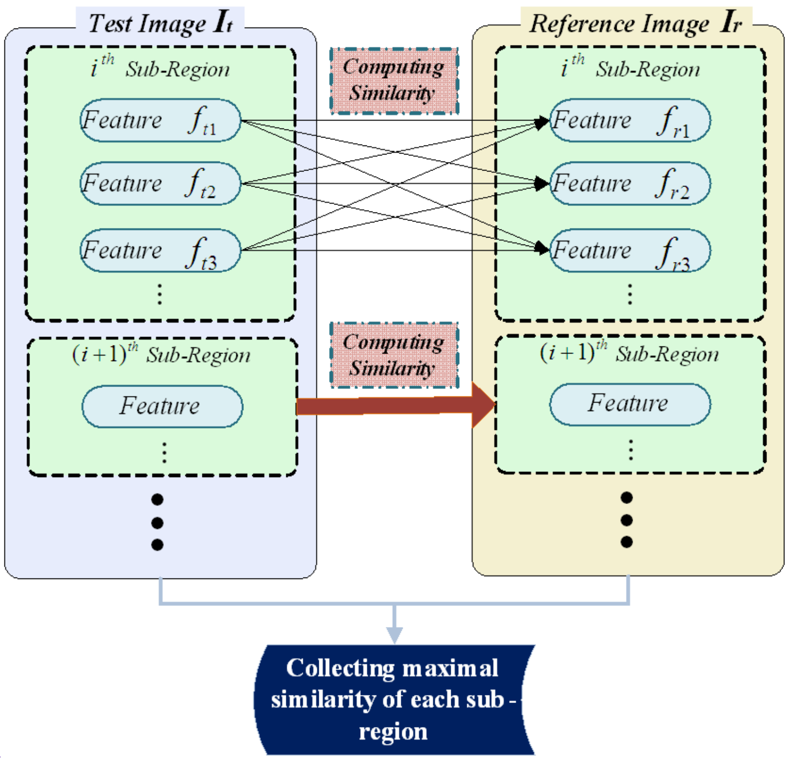


FIGURE 6. Computation of similarities in ith sub-region

Step 3. After collecting the maximal similarity from each sub-region of the test and reference image, the local similarity is computed as the average that S_i multiplies the weight of ith sub-region, w_i, by (4).

$$S_L(I_t, I_r) = \frac{1}{k} \sum\nolimits_{i=1}^{k} (S_i(I_t, I_r) \times w_i) \tag{4}$$

In steps 1 and 2, S_i is chosen as the maximum of the similarities rather than the average. Because most matching results are invalid, the similarities often have a small value such that the average of similarities is reduced. Therefore, choosing the average of the similarities as S_i may induce the poor discrimination between the sub-regions of I_t and I_r. In this article, the global similarity is calculated by combining the inline-point S_{ip} and cosine correlation S_c which are introduced as follows.

The inline-point similarity is computed as (5).

$$S_{ip} = match(I_t, I_r) \tag{5}$$

where $match(I_t, I_r)$ is the number of validly matched features of two images. The feature matching method is the same as [9] and described in the following:

Step 1. For each feature f_t^x of image I_t, estimate the Euclidean distances at all features in image I_r.

Step 2. If the ratio between the nearest and the second nearest distance is less than a threshold, f_t^x is determined to be correctly matched.

The distance ratio is utilized to determine whether the feature is matched because the nearest distance is much shorter than other distances if the feature is correctly matched.

The cosine similarity is similar to the local similarity; however, it is not limited to sub-regions. The correlations of the matched features between I_t and I_r are first computed by (2). Then, the cosine correlation is computed by (6)

$$S_C = \max\left(d(f_t^1, f_r^1), \ldots, d(f_t^1, f_r^y), \ldots, d(f_t^x, f_r^y)\right) \tag{6}$$

where S_C means the maximal cosine correlation, and $d(f_t^1, f_r^1)$ is the correlation of the feature between I_t and I_r. The complete global similarity is the combination of the S_{ip} and S_C by (7).

$$S_G = S_{ip} \times S_C \tag{7}$$

2.3.2. *Integration of local and global similarity.* Finally, the local similarity and global similarity are integrated to avoid wrong classifications caused by the situation that only some local regions of two subjects are very similar. And the final similarity S_{all} is computed by (8) and used for face recognition.

$$S_{all} = S_G \times S_L \tag{8}$$

3. **Experimental Results.** The study reveals that PCA-Based SURF may be suitable for face recognition. We examine it with two face databases and the simulation results are shown as follows. Section 3.1 introduces the CAS-PEAL-R1 face database and the image preprocessing. The ORL (Olivetti) database is addressed in Section 3.2. Experiments with different parameters are designed in Section 3.3. In Section 3.4, we compare the results of the proposed scheme with other methods.

3.1. **CAS-PEL-R1 face database.** CAS-PEAL-R1 database [18,19] is a Chinese face image database that face images (Figure 7) are constrained in many different conditions. In the simulation, two test image sets with expression and accessory variation are chosen to perform face recognition. Each subset used is introduced as follows:

(1) Training set: There are 1200 images of 300 persons in this set. The PCA projection matrix is built by calculating the eigenspace of training set. The number of sub-regions is estimated and constructed by applying K-means algorithm to this set.

(2) Gallery set: This set contains 1040 images of 1040 persons and is utilized to be the reference image set. For evaluating face recognition performance, each inputted test image should be compared with all gallery images and classified to be one of them.

(3) Accessory test image set: This set contains 2285 images of 438 persons with accessory variation and is utilized to evaluate the recognition rate.
(4) Expression test image set: This set contains 1570 images of 377 persons with expression variation and is also utilized to evaluate the recognition rate.

(a) (b) (c) (d)

Figure 7. Sample images of two different test image sets: (a) and (b) different expression, (c) and (d) different accessory

Image preprocessing. Before doing the experiment, the images of CAS-PEAL-R1 database are preprocessed to reduce the unnecessary noise and constrain the size of image to be 75×65. For simulating the same experimental environment with [16], the CSU Face Identification Evaluation System [20] is chosen. In the following, five steps of preprocess procedures are described and the preprocessed images are shown in Figure 8.

Step 1. Convert the value of each pixel in gray level to a floating point value.
Step 2. The image is normalized geometrically to a constant scale based on the given eye coordinates, and the positions of two eyes are moved to a fixed location to reduce the registration errors.
Step 3. The elliptical mask is exploited to crop the face image for discarding part of hair, clothes, and background, and then it reserves the face region from forehead to chin and cheek to cheek.
Step 4. Apply histogram equalization to the reserved region such that the effect of illumination variation is reduced.
Step 5. The pixel values are normalized to have zero mean and unit standard deviation.

3.2. **ORL (Olivetti) face database.** This database [21] contains 40 persons and each one has 10 images with different orientations and facial expressions. Each image size is 112×92 pixels. Figure 9 shows some sample images with variant situation of one subject.

3.3. **Estimation of parameters.** In K-means clustering algorithm, there are two parameters that affect the construction of sub-regions: the number of sub-regions, k, and the manually or randomly initial cluster centers. The initial cluster centers control the distribution of sub-regions and may affect the recognition rate drastically.

The experiment is designed as testing different k ($k = 2 \sim 3$) with the initial values of cluster centers given randomly or manually. There are five tests with initial cluster centers given randomly and one test with initial cluster centers given manually.

Experimental results are shown in Table 1. "M" means the initial cluster centers are given manually, and the number (1 to 5) means the initial cluster centers are given randomly. In Table 1, the variation of recognition is shown with three different cluster types, which are illustrated in Figure 4. The weight of each sub-region is assigned a

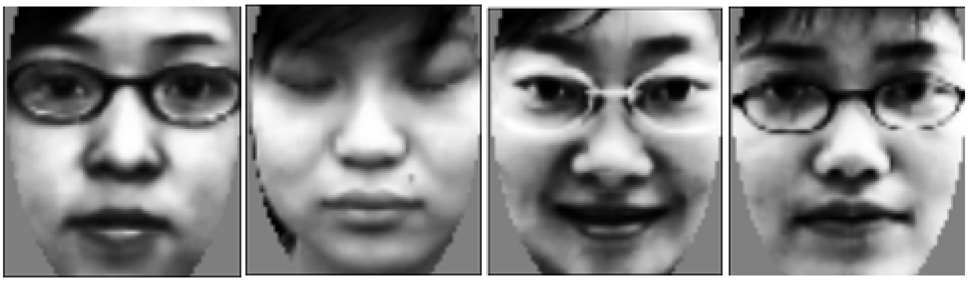


Figure 8. The preprocessed sample image

Figure 9. Some sample images in ORL face database

Table 1. Tests on parameters of K-means algorithm

k	A					
No	M	1	2	3	4	5
Recognition Rate (%)	93.2	92.0	91.7	91.5	92.5	91.2
k	B					
No	M	1	2	3	4	5
Recognition Rate (%)	91.5	91.2	90.7	90.2	90.7	91.0
k	C					
No	M	1	2	3	4	5
Recognition Rate (%)	91.7	90.0	91.5	90.2	90.2	91.2

different weight that can emphasize the sub-region with important features and conduce to the recognition rate. Here, Figure 4(a): $w_1 = \frac{8}{9}$, $w_2 = \frac{1}{9}$, Figure 4(b): $w_1 = \frac{4}{8.5}$, $w_2 = \frac{3.5}{8.5}$, $w_3 = \frac{1}{8.5}$, and Figure 4(c): w_1, w_2, and w_3 are similar to [16]. From the results in Table 1, there are no obvious differences in face recognition between different clustering types and different ways of initializing cluster centers. Therefore we choose the initial cluster centers manually in the performance comparison with other methods.

3.4. **Comparison with other methods.** In this subsection, the proposed scheme is examined to show the performance in face recognition and compared with other methods.

For comparing the results of [16], we simulate the same experimental environment of CAS-PEAL-R1 with [16]. Four hundred images of 400 persons are extracted as the gallery set, and the accessory and expression test image sets of CAS-PEAL-R1 database are selected as two test sets. Then, the accessory set enrolls 291 images to be the Test set A and the expression set enrolls 739 images to be the Test set E. Table 2 shows the simulation and comparison results.

In this table, LBP_CHI algorithm [22] uses LBP method to combine the dissimilarity measures which is Chi square statistic. SIFT_GRID method [11] uses the overlapping regular grids as the matching strategy of SIFT features, and SIFT_CLUSTER method [10] applies the K-means strategy to clustering SIFT features. PCA-SIFT_CLUSTER scheme [16] clusters PCA-SIFT features in matching stage with the K-means algorithm. Proposed-A to Proposed-C are the proposed schemes which combine three different cluster types illustrated in Figure 4.

In the case of ORL database, we select 160 images of frontal face as the gallery set and the remaining images are the test set. The ORL database contains 10 poses within ±20 degrees. For attaining better performance, we choose cluster type A in the feature clustering stage and apply cluster type A to the face recognition problem. Table 3 shows the experimental results and comparison with other methods.

In this table, SIFT_DIRECT [23] is denoted as the approach for face recognition which is based on matching standard SIFT features. SIFT_GRID [11] is denoted as the SIFT features matching with an overlapping regular grid, and SIFT_CLUSTER applies the K-means strategy to clustering SIFT features [10]. Discriminative_SIFT is denoted as the discriminative features' matching method with 16 uniform grids in face recognition [15]. Based on Tables 2 and 3, the experimental results show the proposed scheme outperforms the other methods.

Table 2. Comparison results in CAS-PEAL-R1 database

Methods	*Test Set A (Accessory)*	*Test Set E (Expression)*
LBP_CHI	91.5%	92.4%
SIFT_CLUSTER	93.1%	94.7%
SIFT_GRID	86.9%	87.9%
SIFT_PCA & Cluster	94.9%	95.9%
Proposed-A	**95.5%**	**96.5%**
Proposed-B	**95.0%**	**96.0%**
Proposed-C	**95.1%**	**96.0%**

Table 3. Comparison results in ORL database

Methods	*Test Set (ORL)*
SIFT_DIRECT	95.8%
SIFT_GRID	95.2%
SIFT_CLUSTER	95.0%
Discriminative_SIFT	95.5%
Proposed-A	**96.6%**

4. **Conclusions.** This article presents an effective face recognition method which uses PCA-SURF features to classify face images. The proposed scheme is described as follows. First, the procedure of interest point's detection is applied by carrying out the SURF detector. Second, the feature of the region around the interest point is extracted as the descriptor vector for each interest point. Third, PCA is applied to projecting the high dimensional descriptors to new feature space with low dimensions as PCA-SURF features. Fourth, K-means algorithm clusters the features in the face image. Finally, local and global similarities are computed and combined to classify face images.

Simulation results show that the performance of the proposed scheme is better than other methods. More precisely, PCA-based SURF local descriptors are more robust than original SURF and SIFT local descriptors to the accessory, expression, and pose variations.

In addition, this work has lower computational complexity. The introduction of PCA reduces the dimension of feature space. The fast indexing method is applied in matching stage and also lowers the computation time of feature matching.

Acknowledgment. The research in this article uses the CAS-PEAL-R1 face database collected under the sponsor of the Chinese National Hi-Tech Program and ISVISON Tech. Co. Ltd. The research in this article uses the ORL face database collected under the sponsor of AT&T Laboratories Cambridge.

REFERENCES

[1] M. Turk and A. Pentland, Eigenfaces for recognition, *Journal of Cognitive Neuroscience*, vol.3, no.1, pp.71-86, 1991.
[2] R. O. Duda, P. E. Hart and D. G. Stork, *Pattern Classification*, 2nd Edition, Wiley Interscience, 2001.
[3] P. N. Belhumeur, J. P. Hespanha and D. J. Kriegman, Eigenfaces vs fisherfaces: Recognition using class specific linear projection, *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol.19, no.7, pp.711-720, 1997.
[4] W. S. Lee, H. J. Lee and J. H. Chung, Wavelet-based FLD for face recognition, *Proc. of the 43rd IEEE Midwest Symposium, Circuits and Systems*, vol.2, pp.734-737, 2000.
[5] H. K. Ekenel and R. Stiefelhagen, Analysis of local appearance-based face recognition: Effects of feature selection and feature normalization, *In CVPR Biometrics Workshop*, pp.34-34, 2006.
[6] T. Ojala, M. Pietikainen and T. Maenpaa, Multiresolution gray-scale and rotation invariant texture classification with local binary patterns, *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol.24, no.7, pp.971-987, 2002.
[7] C. Liu and H. Wechsler, A Gabor feature classifier for face recognition, *Proc. of the 8th IEEE International Conference on Computer Vision*, vol.2, pp.270-275, 2001.
[8] S. Yan, H. Wang, X. Tang and T. S. Huang, Exploring feature descriptors for face recognition, *International Conference on Acoustics, Speech, and Signal Processing*, vol.1, pp.629-632, 2007.
[9] D. Lowe, Distinctive image features from scale-invariant keypoints, *Int. J. Computer Vision*, vol.60, no.2, pp.91-110, 2004.
[10] M. Bicego, A. Lagorio, E. Grosso and M. Tistarelli, On the use of SIFT features for face authentication, *Proc. of the Conference on Computer Vision and Pattern Recognition Workshop*, p.35, 2006.
[11] J. Luo, Y. Ma, E. Takikawa, S. Lao, M. Kawade and B. Lu, Person-specific SIFT features for face recognition, *IEEE Int. Conference on Acoustics, Speech and Signal Processing*, pp.593-596, 2007.
[12] 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*, vol.2, no.1, pp.506-513, 2004.
[13] K. Mikolajczyk and C. Schmid, A performance evaluation of local descriptors, *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol.27, no.10, pp.1615-1630, 2005.
[14] H. Bay, A. Ess, T. Tuytelaars and L. Van Gool, Speeded-up robust features, *Computer Vision and Image Understanding*, vol.110, no.3, pp.346-359, 2008.
[15] A. Majumdar and R. K. Ward, Discriminative SIFT features for face recognition, *Canadian Conference on Electrical and Computer Engineering*, pp.27-30, 2009.
[16] S. D. Lin, J.-H. Lin and C.-C. Chiang, Using gradient features from scale-invariant keypoints on face recognition, *International Journal of Innovative Computing, Information and Control*, vol.7, no.4, pp.1639-1649, 2011.
[17] *http://www.maths.lth.se/matematiklth/personal/petter/surfmex.php.*
[18] B. Cao, S. Shan, X. Zhang and W. Gao, Baseline evaluations on the CAS-PEAL-RI face database, *Advances in Biometric Person Authentication, LNCS*, vol.3338, pp.370-378, 2004.
[19] W. Gao, B. Cao, S. Shan, X. Chen, D. Zhou, X. Zhang and D. Zhao, The CAS-PEAL large-scale Chinese face database and baseline evaluations, *IEEE Trans. System, Man and Cybernetics (Part A)*, vol.38, no.1, pp.149-161, 2008.
[20] D. S. Bolme, J. R. Beveridge, M. Teixeira and B. A. Draper, The CSU face identification evaluation system: Its purpose, features and structure, *Proc. of Int. Conf. Vision Systems*, pp.304-311, 2003.
[21] *http://www.cl.cam.ac.uk/research/dtg/attarchive/facedatabase.html.*

[22] T. Ahonen, A. Hadid and M. Pietikainen, Face recognition with local binary patterns, *Proc. of European Conference on Computer Vision*, pp.469-481, 2004.
[23] M. Aly, Face recognition using SIFT features, *CNS/Bi/EE Report 186*, pp.3277-3280, 2006.