HIGH TECHNOLOGY LETTERS | Vol. 28 No. 3 | Sep. 2022 | pp. 280-287

doi:10.3772/j.issn.1006-6748.2022.03.007

Research on will-dimension SIFT algorithms for multi-attitude face recognition①

SHENG Wenshun (圣文顺)②*, SUN Yanwen**, XU Liujing*

(* Pujiang Institute, Nanjing Tech University, Nanjing 211200, P. R. China)

(** School of Information Engineering, Nanjing Audit University, Nanjing 211815, P. R. China)

Abstract

The results of face recognition are often inaccurate due to factors such as illumination, noise intensity, and affine/projection transformation. In response to these problems, the scale invariant feature transformation (SIFT) is proposed, but its computational complexity and complication seriously affect the efficiency of the algorithm. In order to solve this problem, SIFT algorithm is proposed based on principal component analysis (PCA) dimensionality reduction. The algorithm first uses PCA algorithm, which has the function of screening feature points, to filter the feature points extracted in advance by the SIFT algorithm; then the high-dimensional data is projected into the low-dimensional space to remove the redundant feature points, thereby changing the way of generating feature descriptors and finally achieving the effect of dimensionality reduction. In this paper, through experiments on the public ORL face database, the dimension of SIFT is reduced to 20 dimensions, which improves the efficiency of face extraction; the comparison of several experimental results is completed and analyzed to verify the superiority of the improved algorithm.

Key words: face recognition, scale invariant feature transformation (SIFT), dimensionality reduction, principal component analysis-scale invariant feature transformation (PCA-SIFT)

0 Introduction

Face recognition is an important branch of computer vision research. It is a subject that belongs to both the field of biometric recognition and artificial intelligence, and also one of the most successful applications of image analysis and understanding, involving computer graphics, pattern recognition, artificial intelligence, image processing and other related technologies[1-3]. The feasibility technology of face recognition has been studied for more than 30 years, and has been widely used in business, security, identity authentication, law enforcement and so on. This technology has attracted more and more attention, and has gradually become a dynamic research field.

Face recognition can be divided into frontal face recognition and multipose face recognition according to whether the pose changes or not[4]. The former has high requirements for frontal face posture and is very sensitive to posture changes, the robustness[5] of the system is often difficult to guarantee; the latter uses multiple face images of multiple vision to build a multi-sample and multi-posetraining set[6], and adds a feature fusion process after feature extraction[7], so the recognition rate is high and more universal. However, further in-depth researches and improvements are needed for dynamic face recognition and detection as well as for faces under complex multi background. Face recognition is an important research field in biometric recognition technology. Due to the changes of emotion and environment in the process of facial image acquisition, such as multi pose expressions, changes in light intensity, change in noise intensity, rotation, scaling scale, occlusions on the face[8-9] (glasses, sunglasses, hats, wound dressings, etc.), age change, accurate and powerful face recognition system is facing great challenges. The scale invariant feature transformation (SIFT) algorithm[10-12] is based on the potential of the image on the scale space to be invariant to the scale and selection points of interest, which is independent of image size, scale scaling and rotation[13]. The algorithm solves the above problem by using Gaussian differential function to find extreme points. SIFT feature is based

① Supported by the National Natural Science Foundation of China (No. 61571222), the Natural Science Research Program of Higher Education Jiangsu Province (No. 19KJD520005), Qing Lan Project of Jiangsu Province (Su Teacher's Letter [2021] No. 11) and Jiangsu Graduate Scientific Research Innovation Program (No. KYCX21_1944).

② To whom correspondence should be addressed. E-mail: gisma@qq.com.

Received on Aug. 30, 2021

on some local appearance interest points on the object. The processed image can detect the face under the changes of illumination intensity, noise intensity and microscopic angle of view, with high stability. In addition, the detection rate of partial object masking described by SIFT features is also quite high, and even more than three sift object features are enough to calculate the position and direction. However, although SIFT algorithm can deal with face clearly, it has high dimensionality, high computational complexity and cumbersome computational process, which seriously affects the realtime performance of face recognition[14]. The classical SIFT algorithm has a large amount of computation, and each feature point contains $4 \times 4 \times 8 = 128$ feature description vectors. Extracting too many feature points in areas with complex details is prone to produce false matching. For the realtime performance of target tracking, the efficiency of SIFT algorithm is extremely low. In this paper, a SIFT feature extraction and dimensionality reduction algorithm based on principal component analysis (PCA)[15] is proposed. On the original basis of the SIFT algorithm in face recognition, the dimensionality reduction algorithm is used to reduce the dimensionality, simplify the tedious calculation process and reduce the calculation complexity, and improve the efficiency of face recognition. The improved SIFT algorithm has obvious advantages over the original SIFT algorithm in that it can speed up the processing of massive data while maintaining the illumination invariance, rotation invariance, scale invariance and occlusion invariance of the original algorithm.

1 Face recognition and SIFT algorithm principle

1.1 Face recognition

The face recognition system can be roughly divided into the following four stages, namely, face image acquisition and detection, face image preprocessing, face image feature extraction, face image matching and recognition[16]. Firstly, the basic process is to collect images containing facial information through cameras, scanners, video streams, etc., and use face detection technology to detect or track the faces in the images, judge the location of the faces in the images, determine and complete accurate positioning of faces. Secondly, the face image preprocessing technology is used to denoise the image obtained in the previous stage of positioning and normalize the illumination to improve the face quality, so as to facilitate the extraction of face features in the next stage. Thirdly, the face image feature extraction technology is used to extract features of the face image after the second stage of correction preprocessing, that is, to extract the features of the main parts of the face image such as eyes, ears, mouth, and nose, and perform classification and recognition. Finally, using the face image matching and recognition technology, the feature data of the face image extracted in the previous stage and the feature data collected and extracted from the face database image are searched, recognized and matched, so as to distinguish, judge and recognize the human identity. The specific process is shown in Fig. 1.

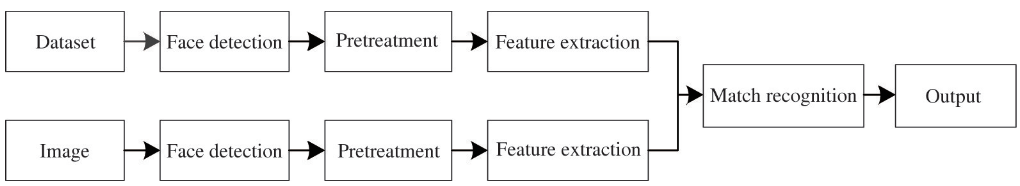


Fig. 1 The process of face recognition

As an important biological feature, human face has the following advantages in the fields of electronic information security, security protection, human trace tracking, human-computer interaction and so on.

(1) Universality: a biological characteristic possessed by all.

(2) Directness and operability: it is convenient to obtain the face exposed, and the cost of acquisition equipment is low. The face image can be obtained by only the camera, and the operation is simple.

(3) Concealment: undetectable to the identifier at the time of authentication.

(4) Non-contact: it is obtained in a non-contact manner through shooting equipment, which is easy to be accepted and promoted.

(5) Security and reliability: in addition to medical factors, facial features are stable, not easy to lose and difficult to forge.

1.2 Classical SIFT algorithm

1.2.1 Basic principle of algorithm

Based on the image in scale space, SIFT algorithm uses Gaussian differential function to identify potential points of interest that are invariant to scale and selection. It is a very stable local feature descriptor algorithm. The local feature descriptor extracted by SIFT algorithm has strong tolerance for many factors such as the change of illumination intensity, noise intensity, rotation, scaling scale, change of micro vision angle and so on[17].

1.2.2 SIFT algorithm steps

(1) Establish scale space through Gaussian convolution.

(2) Construct a Gaussian difference pyramid to detect extreme points, and verify the extreme points by traversal.

(3) Delete extreme points of edge response and low contrast.

(4) Determine the direction of the feature point through the gradient histogram.

(5) Create feature descriptors.

The basic flow of SIFT algorithm implementation is shown in Fig. 2.

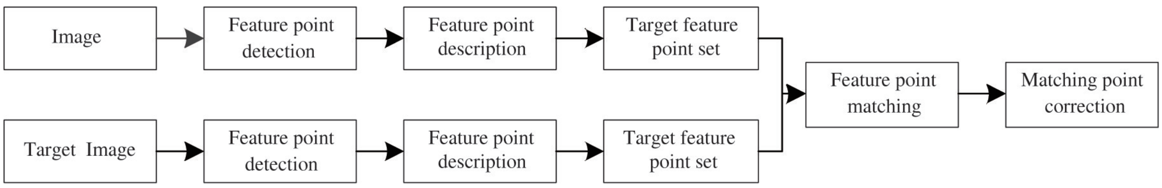


Fig. 2 Basic flow of SIFT algorithm

2 Dimension reduction algorithm and improved SIFT algorithm based on dimension reduction algorithm

2.1 Dimensionality reduction algorithm

When processing an image, a large number of feature points or some more complex feature points are extracted to improve accuracy, but each feature point will generate 128-dimensional feature descriptor. Assuming that the database needs to store an image, SIFT algorithm will decompose it into 400 feature points when processing the image, which will occupy a lot of memory and slow down the retrieval speed, which will seriously affect the efficiency. In order to improve this characteristic of SIFT algorithm, it is necessary to introduce dimensionality reduction technology to reduce the dimension of 120-dimensional feature vector to 64 dimensions, or even 32 dimensions. The commonly used dimensionality reduction techniques are principal component analysis (PCA) algorithm and linear discriminant analysis (LDA) algorithm. In this study, the PCA algorithm is more appropriate.

PCA is one of the most commonly used linear dimensionality reduction techniques[18], which takes many samples of relevant input features as inputs and outputs orthogonal features as linear combinations of input features. It is an effective method for information processing, compression and extraction, and has high efficiency in dimensionality reduction and feature extraction. In essence the PCA method is based on the K-L transformation principle[19], which extracts the features of the data to form the feature space, and then projects the test set to the feature space to obtain the projection coefficient[20]. In other words, the high-dimensional original image is mapped to the low-dimensional space through the projection method, retaining as many pixels of the original image as possible and removing the pixels with a small amount of information. The larger the variance of the mapped data, the more feature points of the original image are retained. Studies have proved that PCA is a linear dimensionality reduction technique that has the least number of pixels lost in the original image. Therefore, PCA algorithm is used to reduce the dimensionality of the image.

The main idea of PCA is to construct a series of linear combinations of original features to form low dimensional features. In this way, it can not only remove the correlation of the data, but also keep the variance information of the original data to the greatest extentafter dimensionality reduction.

The core of the algorithm is to find the maximum variance. Suppose that $\boldsymbol{W}$ is the projection matrix, then the projection of the image x in the new coordinates is $\boldsymbol{W}^{\mathrm{T}}x$, the variance is $\boldsymbol{W}^{\mathrm{T}}xx^{\mathrm{T}}\boldsymbol{W}$, and the objective function is

$$\arg\max \operatorname{tr}(\boldsymbol{W}^{\mathrm{T}}\boldsymbol{X}\boldsymbol{X}^{\mathrm{T}}\boldsymbol{W}) \quad \text{s.t. } \boldsymbol{W}^{\mathrm{T}}\boldsymbol{W} = I \tag{1}$$

Then use the Lagrangian formula to solve Eq. (1):

$$\boldsymbol{X}\boldsymbol{X}^{\mathrm{T}}\boldsymbol{W} = (-\lambda)\boldsymbol{W} \tag{2}$$

Among them $-\lambda$ is the eigenvalue of $\boldsymbol{X}\boldsymbol{X}^{\mathrm{T}}$, to reduce the image from n-dimension to k-dimension, the eigenvectors corresponding to k eigenvalues and the matrix $\boldsymbol{W}$ composed of k eigenvectors are needed.

2.2 Reduced dimensional improvement of the SIFT algorithm

2.2.1 Principle of PCA-SIFT algorithm

Although SIFT algorithm is applied to face recognition to solve the problems of image rotation, scale scaling, illumination changes, noise interference and so on, due to the existence of high-dimensional (128-

dimensional) feature vectors, the computational complexity is high, the recognition process takes a long time and the realtime performance is not high. In order to solve this problem, a SIFT algorithm based on PCA dimensionality reduction is proposed. Using the PCA algorithm with the function of filtering feature points, the feature points extracted in advance by SIFT algorithm are filtered, and the high-dimensional data are projected into the low-dimensional space to remove redundant feature points, thereby changing the way of generating feature descriptors and achieving the effect of dimensionality reduction, as shown in Fig. 3.

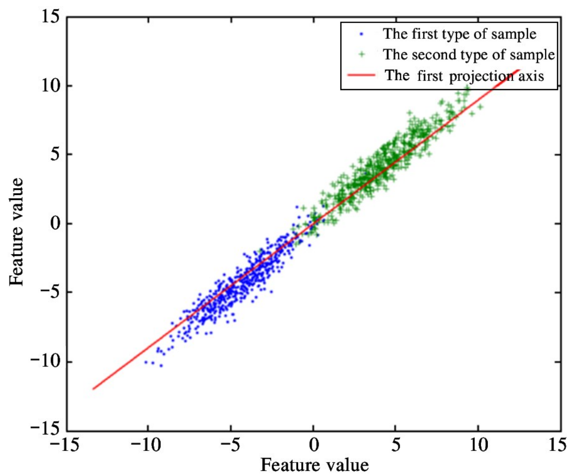


Fig. 3 Projection target of PCA algorithm

2.2.2 Calculate the projection matrix

The calculation of the PCA-SIFT projection matrix mainly involves the following steps.

(1) Use the SIFT algorithm to extract the feature points in the training set, and the number is recorded as n.

(2) Divide a 41×41 neighborhood window and rotate the coordinate axis to the main direction of each feature point when generating a feature descriptor for each feature point. Remove the edge pixels of the four corners, calculate the gradient of each remaining pixel in the horizontal and vertical directions, and obtain a description vector of $39 \times 39 \times 2 = 3042$.

(3) Calculate the average vector of the 3042-dimensional description vector of the n feature points, and subtract the average vector from the 3042-dimensional vector of the n feature points, and put the resulting difference into the $n \times 3042$ matrix $\boldsymbol{W}$.

(4) Calculate the covariance matrix Cov$\boldsymbol{W}$ of the matrix $\boldsymbol{W}$, and calculate the eigenvalues and eigenvectors of the covariance matrix Cov$\boldsymbol{W}$[21].

(5) Sort the obtained eigenvalues from large to small, and select the eigenvectors corresponding to the first k eigenvalues as the principal component directions to form a $k \times 3042$ projection matrix, where k is the dimension of the PCA-SIFT feature descriptor[22].

The above steps (3) − (5) are pre-calculated, that is, the projection matrix is calculated in advance through the same type of image set using the PCA principle. After the feature points in the image are calculated to obtain the 3042-dimensional feature descriptor, it only needs to be multiplied by the projection matrix to achieve the effect of dimensionality reduction. Regarding the selection of n of the projection matrix, the value of n can be fixed as needed, or the size of n can be automatically determined according to the percentage of the eigenvalue energy value of the covariance matrix. The flow chart for its generation of PCA-SIFT descriptors is shown in Fig. 4.

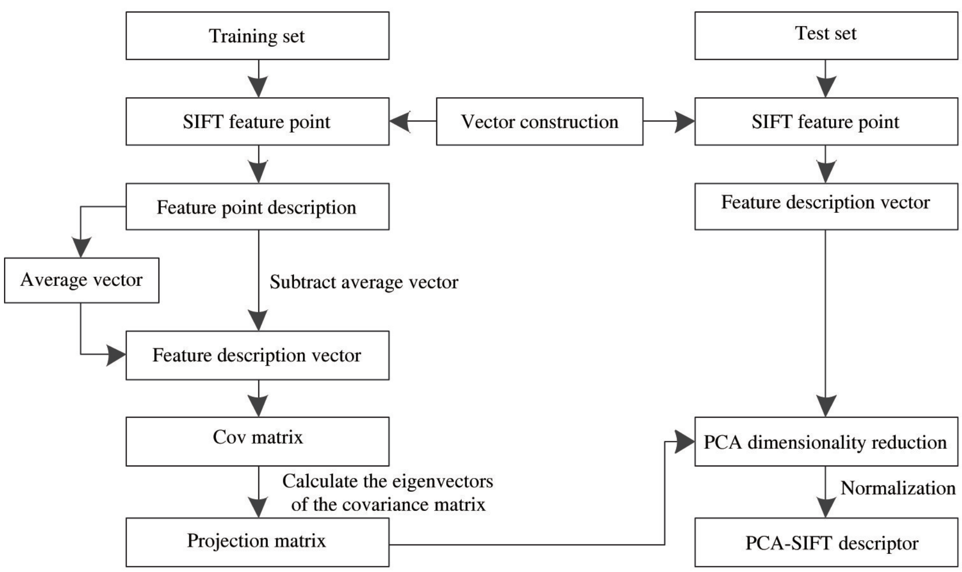


Fig. 4 Flow chart for generating PCA-SIFT descriptors

2.2.3 Generate PCA-SIFT feature descriptor

The steps to generate the PCA-SIFT feature descriptor are as follows.

(1) Rotate the coordinate axis to the main direction of the feature point, divide a 41 × 41 neighborhood window, and calculate the horizontal and vertical gradient of each pixel after removing the edge corner pixels, that is, a 39 × 39 × 2 = 3042 dimensional description vector.

(2) Multiply the vector by the projection matrix obtained in the previous section to obtain a k-dimensional eigenvector.

3 Discussion

In this paper, Python is used as the experimental test language, and the ORL face database[23] is selected for experimental comparison. The ORL face database contains 40 people, each of whom has 10 gray images with different expressions, gestures or occlusion to varying degrees. The number of the pictures is 92 112. This is currently one of the most used face databases for research on face recognition. Fig. 5 shows some of the images in the ORL face database.

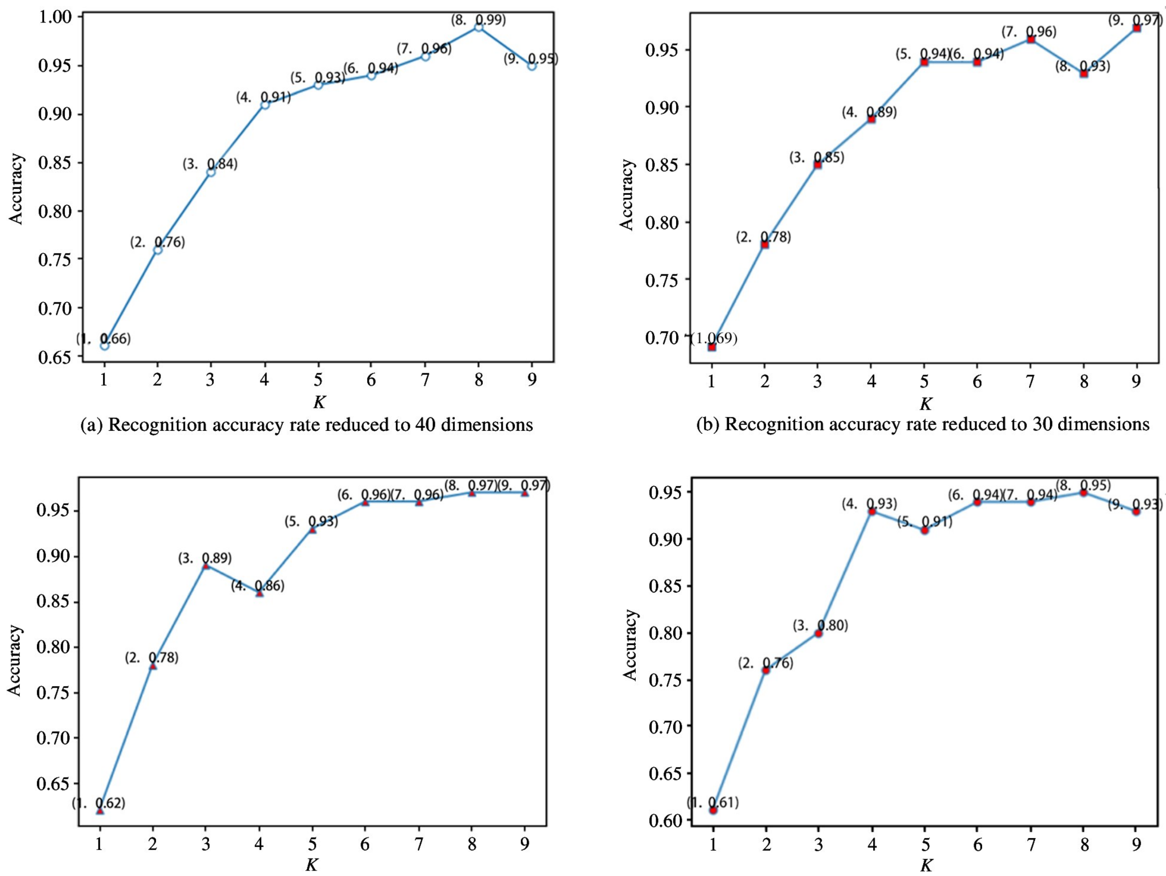
Fig. 5 Partial face images in ORL face database

3.1 Determination of the optimal dimension *k* value

Before image processing, PCA algorithm is selected to reduce the dimensionality of the image, and the accuracy rates when the dimensionality is reduced to 40, 30, 20, and 10 respectively are shown in Fig. 6.

In summary, the recognition rates of different dimensions are shown in Fig. 7.

In this article, $k = 20$ is set, which means that the 128-dimensional feature descriptor of SIFT will be reduced to the 20-dimensional PCA-SIFT feature descriptor, and then the experimental results are analyzed and compared.



(a) Recognition accuracy rate reduced to 40 dimensions

(b) Recognition accuracy rate reduced to 30 dimensions

(c) Recognition accuracy rate reduced to 20 dimensions

(d) Recognition accuracy rate reduced to 10 dimensions

Fig. 6 Comparison of recognition accuracy of different reduced dimensions

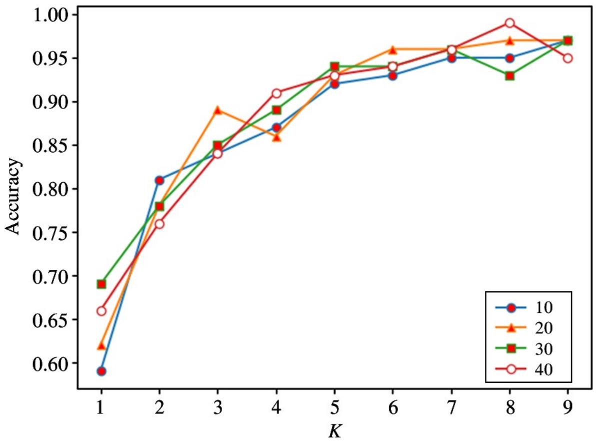


Fig. 7 Synthetic graph of recognition accuracy of different dimensions

3.2 Algorithm test and algorithm comparison

In view of the influence of different expressions, posture changes, obstructions, scaling and other factors, the following four sets of experiments show the effects of the algorithm, as shown in Fig. 8.

The algorithm was tested on the ORL face library by selecting i ($i = 1, 2, \cdots, 8, 9$) images from each person as training samples, training the support vector machine (SVM) under the optimal parameters to obtain the SVM model, and the remaining j ($j = 10 - i$) images as test samples, labelled in turn as $i + 1$ images, $i + 2$ images, $\cdots$, 10 images. The test results are shown in Table 1.

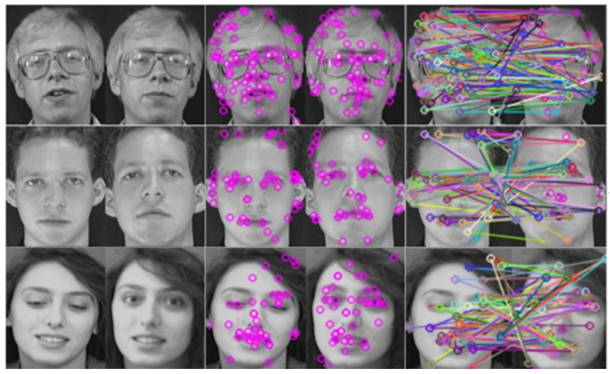


Fig. 8 Generating rendering of feature descriptor under multiple factor interference

Table 1 Test results (the recognition rate/%)

Training sample	Image 2	Image 3	Image 4	Image 5	Image 6	Image 7	Image 8	Image 9	Image 10
1 piece /person	90.00	90.00	80.00	90.00	80.00	90.00	90.00	90.00	80.00
2 piece /person		100.00	90.00	90.00	90.00	90.00	90.00	100.00	80.00
3 piece /person			100.00	90.00	90.00	90.00	90.00	100.00	80.00
4 piece /person				100.00	100.00	90.00	90.00	90.00	90.00
5 piece/person					100.00	100.00	100.00	90.00	90.00
6 piece/person						100.00	100.00	100.00	100.00
7 piece /person							100.00	100.00	100.00
8 piece /person								100.00	100.00
9 piece /person									100.00

It can be seen from Table 1 that when the number of training samples exceeds 6 pieces/person, the recognition rate is 100%; when the number of training samples is only 1 piece/person, the recognition rate can reach 85%, and the average recognition rate of the algorithm is 97%. The average recognition rate is better than the face recognition rate in Refs[24-27]. It can be seen that the algorithm has a higher recognition rate than the general SVM algorithm and PCA algorithm. The comparison of different face recognition algorithms is shown in Table 2.

Table 3 shows the comparison of the advantages and disadvantages of the SIFT algorithm and the PCA-SIFT algorithm.

The comparison of the operation time of the two algorithms is shown in Table 4. The comprehensive time-consuming of the PCA-SIFT algorithm is significantly lower than that of the SIFT algorithm.

Table 2 Comparison of different face recognition algorithms

Face recognition algorithm	Recognition rate/%
KICA + PCA	93.00
PCA + SVM	93.50
PCA + PSO	91.00
PCA + AdaBoost	96.33
SIFT	90.13
PCA + SIFT	97.00

4 Conclusions

By comparing the PCA-SIFT algorithm with the traditional SIFT algorithm, the descriptors obtained by the PCA-SIFT algorithm under the conditions of rotation, scale transformation, perspective transformation,

Table 3 Comparison of advantages and disadvantages between SIFT algorithm and PCA-SIFT algorithm

Algorithm	Channel	Advantage	Disadvantage
SIFT	128	Easy to understand, easy to develop, and applicable to a wide range	High dimensionality, complex and tedious calculation
PCA-SIFT	Variable, can be 20 or less	The dimensionality is reduced while basically maintaining the invariance, thereby reducing the calculation time	The projection matrix is only suitable for images at the same scale

Table 4 Time consuming comparison between SIFT algorithm and PCA-SIFT algorithm

Algorithm	SIFT		PCA-SIFT	
Time/s	Generate feature descriptor 1.56	Image matching 2.31	Generate feature descriptor 1.63	Image matching 0.56
Sum/s	3.87		2.19	

and feature recognition and matching, the time consumed to generate feature descriptors is almost the same as that of traditional SIFT algorithms. However, the time consumed in image matching is much less than that of traditional SIFT algorithms. It can be seen that the descriptors generated by PCA-SIFT are of higher quality and faster.

PCA is used to reduce the dimension of the SIFT algorithm, which greatly improves the efficiency and speed of face recognition. Face recognition can be applied to security systems such as surveillance to improve the real-time as well as timeliness of the system.

References

[1] JAYASIMHA Y, REDDY R. A robust face emotion recognition approach through optimized SIFT features and adaptive deep belief neural network[J]. *Intelligent Decision Technologie*, 2019, 13(3): 379-390

[2] RUAN S, TANG C, XU Z, et al. Multi-pose face recognition based on deep learning in unconstrained scene[J]. *Applied Sciences*, 2020, 10(13): 4669-4685

[3] GUAN Y, FANG J, WU X. Multi-pose face recognition using cascade alignment network and incremental clustering[J]. *Signal, Image and Video Processing*, 2020, 15(1): 63-71

[4] HOSSEIN-NEJAD Z, AGAHI H, MAHMOODZADEH A. Image matching based on the adaptive redundant keypoint elimination method in the SIFT algorithm[J]. *Pattern Analysis and Applications*, 2021, 24(1):669-683

[5] SHI Q, LIU Y, XU Y C. Image feature point matching method based on improved BRISK[J]. *International Journal of Wireless and Mobile Computing*, 2021, 20(2): 132-138

[6] GAO Z Y. Basic principle and application of face recognition technology[J]. *Electronic Production*, 2019, 1848(1): 41-42 (In Chinese)

[7] QIAO M Y, LIANG X, CHEN M J. Improved SIFT algorithm based on image filtering[J]. *Journal of Physics: Conference Series*, 2021, 1848(1): 1-6

[8] LIN W P, TSAI S. Research on volleyball image classification based on artificial intelligence and SIFT algorithm [J]. *Mathematical Problems in Engineering*, 2021, 39(9): 1-10

[9] LI H L, WANG Z, HUANG Y J. Face recognition method based on image features and face pose[J]. *Science Technology and Engineering*, 2018, 18(31): 195-199 (In Chinese)

[10] ALMUTLAQ R A, ALJUTAILI D S, ALHARBI S A, et al. The impact of using SR-SIFT algorithm on various banknotes[J]. *International Journal of Data Science*, 2020, 5(2): 151-159

[11] LIN Z D, DU Y Z. Performance comparison of sift and surf image matching algorithms under specific conditions [J]. *Journal of Xiamen University of Technology*, 2019, 27(3): 30-36 (In Chinese)

[12] LI Y S, FAN LD, XIE W X. TGSIFT: robust SIFT descriptor based on tensor gradient for hyperspectral images [J]. *Chinese Journal of Electronics*, 2020, 29(5): 916-925

[13] SRIKANTH M V, PRASAD V, PRASAD S K. An improved firefly algorithm-based 2-D image thresholding for brain image fusion[J]. *International Journal of Cognitive Informatics and Natural Intelligence*, 2020, 14(3): 60-96

[14] LI H. Feature extraction of face image based on convolution neural network[J]. *Journal of Henan University of Engineering*, 2019, 31(1): 66-70 (In Chinese)

[15] WANG C, YANG F, WANG H, et al. Urban house detection using SAM and SIFT on hyperspectral remote sensing images[J]. *Journal of Physics: Conference Series*, 2019, 1237(3): 29-33

[16] SIMPSON E A, MAYLOTT S E, LEONARD K, et al. Face detection in infants and adults: effects of orientation and color[J]. *Journal of Experimental Child Psychology*, 2019, 186(1): 17-32

[17] YANG C. Face Image Beautification and 3D Reconstruction Application Based on Feature Fusion[D]. Hangzhou: School of Informatics Science and Technology, Zhejiang University of Technology, 2019:20-26 (In Chinese)

[18] NOOR D, LI Y, LI Z, et al. Multi-scale gradient image super-resolution for preserving key SIFT points in low-resolution images[J]. *Signal Processing: Image Communication*, 2019, 78(1): 236-245
[19] ZHOU J X, XU M Y, LIU J Q. Image registration method based on improved SIFT algorithm[J]. *Industrial Control Computer*, 2019, 32(5): 97-98 (In Chinese)
[20] YU C Y, YAO C T, PEI M T, et al. Diffusion-based kernel matrix model for face liveness detection[J]. *Image and Vision Computing*, 2019, 89(1): 88-94
[21] CHEN Z H, WU Y D, CAI G R, et al. Face quality assessment algorithm based on texture feature fusion[J]. *Journal of Jimei University*, 2018, 23(4): 312-320 (In Chinese)
[22] DALAL S, VISHWAKARMA V P. Optimization of weights in ELM for face recognition[J]. *Journal of Information and Optimization Sciences*, 2021, 42(6): 1337-1352
[23] DU M Y. Mobile payment recognition technology based on face detection algorithm[J]. *Concurrency and Computation: Practice and Experience*, 2018, 30(22): 4655-4663
[24] CAO W F. Face recognition based on hierarchical SVM and KICA[J]. *Journal of Nanjing Institute of Technology*, 2019, 17(3): 76-79 (In Chinese)
[25] ZHU Q J, WANG H L, ZHANG G H. Research on face recognition algorithm based on fast PCA and K-CV optimized SVM[J]. *Journal of Hubei Minzu University (Natural Science Edition)*, 2021, 39(2): 193-198 (In Chinese)
[26] LIN Z M. Research on an improved PSO face recognition algorithm[J]. *Natural Science Journal of Harbin Normal University*, 2020, 36(4): 51-55 (In Chinese)
[27] WANG H Y, MENG X Y, ZHAO Y, et al. Reaserach and implementation of embedded face recognition based on AdaBoost and PCA[J]. *Transducer and Microsystem Technologies*, 2017, 36(6): 59-62 (In Chinese)

SHENG Wenshun, born in 1979. He received his M. S. degree from Wuhan University in 2014. He also received his B. S. degree from Dalian University of Technology in 2002. His research interests include image processing, computer vision and machine learning.