

# Face Recognition Method Based on Improved Gabor Wavelet Transform Algorithm

Jie-sheng Wang, Yan-lang Ruan, Bo-wen Zheng, Shu-zhi Gao

**Abstract**—A face recognition method was proposed based on the improved Gabor wavelet transform algorithm. Firstly, the optimized SSR algorithm was used to carry out the preprocessing on face images. Then the Gabor wavelet transform and the discrete cosine transform (DCT) are combined to form the improved Gabor wavelet transform, which is used to realize the image feature extraction and dimension reduction. Finally, the nearest neighbor distance detector and Euclidean distance are used as the classifier to realize the recognition and classification of face images, which can effectively shorten the recognition time and maintain a considerable recognition rate.

**Index Terms**—face recognition, gabor wavelet transform, discrete cosine transform

## I. INTRODUCTION

FACE recognition is one of the most active research topics in the computer vision field in recent years [1-2]. After many years of efforts, the automatic face recognition has achieved gratifying results in various environments. The research of face recognition began in the middle and late 1960s [3-4]. Early face recognition algorithms are mainly based on geometric features [5], template matching [6], and the learning-based methods [7-8]. The methods based on geometric features mainly used the geometrical parameters, such as the size, location, angle and distance of facial features, to carry out face identifications [9]. The template matching method is used to identify the face by comparing the relevant matching images with the standard template.

With the advent of high-performance computers, scholars are more open-minded in the research of face recognition, and it has greatly accelerated the speed of face recognition

Manuscript received June 15, 2018; revised August 29, 2018. This work was supported by the Basic Scientific Research Project of Institution of Higher Learning of Liaoning Province (Grant No. 2017FWDF10), the Project by Liaoning Provincial Natural Science Foundation of China (Grant No. 20180550700) and the Opening Project of National Financial Security and System Equipment Engineering Research Center (Grant No. USTLKFGJ201502).

Jie-Sheng Wang is with the School of Electronic and Information Engineering, University of Science and Technology Liaoning, Anshan, 114051, PR China; National Financial Security and System Equipment Engineering Research Center, University of Science and Technology Liaoning. (phone: 86-0412-2538246; fax: 86-0412-2538244; e-mail: wang\_jiesheng@126.com).

Yan-lang Ruan is with the School of Electronic and Information Engineering, University of Science and Technology Liaoning, Anshan, 114051, PR China (e-mail:ruanyanlang@163.com).

Bo-Wen Zheng is with the School of Electronic and Information Engineering, University of Science and Technology Liaoning, Anshan, 114051, PR China (e-mail:18842038281@163.com).

Shu-Zhi Gao is with Shenyang University of Chemical Technology, Shenyang, 110000, PR China (e-mail: szg6868@126.com).

research. From 1990 to 1997, the face recognition has formed several mainstream research directions: Eigenface method [10] based on the principal component analysis, Fisherface method [11] based on Fisher linear discriminant analysis, elastic graph matching method [12] and local feature analysis method [13]. As a kind of successful face recognition technology, the Eigenface method has raised the second climax of face recognition research. Along with the method of eigenface, the multi-subspace face recognition method based on singular value decomposition, independent component analysis, factor encoding and kernel are proposed. These methods are better than the traditional Eigenface method.

After 1998, scholars began to focus on the extraction of facial features under non-ideal conditions (such as skin color, posture, light, facial expressions, etc.). At this stage, a lot of new or improved algorithms emerged: two-dimensional principal component analysis method, two-dimensional linear discriminant analysis method, face detection method based on Adaboost, method based on deformable parts model (DPM), method based on convolutional neural network [14], method based on sparse representation [15], etc. Similar to the eigenface method, there are many different transform methods to extract the global features for face recognition: factor encoding [16], wavelet transform [17], discrete cosine transform [18], etc. On the choice of classifiers, the commonly used classifiers are based on neural network [19-20], support vector machine [21], nearest neighbor classifier [22], hidden Markov model [23], etc.

The kernel function of the Gabor wavelet is very similar to the response of the primary visual cortical sensory field cells of mammals, which can extract the features of target images in different spatial positions, frequencies and directions [24]. The feature extracted by Gabor wavelet can effectively overcome the global interference, such as scale, angle, illumination and so on. So the Gabor wavelet transform is widely used in the field of texture feature analysis and target recognition [25-26]. The earliest Gabor wavelet-based face recognition method is to use a set of two-dimensional Gabor wavelet filters (feature extraction Gabor wavelet filter group)[27] for the extraction of face image feature subspace data. Because of its excellent recognition performance, the face recognition method based on Gabor wavelet has received great attention from the academic circle, and a series of improved algorithms have emerged. This paper proposes a combination of Gabor wavelet transform and DCT algorithm [28]. The image compression is carried out based on 2D DCT algorithm and the Gabor wavelet transform processes the face images, which greatly reduces the workload of Gabor wavelet filter, effectively reduces the work time and improves the

recognition rate.

## II. FACIAL IMAGE PREPROCESSING

### A. Histogram Equalization

Histogram equalization, also known as gray-scale equalization, refers to the output of the balanced histogram, that is to say each gray-scale pixel is almost equal. After the equalization of the picture, the pixel distribution is symmetrical, and the contrast ratio and the dynamic range are larger than that before the treatment. The distance within the uninterrupted interval of 0-1 is considered first. The normalized histogram of images can be represented by a probability density function.

$$p(x), 0 \leq x \leq 1 \quad (1)$$

In the 0-1 interval, the sum of all the probabilities is 1.

$$\int_{x=0}^1 p(x) = 1 \quad (2)$$

The function before conversion is  $p_r(r)$ , the converted function is  $p_s(s)$ , and the gray mapping relation is  $s = f(r)$ . The relationship among them can be described as follows.

$$p_s(s) = p_r(r) \cdot \frac{dr}{ds} \quad (3)$$

The converted function satisfies  $p_s(s) = 1, 0 \leq s \leq 1$ , where the histogram is balanced at this time. The following equation should be met.

$$p_r(r) = \frac{dr}{ds} \quad (4)$$

Integrate  $r$  on both sides to obtain:

$$s = f(r) = \int_0^r p_r(\mu) d\mu \quad (5)$$

Eq. (5) shows the cumulative distribution function of an image. Eq. (5) is assumed to be the cumulative distribution function in the gray scale at this stage of 0-1. If it has all gray scale [0,255], Eq. (5) is multiplied by the  $D_{\max}$ . Of course, for images, this coefficient is obviously 255. Therefore, the transformation relationship can be expressed as follows.

$$D_B = f(D_A) = D_{\max} \int_0^{D_A} p_{D_A}(\mu) d\mu \quad (6)$$

where  $D_B$  is the converted gray value and  $D_A$  is the gray value before conversion.

If the gray level is discrete, it can be expressed as:

$$D_B = f(D_A) = \frac{D_{\max}}{A_0} \sum_{i=0}^{D_A} H_i \quad (7)$$

Where  $A_0$  is the entire area of the image, also known as a collection of all pixels.  $H_i$  is the number of pixels in the  $i$  th gray level. Histogram equalization effect is shown in Fig. 1.

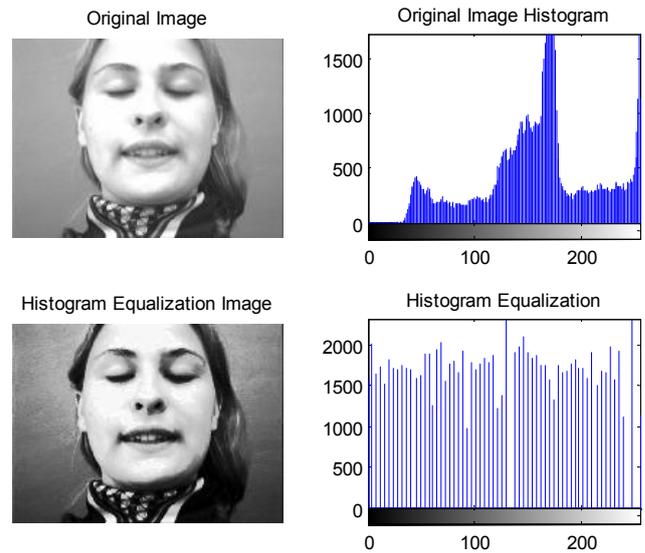


Fig.1 Performance comparison based on Histogram equalization method.

### B. Image Preprocessing Based on SSR Algorithm

#### (1) Retinex Algorithm

According to the theory proposed by Edwin Land, an image  $S(x, y)$  is composed of two parts, the reflection image  $R(x, y)$  and the incident image  $L(x, y)$ , whose principle is shown shown in Fig. 2. The image seen by the observer is the combination of the incident image and the reflection image, that is  $S(x, y)$ , which can be expressed as follows.

$$S(x, y) = R(x, y) \times L(x, y) \quad (8)$$

The essence of Retinex theory is to get the reflection property R through the image S, that is to remove the incident light L, thus obtaining the original appearance of the object.

#### (2) SSR Algorithm

According to the related theory of center-wrapping theorem, Jobsen proposed a single-scale Retinex algorithm, and its brief steps are described as follows.

Step 1: By taking the logarithm on the original image, the reflected light can be separated from the Irradiation light.

$$S(x, y) = r(x, y) + l(x, y) = \log(R(x, y)) + \log(L(x, y)) \quad (9)$$

Step 2: The original image  $F(x, y)$  was convolved with a Gaussian function, which can also be called low-pass filtering. The function  $D(x, y)$  after filtering is obtained.  $F(x, y)$  represents the Gaussian filter function.

$$D(x, y) = S(x, y) * F(x, y) \quad (10)$$

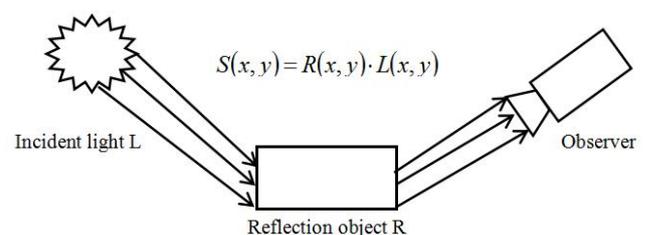


Fig. 2 Retinex theoretical schematic.

Step 3: The logarithmic function is adopted to carry out the subtraction operation, that is to use the previous image to subtract the low-pass filtered image to obtain the high-frequency enhanced image  $G(x, y)$ .

$$G(x, y) = S(x, y) - \log(D(x, y)) \quad (11)$$

Step 4: Take the anti-logarithm on  $G(x, y)$  to obtain the enhanced image.

$$R(x, y) = \exp(G(x, y)) \quad (12)$$

Step 5: Take the contrast enhancement with  $R(x, y)$  to realize the required effect.

Several typical central surround functions  $F(x, y)$  are expressed as follows.

a.  $F(x, y) = \frac{1}{r^2}$ , where  $r = \sqrt{x^2 + y^2}$ .

b.  $F(x, y) = 2 / (a + \left(\frac{r^2}{c_1^2}\right))$ .

c.  $F(x, y) = e^{-\frac{|r|}{c_2}}$ .

d.  $F(x, y) = e^{-\frac{r^2}{c_1}}$ .

The SSR algorithm is processed, and the processing results are shown in Fig. 3. The SSR algorithm is applied to the original image and the histogram equalization is used to process the picture. The results are shown in Fig. 4.

It can be seen from the above figures that after the original image is processed by SSR algorithm, the illumination effect can be obviously improved, which is helpful for the later identification process.

### (3) Optimization of SSR Algorithm

In the computer vision field, the Difference of Gaussian (DOG) is an enhanced algorithm to reduce the blurred image of a gray image by using another fuzzy image of an original gray image. The ambiguity of the blurred image is reduced by DOG algorithm.



Fig. 3 Image processing results by using SSR algorithm.



Fig. 4 Histogram equalization by using SSR algorithm.

One-dimensional Gaussian difference function is expressed as:

$$f(x, \mu, \sigma_1, \sigma_2) = \frac{1}{\sigma_1 \sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma_1^2}\right) - \frac{1}{\sigma_2 \sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma_2^2}\right) \quad (13)$$

Two-dimensional Gaussian difference function is expressed as:

$$f(u, v, \sigma) = \frac{1}{2\pi\sigma} \exp\left(-\frac{u^2 + v^2}{2\sigma^2}\right) - \frac{1}{2\pi K^2 \sigma^2} \exp\left(-\frac{u^2 + v^2}{2K^2 \sigma^2}\right) \quad (14)$$

Eq. (14) replaces the above mentioned Gauss low-pass filter  $F(x, y)$ . By adjusting the parameters reasonably, the improved SSR algorithm is obtained. The Gaussian difference filter replaces the Gaussian low-pass filter. The experimental results are shown in Fig. 5. The original image is tested with the improved SSR algorithm, and then the histogram is equalized. The experimental results are shown in Fig. 6. According to the experimental results, the improved SSR algorithm has better processed effect on the light problem of the original image than direct usage of the SSR algorithm.

### (4) Other Image Preprocessing Methods

Other image preprocessing methods include noise processing, area segmentation, edge detection, etc., and are not described in details herein. The experimental results of these pretreatment methods are shown in Fig. 7 to Fig. 12.



Fig. 5 Improved SSR algorithm.



Fig. 6 Histogram equalization based on the improved SSR algorithm.



Fig. 7 Salt and pepper noise processing.



Fig. 8 Region segmentation.



Fig. 9 Boundary extraction.

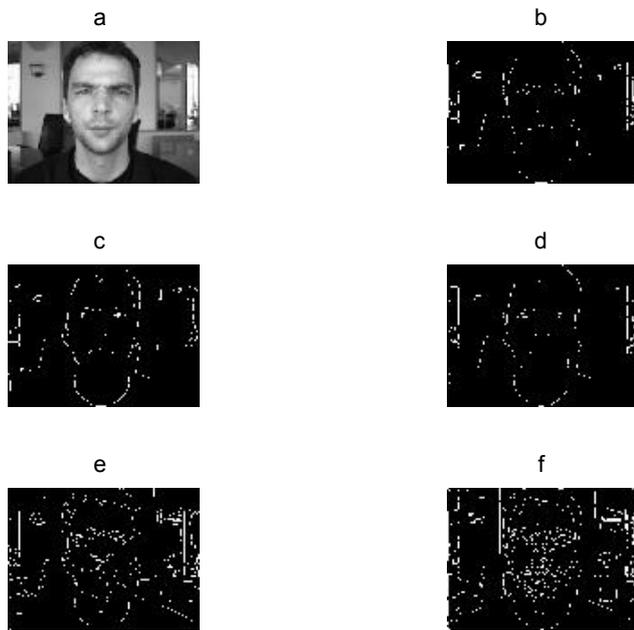


Fig. 10 Edge detection.



Fig. 11 Grayscale corrosion expansion.



Fig. 12 Corner detection.

### III. EIGENFACE EXTRACTION

#### A. Eigenface Extraction Based on Gabor Transform

The related properties of Gabor function are close to those of human eyes in terms of the features of objects, which can effectively find the edge and local features of texture image at

different frequencies and in different directions. Gabor filter is a special type of Fourier transform with Gaussian window. It is the only function which can get the best combined resolution in time domain and frequency domain under uncertainty principle. Gabor filter is a feasible method to extract the eigenface of face images.

The corresponding response function of Gabor filter modulated by two-dimensional Gauss function is described as follows.

$$g(x, y, k_x, k_y) = \exp\left[-\frac{1}{2\sigma^2}(x^2 + y^2) + j(k_x x + k_y y)\right] \quad (15)$$

where  $\delta$  is the standard deviation of Gaussian function,  $k_x$  is the relative frequency component in the x-axis direction,  $k_y$  is the relative frequency component in the y-axis direction,  $k = \sqrt{k_x^2 + k_y^2}$  is the relative frequency of the warp, and the

Gabor filter direction can be expressed as  $\theta = \arctan\left(\frac{k_y}{k_x}\right)$ .

Because the filters can have different center frequencies, the periodic characteristics of the Eigenface itself can be expressed just by the center frequency of the Gabor filter. So we can select a set of band-pass narrowband filters with different frequencies and extract them by filtering Eigenvalues of face images. If N Gabor filters are used to extract the features of a picture, a feature vector can be obtained, which can be expressed as:

$$g_{(i,j)}(a, k, \theta, \sigma) = \left\| \sum_{x=0}^{m-1} \sum_{y=0}^{q-1} f(x, y) g(i-x, j-y, a, k, \theta, \sigma) \right\|^2 \quad (16)$$

where  $f(x, y)$  is the original image and  $g_{(i,j)}(a, k, \theta, \sigma), k = 0, 1, 2, \dots, N$  is the eigenvalue at the center point  $(x, y)$  of the Gaussian window.

Five different frequencies of Gabor filters are  $1, 1/\sqrt{2}, 1/2, 1/2\sqrt{2}$  and  $1/4$  and eight different directions are  $\pi/4, 3\pi/8, \pi/2, 5\pi/8, 3\pi/4, 7\pi/8, 0, \pi/8$ . The eigenface Extraction results shown in Fig. 13 to Fig. 18.

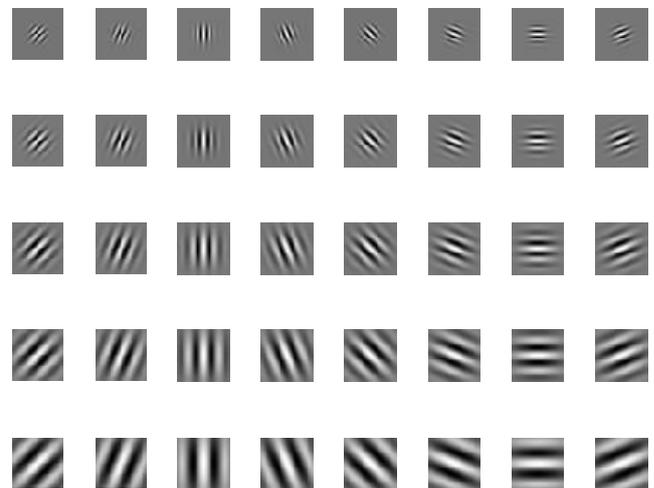


Fig. 13 Real part of Gabor filter.

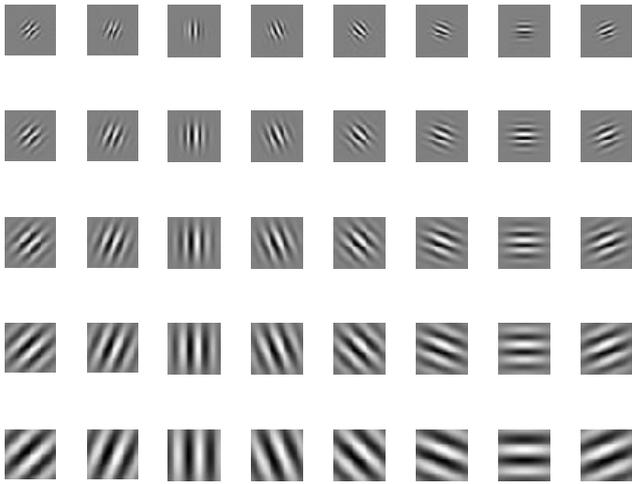


Fig. 14 Imaginary part of Gabor filter.

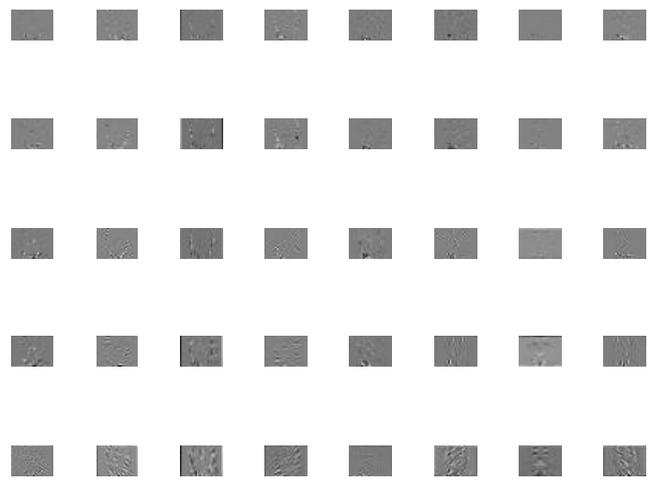


Fig. 17 Imaginary part of feature face based on Gabor filter.

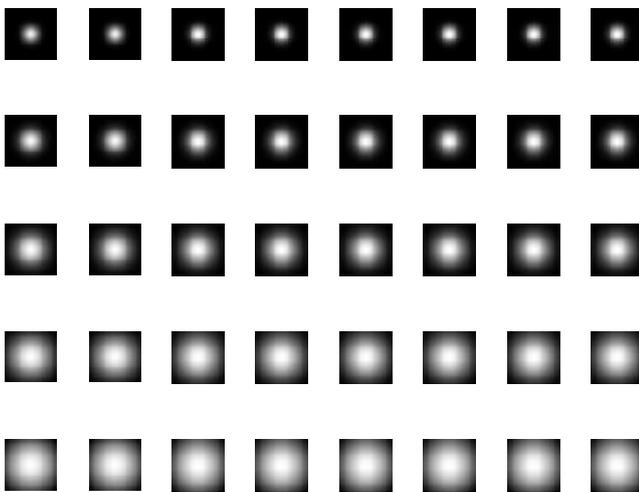


Fig. 15 Gabor filter frequency.

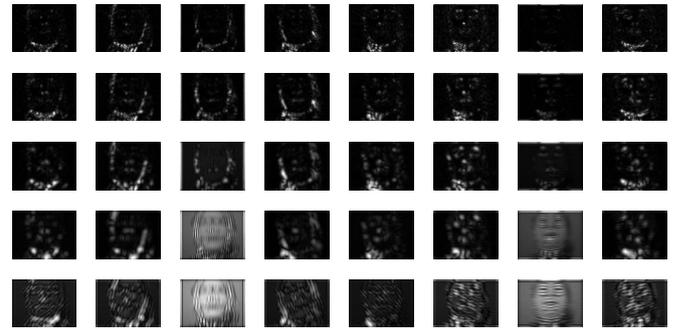


Fig. 18 Combination of real part and imaginary part.

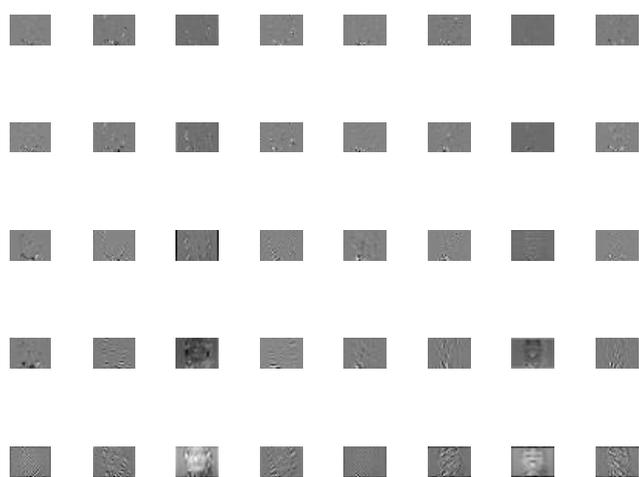


Fig. 16 Real part of feature face based on Gabor filter.

**B. Wavelet Transform**

The advantage of WT relative to Fourier transform is described as follows. Because wavelet transform is a local transformation between time domain and frequency domain, it can effectively extract the information in the signals through the translation and expansion of these operations on the function. The signal multi-scale refinement analysis can solve the problem many Fourier transform can not solves.

Therefore, the wavelet transform known as "mathematical microscope". The discrete wavelet transform is defined as follows.

$f$  is a measurable function and  $L^2(R)$  is a two-dimensional signal space of  $f$  on  $R$ . The required restriction of  $f$  is:

$$\int_{-\infty}^{\infty} |f(x)|^2 dx < \infty \tag{17}$$

An integral transformation  $W_\psi$  is defined on  $L^2(R)$ .

$$(W_\psi f)(b, a) = |a|^2 \int_{-\infty}^{\infty} f(x) \overline{\psi\left(x - \frac{b}{a}\right)} dx, f \in L^2(R) \tag{18}$$

$$\int_{-\infty}^{\infty} \psi(x) dx = 0 \tag{19}$$

In the above formula, the integral transform  $W_\psi$  is a connected wavelet transform, which is related to the basis wavelet  $\Psi$ .

Assuming the basis wavelet function  $\psi \in L^2(R) \cap L^1(R)$  and  $F\{\hat{\psi}(0)\} = 0$  (where  $\hat{\psi}$  is the Fourier transform of  $\Psi$ ),  $\Psi$  obtains a family of functions after translation and scaling transformations, which are described as follows.

$$\psi_{a,b} = |a|^{-\frac{1}{2}} \psi\left(x - \frac{a}{b}\right) \quad (a, b \in R, b \neq 0) \tag{20}$$

where,  $\psi_{a,b}$  is a continuous wavelet,  $a$  is a translation factor

and  $b$  is a scaling factor.

Discrete wavelet transform is used to discretize  $a$  and  $b$ , where  $a = a_0^{-m}$ ,  $b = nb_0 a_0^{-m}$ ,  $a_0 > 1$ ,  $b_0 \neq 0$ . So the following transformation formula are obtained.

$$\psi_{m,n}(x) = a_0^{\frac{m}{2}} \psi(a_0^m x - nb_0) \quad (m, n \in Z) \quad (21)$$

where  $Z$  is an integer set.

When applying wavelet transform on image processing, the two-dimensional functions (wavelet functions and scaling functions) must be used. The transformation form one-dimensional to two-dimensional can be expressed as:

$$\begin{aligned} \phi(x, y) &= \phi(x)\phi(y) \\ \psi^{(1)}(x, y) &= \phi(x)\phi(y) \\ \psi^{(2)}(x, y) &= \phi(x)\phi(y) \\ \psi^{(3)}(x, y) &= \phi(x)\phi(y) \end{aligned} \quad (22)$$

After two-dimensional transform, the wavelet transform is adopted to decompose or reconstruct the images. The detailed process of wavelet decomposition can be expressed succinctly in Fig. 19.

In Fig. 19, L is the low-frequency signal of the image and H is the high-frequency signal of the image. After the first decomposition, the original signal is decomposed into LL, LH, HL and HH, that is to say the original signal is decomposed into low frequency signal, mixed signal including low frequency and high frequency, and high frequency signal. In general, for a variety of images, their basic information is contained in the low-frequency sub-band, and some of the details of the image information or edge information is included in the high-frequency sub-band. So according to the above diagram, the low-frequency sub-band LL, which can be called an analysis sub-band, covers almost all the basic information of the image, and the other three sub-bands (LH, HH and HL) can be called detail signals and cover the image details information. The wavelet transform effect is shown in Fig. 20 and Fig. 21.

### C. Gabor Wavelet Transform

The essence of Gabor wavelet transform is actually the wavelet transform based on Gabor function, which can be used to analyze various kinds of images. The main principle of Gabor wavelet transform can be described as follows. Gabor function itself constitutes a non-orthogonal basis. The given basis function and its expansion will be able to get a relatively localized frequency description. For this reason, the Gabor wavelet transform method can extract the eigenfaces of face images, that is to say that the usage of a set of filters with different scales can obtain the local features with different scales of images.

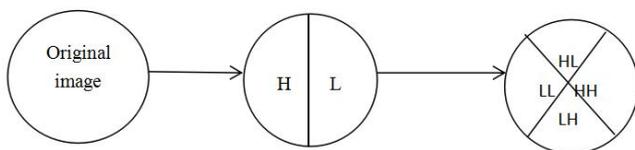


Fig. 19 Wavelet transform.



Fig. 20 Primary wavelet transform.

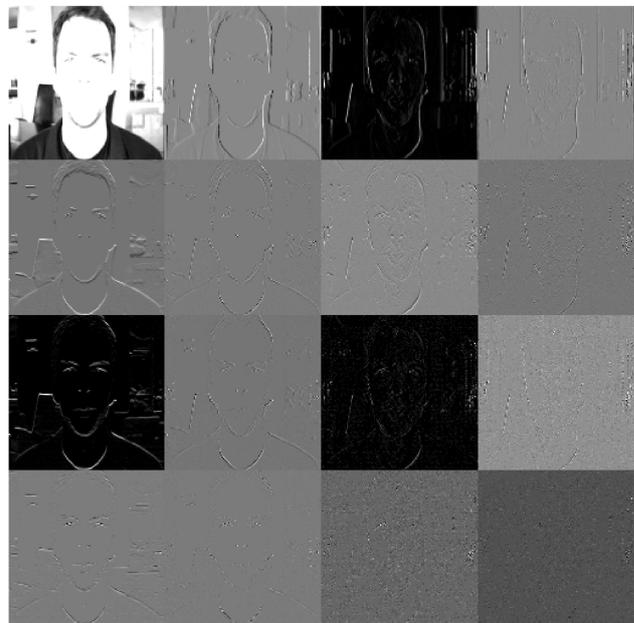


Fig. 21 Second wavelet transform.

The two-dimensional Gabor function expression is described as follows.

$$g(x, y) = \left[ \frac{1}{2\pi\sigma_x\sigma_y} \right] \exp \left[ -\frac{1}{2} \left( \frac{x^2}{\sigma_x^2} + \frac{y^2}{\sigma_y^2} \right) + 2\pi j Wx \right] \quad (23)$$

$g(x, y)$  is a basis function, and then expanded. This process is to carry out a certain degree of rotation and scaling expansion transform to obtain the Gabor wavelets. The specific expression function can be expressed as:

$$g_{m,n}(x, y) = a^{-m} G(x', y'), \quad a > 1, \quad m, n \in Z \quad (24)$$

where,  $(x', y') = a^{-m} (x \cos \theta + y \sin \theta, -x \sin \theta + y \cos \theta)$ ,  $\theta = \frac{n\pi}{M}$ ,  $M$  represents the number of directions, and  $a^{-m}$  represents a scale factor.

By changing the values of  $M$  and  $n$ , a set of Gabor wavelet filters with different directions and scales can be obtained. However, the obtained wavelet clusters are non-orthogonal to each other. If they are used to filter the images, there will be a lot of redundant information. When Gabor wavelet filter is adopted to extract facial features, the most crucial step is how to design these parameters of the function. If the parameters are designed reasonably, the redundant information of the images can be made smallest or best. Next these parameters can be derived one by one.

Assume  $U_L$  is the minimum filter center frequency,  $U_H$  is the filter center frequency maximum,  $M$  is the number of Gabor wavelet filter directions,  $S$  is the number of Gabor wavelet filter scales. Seen from the  $U_H = \alpha^{s-1}U_L$ , the scale parameter can be calculated by:

$$\alpha = \left(\frac{U_H}{U_L}\right)^{\frac{1}{s-1}} \quad (25)$$

$$U_H - U_L = t + 2\alpha t + 2\alpha^2 t + \dots + 2\alpha^{s-2} t + \alpha^{s-1} t = \frac{\alpha+1}{\alpha-1}(\alpha^{2s-1} - 1) \quad (26)$$

Because the Gaussian function half amplitude of the standard deviation  $\delta$  is  $\sigma\sqrt{2\ln 2}$ . Therefore, the corresponding maximum filter half amplitude is  $\alpha^{s-1}t = \sigma_u\sqrt{2\ln 2}$ . Substitute this relationship into Eq. (25) and (26) to obtain:

$$\sigma_u = \frac{(\alpha-1)U_H}{(\alpha+1)\sqrt{2\ln 2}} \quad (27)$$

According to the prior knowledge, the tangent angle of two adjacent ovals is  $\varphi = \pi / M$ , so there is:

$$\frac{(u-U_H)^2}{2\ln 2\sigma_u^2} + \frac{v^2}{2\ln 2\sigma_v^2} = 1 \quad (28)$$

Because of  $v = \tan \frac{\varphi}{2} u$ , so obtain:

$$\left(\sigma_v^2 + \tan^2 \frac{\varphi}{2} \sigma_u^2\right)u^2 - 2U_H\sigma_v^2u + U_H^2\sigma_v^2 - 2\ln 2\sigma_u^2\sigma_v^2 = 0 \quad (29)$$

In the equation with respect to  $u$ , the condition that it can have real solutions is:

$$\sigma_v = \tan \frac{\varphi}{2} \sqrt{\frac{U_H^2}{2\ln 2} - \sigma_u^2} \quad (30)$$

Based on Eq. (29) and (30), obtain:

$$\sigma_v = \tan\left(\frac{\pi}{2M}\right) \left[ U_H - 2\ln 2 \left( \frac{\sigma_u^2}{U_H} \right) - 2\ln 2 \left( \frac{\sigma_v^2}{U_H} \right) \right] \left[ 2\ln 2 - \frac{(2\ln 2)^2 \sigma_H^2}{U_H^2} \right]^{\frac{1}{2}} \quad (31)$$

After a series of above formulas derivation, the relationship among parameters of Gabor wavelet filter is very

obvious. As long as we determine these five parameters ( $\omega$ ,  $S$ ,  $M$ ,  $U_H$  and  $U_L$ ), other parameters of the filter can be easily calculated.

Five different frequencies of Gabor filters are  $1$ ,  $1/\sqrt{2}$ ,  $1/2$ ,  $1/2\sqrt{2}$  and  $1/4$  and eight different directions are  $\pi/4$ ,  $3\pi/8$ ,  $\pi/2$ ,  $5\pi/8$ ,  $3\pi/4$ ,  $7\pi/8$ ,  $0$ ,  $\pi/8$ . The eigenface Extraction results shown in Fig. 22 to Fig. 27.

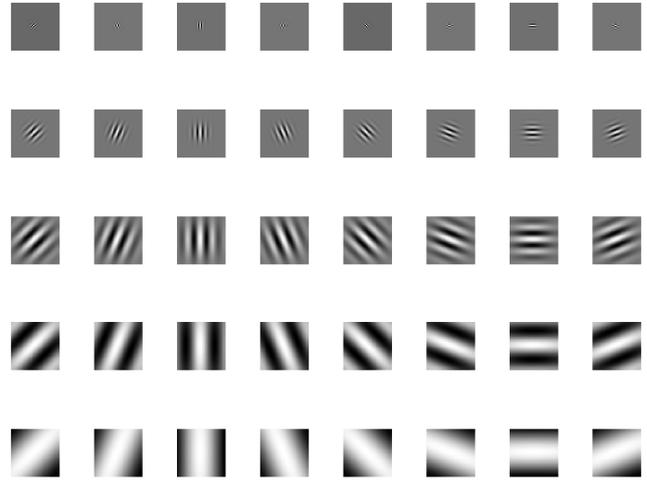


Fig. 22 Real part of Gabor wavelet filter.

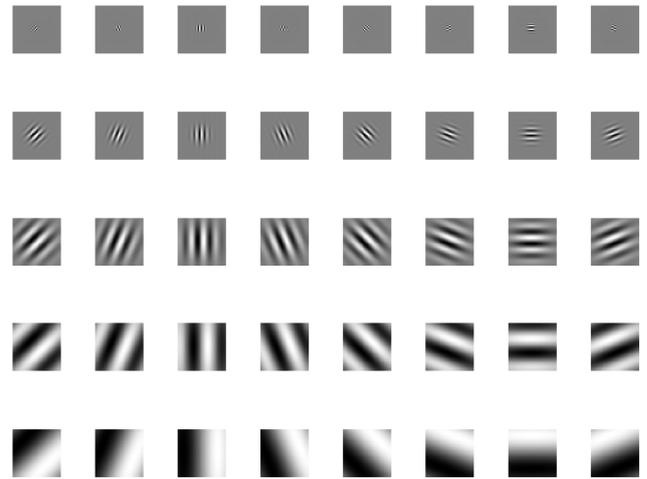


Fig. 23 Imaginary part of Gabor wavelet filter.

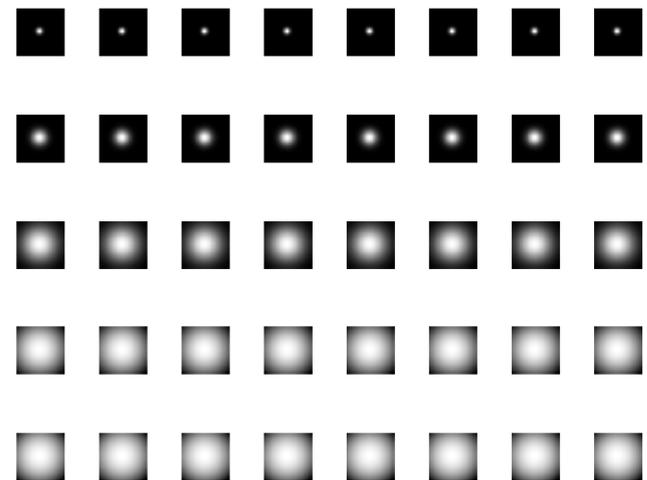


Fig. 24 Gabor wavelet filter frequency.

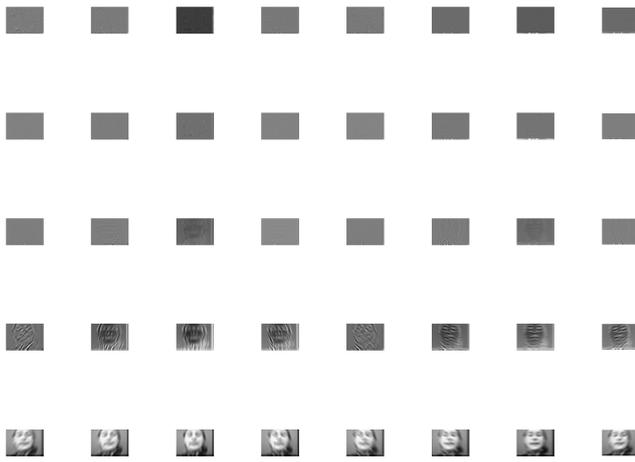


Fig. 25 Real part of feature face based on Gabor wavelet filter.

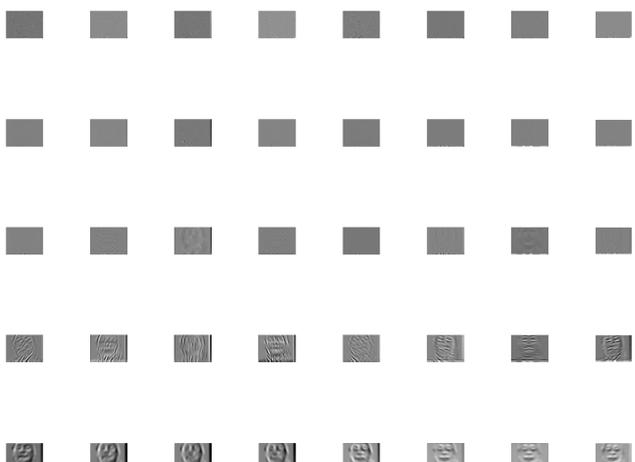


Fig. 26 Imaginary part of feature face based on Gabor wavelet filter.

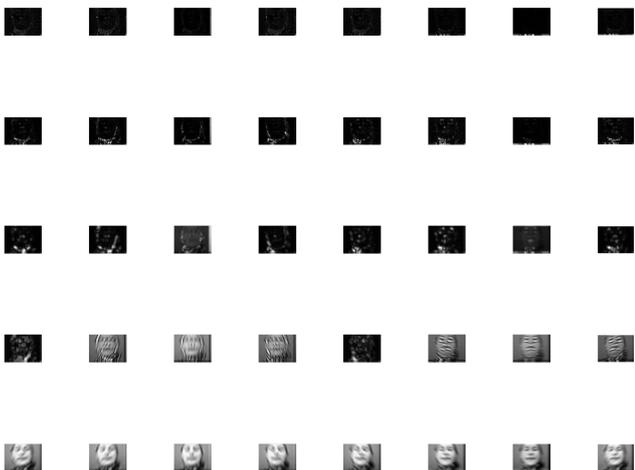


Fig. 27 Combination of real part and imaginary part.

It can be seen from the above figures that the 40 eigenfaces obtained by Gabor wavelet transform have a clearer outline. In addition to retaining most of the original image information, the original image is also effectively compressed. Compared with the Gabor-transformed feature Face, the performance is better, so it is more conducive to the late identification process.

These 40 Gabor filters have a total of 80 Gabor templates because each Gabor filter consists of real and imaginary parts.

In this article, the size of the Gabor template is 39\*39. The the essence of Gabor wavelet feature extraction process is essentially to use these Gabor template to carry out convolution operation with the image, which will produce 40 real response maps and 40 real response maps. The real part and the imaginary part are merged to form 40 total response graphs (39\*39). Finally, each response graph (39\*39) is straightened, that is to say the 39\*39 matrix is converted into a vector of 1\*1521. In this way, the number of features of each face (39\*39) is 40 \*1521.

#### IV. FACE RECOGNITION BASED ON IMPROVED GABOR WAVELET TRANSFORM

##### A. Two-dimensional Discrete Cosine Transform (2D DCT)

The two-dimensional discrete cosine transform is defined as follows:

$$F(0,0) = \frac{1}{N} \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} f(x,y) \quad (32)$$

$$F(0,v) = \frac{\sqrt{2}}{N} \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} f(x,y) \cdot \cos\left(\frac{(2x+1)v\pi}{2N}\right) \quad (33)$$

$$F(u,0) = \frac{\sqrt{2}}{N} \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} f(x,y) \cdot \cos\left(\frac{(2y+1)u\pi}{2N}\right) \quad (34)$$

$$F(u,v) = \frac{\sqrt{2}}{N} \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} f(x,y) \cdot \cos\left(\frac{(2y+1)u\pi}{2N}\right) \cdot \cos\left(\frac{(2x+1)v\pi}{2N}\right) \quad (35)$$

where,  $f(x,y)$  represents the image space matrix,  $(x,y)$  is the current image pixel position,  $F(u,v)$  is the transform coefficient matrix,  $u, v = 1, 2, \dots, N-1$ . The two-dimensional discrete transformation is actually to carry out one-dimensional discrete transformation twice. Fig. 28 is the effect of DCT transformation:



(a) Original image



(b) Transformed image

Fig. 28 Results of DCT.

Seen from Fig. 28, the energy of the converted DCT coefficients is mainly concentrated in the upper left corner and the rest of the coefficients are close to zero. This shows that DCT has the advantage of being suitable for image compression characteristic. The thresholding operation is carried out on the transformed DCT coefficient and the coefficients smaller than a certain value are set zero, which is the quantization process in the image compression. Then the inverse DCT operation is performed to obtain the compressed image. By the same token, if the Gabor filtered image is processed with 2D DCT, its effect is show in Fig. 29.

It can be seen from Fig. 29 that after 2D DCT processing, most of the energy of the image is concentrated in the DC part, that is to say in the upper left corner of the figure, and the correlation between the images can be well removed, which makes the image coding simple and the feature extraction easy. Thus a small amount of information in the upper left corner need be extracted as the effective information of the image, which can greatly compress the image. This is the principle of dimensionality reduction in DCT.

### B. Face Recognition Based on Improved Gabor Wavelet Transform

According to the principle of Gabor wavelet, Gabor filter processes every pixel in the entire image, which makes Gabor filter recognition rate higher. However, not every pixel is a feature point of image, and the feature point of the image is only a small part of the image pixels. Therefore, the Gabor filter has done a great deal of useless work, which will results in a great increase of recognition time. So when using the Gabor filter, how to reduce the recognition time must be considered.

When adopting 2D DCT to carry out image processing, the processed image is mainly concentrated in the DC part. If by using 2D DCT for image processing, a few part of the interception of the upper left corner of the image can be removed a lot of useless information, and leave the main information, which will also greatly reduce the workload of the Gabor filter and the recognition time. At the same time, 2D DCT can remove the correlation between images to a certain extent, which is helpful to improve the recognition rate.

In this paper, the idea of improving Gabor wavelet filter is to process the image by 2D DCT, intercept a small portion of the upper left corner, compress the image effectively, then extract the feature by Gabor wavelet. Thereby the recognition time will be greatly reduced and the recognition rate will be improved to some extent.

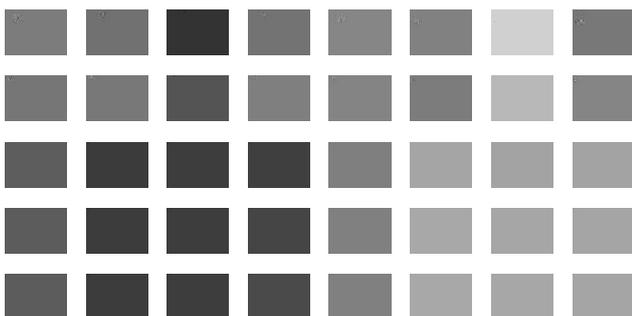


Fig. 29 DCT processing after filtering.

## V. FACE RECOGNITION CLASSIFICATION

### A. Classifier

In this paper, the minimum distance classifier is adopted to realize the face recognition classification. The minimum distance classifier is one of the most commonly used distance classifiers. Firstly, the class centers of the known samples are calculated. Then the sample is classified to the class in relation to the nearest class center.

In an  $n$ -dimensional space,  $A$  is defined as the name of the category,  $X_A$  as the feature set of sample  $A$ ,  $x_{An}$  as the feature set of the  $n$ th dimension of category  $A$ ,  $\mu_A$  as the mean of  $A$ , and  $\mu_{An}$  as the mean of the  $n$ th feature set. The minimum distance classifier firstly calculates the mean of each dimension of each known class  $X_A = (x_{A1}, x_{A2}, x_{A3}, \dots, x_{An})$  to form an average vector  $\mu_A = (\mu_{A1}, \mu_{A2}, \mu_{A3}, \dots, \mu_{An})$ . Similarly, the mean  $\mu_B = (\mu_{B1}, \mu_{B2}, \mu_{B3}, \dots, \mu_{Bn})$  belonging to another class  $X_B = (x_{B1}, x_{B2}, x_{B3}, \dots, x_{Bn})$  is calculated. Then the distances  $d(x, \mu_A)$  and  $d(x, \mu_B)$  form  $X$  to  $X_A$  and  $X_B$  are calculated.

### B. Distance of Classifier

After the face samples have been subjected to the feature extraction process, the feature space based on these feature vectors is finally formed. Therefore, how to calculate the similarity between the samples can generally be calculated by calculating the distance or angle. Given the vectors  $x$  and  $y$ , the common similarity measures includes:

#### 1. Euclidean distance

$$d(x, y) = \|x - y\| = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \quad (36)$$

#### 2. Block Distance

$$d(x, y) = \sum_{i=1}^n |x_i - y_i| \quad (37)$$

#### 3. Mahalanobis distance

$$d(x, y) = (x - y)^T C^{-1} (x - y) \quad (38)$$

where,  $C$  is the covariance matrix of the pattern, which can also be simplified as:

$$d(x, y) = \sum_{i=1}^n \frac{x_i y_i}{\sqrt{\lambda_i}} \quad (39)$$

where  $\lambda_i$  is the variance of the  $i$ th component.

#### 4. Cosine angle

$$d(x, y) = \frac{x \cdot y}{\|x\| \|y\|} = \frac{\sum_{i=1}^n x_i y_i}{\sqrt{\sum_{i=1}^n x_i^2} \sqrt{\sum_{i=1}^n y_i^2}} \quad (40)$$

In this paper, the Euclidean distance is chosen as the classifier distance.

VI. SIMULATION EXPERIMENT AND RESULT ANALYSIS

Experimental environment: Windows 10 64-bit operating system, the processor is Intel (R) Core i5 7th Gen, system memory 8G, MATLAB R2016a simulation software. The simulation experiments are carried out according to the following procedure.

Step 1: The BioID face database is selected to carried out simulation experiments. PCA algorithm, Gabor wavelet algorithm and improved Gabor wavelet algorithm are used to recognize the face database respectively, and the recognition rate and recognition time are compared.

Step 2: The ORL face database is chosen to proceed simulation experiments. Then PCA algorithm, Gabor wavelet algorithm and Gabor wavelet algorithm are used respectively to recognize the face database, , and the recognition rate and recognition time are compared.

Step 3: Finally, experiments were carried out in the BioID face database to analyze the relationship between the selected area and the recognition rate after 2D DCT treatment.

A. Experiment 1

The face images belonging to five different people in BioID face database are selected as test images. Each selected test image is in corresponding to two training images, i.e. there are ten pieces of five different people for training images. The PCA algorithm, the Gabor algorithm, the wavelet Gabor wavelet algorithm are adopted to carry out the recognition experiments on the face database respectively. Then 10, 15 and 20 different face images are selected as test images and the corresponding 20, 30 and 40 face images as training images. The selected BioID library face expression is not very rich and there is not too much inconsistency, and did not wear glasses and other effects, The part faces in the BioID face database are shown in Fig. 30. The simulation experimental results are shown in Fig. 31 and Fig. 32.



Fig. 30 Some face images in BioID face database.

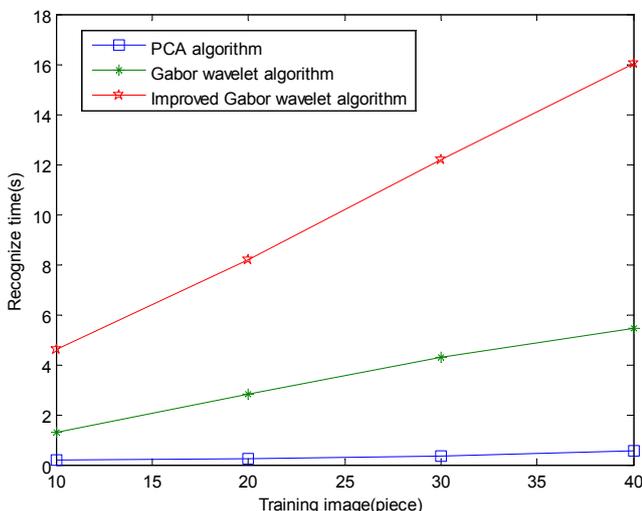


Fig. 31 Recognition time.

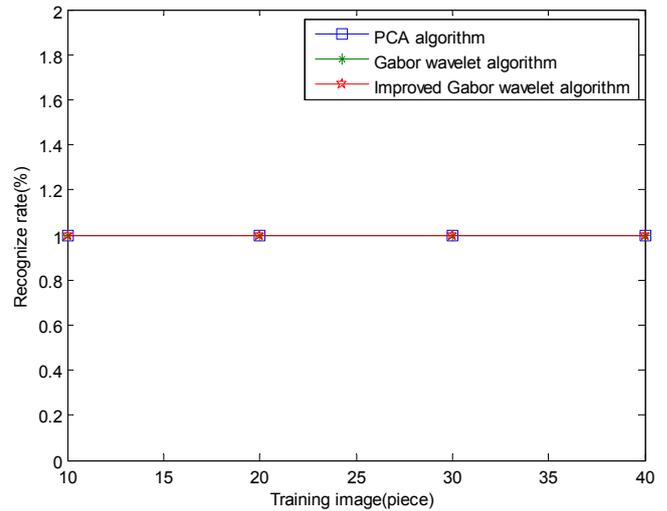


Fig. 32 Recognition rate.

It can be seen from the experimental results that the recognition rates of the three different recognition algorithms are both 100%, which is because the selected BioID face database is relatively easy to identify. For the recognition time, the PCA recognition method is the fastest, the Gabor wavelet transform face recognition is the slowest, and the improved Gabor wavelet recognition recognition speed faster than the Gabor wavelet transform face recognition. Especially for 40 training images, the improved algorithm is faster than the average recognition time of the non-improved algorithm by more than 10s, which fully shows that the proposed algorithm is very effective.

B. Experiment 2

The ORL face database is adopted to carry out simulation experiments on the above three different algorithms. The face expression and the body attitude in ORL face database are very rich, and some people wear glasses. The part faces in the ORL face database are shown in Fig. 33.

The simulation experimental results are shown in Fig. 34 and Fig. 35. Seen form the above simulation experiments results, the PCA face recognition time is still the fastest, the Gabor wavelet transform method is the slowest, and the Gabor wavelet recognition time is improved significantly faster than did not change. For the recognition rate, the recognition rate of the improved Gabor wavelet method is significantly higher than the other two algorithms. In the ORL face database, the effectiveness of the improved algorithm is also proved.



Fig. 33 ORL face database.

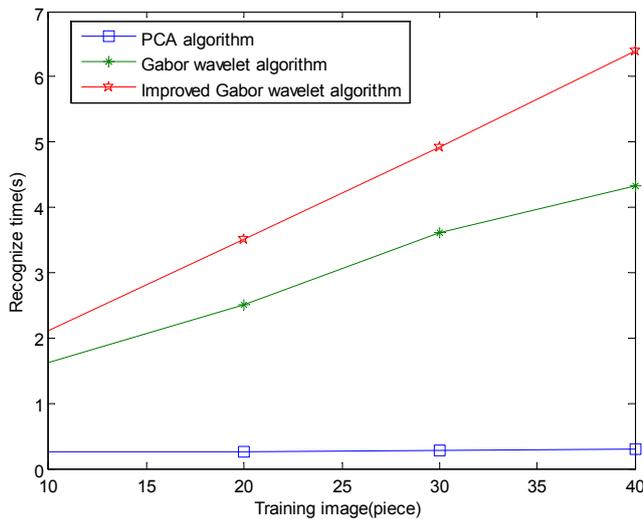


Fig. 34 Recognition time.

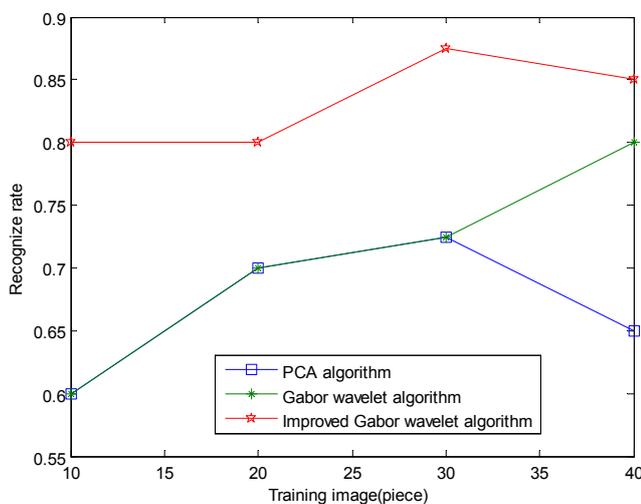


Fig. 35 Recognition rate.

C. Experiment 3

After 2D DCT processing, the size of the upper left corner area has a great relationship with the recognition rate and recognition time. Therefore, the area of the interception must be studied in order to achieve the best recognition time and recognition rate. BioID face database: considering the upper left corner of the vertex, the square areas are selected with the length of the original side length 1/100, 1/90, 1/80, 1/70, 1/60, 1/57, 1/55, 1/40, 1/20, 1/10 and 1/5. The experimental results are shown in Fig. 36 and Fig. 37.

Seen from the above figures, the cut-off critical point that can be effectively identified is 1/57. It is hardly recognizable on the left side of the cut-off point, and the recognition rate is still 100% on the right side of the cut-off point. However, when the interception point from 0 to 1/50, the recognition time changes very little, so the interception point can be set at 1/50. At this time, the average recognition time is only 4.472s and the recognition rate is still 100%. ORL face database: considering the upper left corner of the vertex, the square areas are selected with the length of the original side length 1/20, 1/15, 1/10, 1/5, 1/4, 1/3, 1/2 and 1. The experimental results are shown in Fig. 38 and Fig. 39.

From 1 to 1/2, the recognition time varies greatly. Whereas

form 0 to 1/2, the recognition time is relatively small, and at 1/2 the recognition rate can still be as high as 80%. So 1/2 interception can be used as the best interception point for ORL face database, where the average recognition time is 4.806s and the recognition rate is 80%.

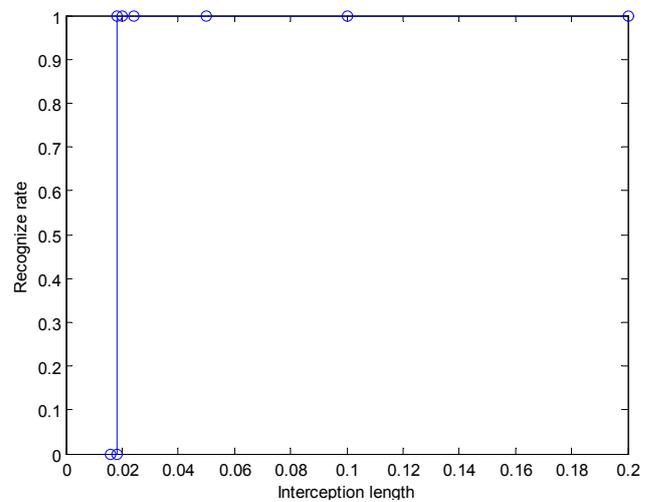


Fig. 36 Recognition rate and interception length.

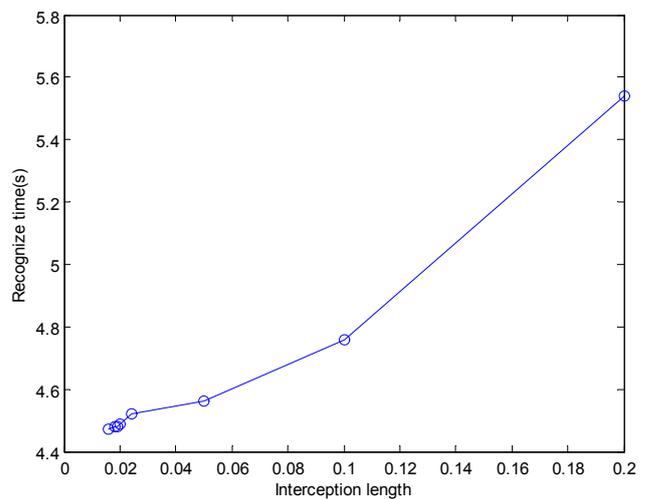


Fig. 37 Recognition time and interception length.

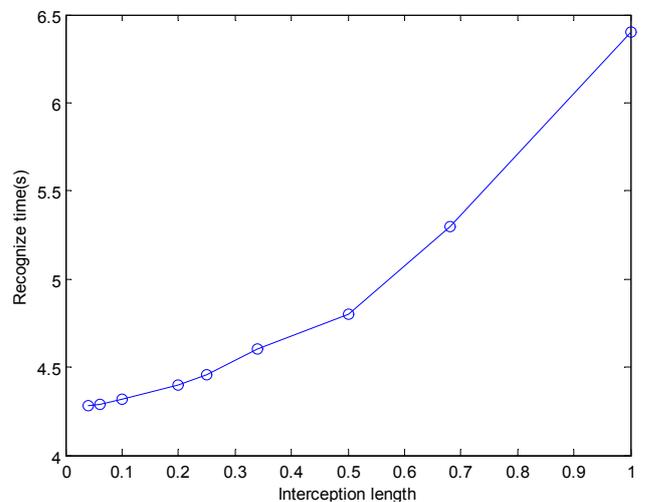


Fig. 38 Recognition rate and interception length.

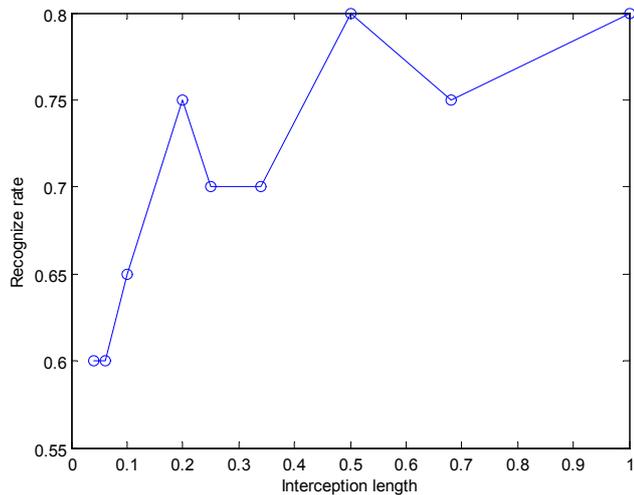


Fig. 39 Recognition time and interception length.

## VII. CONCLUSIONS

This paper discusses the recognition principle and recognition process of face recognition. Before recognition, it is necessary to preprocess the face original image. The pretreatment method is to process the original image by using SSR algorithm. However, the direct use of SSR algorithm for image processing is not very effective, so we improve the SSR algorithm, which is replaced by a Gaussian low-pass filter, and then processed images with a histogram Equalization again, so that after the improvement, the treatment effect is much better, that is to say that the effect of light treatment is obvious. The improved Gabor wavelet transform is used to realize recognition face. The use of 2D DCT is adopted to compress the original image so that recognition rate and recognition time are better than directly using Gabor wavelet transform. The classifier selects the nearest neighbor classifier and the Euclidean distance to classify the image. This reduces the time and increases the effective recognition rate.

## REFERENCES

- [1] D. J. Robertson, E. Noyes, A. J. Dowsett, J. Rob, and B. A. Mike, "Face recognition by metropolitan police super-recognisers," *Plos One*, vol. 11, no.2, pp. e0150036, 2016.
- [2] W. Hwang, X. Huang, K. Noh, and J. Kim, "Face recognition system using extended curvature gabor classifier bunch for low-resolution face image," *Pattern Recognition*, vol. 48, no.4, pp. 1247-1260, 2015.
- [3] C. Ding, and D. Tao, "Robust face recognition via multimodal deep face representation," *IEEE Transactions on Multimedia*, vol. 17, no.11, pp. 2049-2058, 2015.
- [4] Y. Xu, Z. Li, B. Zhang, J. Yang, and J. You, "Sample diversity, representation effectiveness and robust dictionary learning for face recognition," *Information Sciences*, vol. 375, no.C, pp. 171-182, 2017.
- [5] A. F. Abate, M. Nappi, D. Riccio, and G. Sabatino, "2D and 3D face recognition: a survey," *Pattern Recognition Letters*, vol. 28, no.14, pp. 1885-1906, 2007.
- [6] G. J. Hou, G. D. Wang, Z. K. Pan, B. X. Huang, H. Yang, and T. Yu, "Image enhancement and restoration: state of the art of variational retinex models," *IAENG International Journal of Computer Science*, vol. 44, no.4, pp. 445-455, 2017.
- [7] A. A. Mohammed, R. Minhas, Q. M. J. Wu, and M. A. Sid-Ahmed, "Human face recognition based on multidimensional PCA and extreme learning machine," *Pattern Recognition*, vol. 44, no.10, pp. 2588-2597, 2011.
- [8] Z. Huang, R. Wang, S. Shan, and X. Chen, "Face recognition on large-scale video in the wild with hybrid Euclidean-and-Riemannian metric learning," *Pattern Recognition*, vol. 48, no.10, pp. 3113-3124, 2015.
- [9] L. A. Cament, F. J. Galdames, K. W. Bowyer, and C. A. Perez, "Face recognition under pose variation with local Gabor features enhanced by active shape and statistical models," *Pattern Recognition*, vol. 48, no.11, pp. 3371-3384, 2015.
- [10] A. De, A. Saha, and M. C. Pal, "A human facial expression recognition model based on eigen face approach," *Procedia Computer Science*, vol. 45, pp. 282-289, 2015.
- [11] K. C. Kwak, and W. Pedrycz, "Face recognition using a fuzzy fisherface classifier," *Pattern Recognition*, vol. 38, no.10, pp. 1717-1732, 2005.
- [12] A. Tefas, C. Kotropoulos, and I. Pitas, "Face verification using elastic graph matching based on morphological signal decomposition," *Signal Processing*, vol. 82, no.6, pp. 833-851, 2002.
- [13] Y. Choi, H. I. Kim, and M. R. Yong, "Local feature learning using deep canonical correlation analysis for heterogeneous face recognition," *Journal of Korea Multimedia Society*, vol. 19, no.5, pp. 848-855, 2016.
- [14] X. Yin, and X. Liu, "Multi-task convolutional neural network for pose-invariant face recognition," *IEEE Transactions on Image Processing*, vol. 27, no.2, pp. 964-975, 2018.
- [15] Y. Shen, M. Yang, B. Wei, C. T. Chou, and W. Hu, "Learn to recognise: exploring priors of sparse face recognition on smartphones," *IEEE Transactions on Mobile Computing*, vol. 16, no.6, pp. 1705-1717, 2016.
- [16] Z. S. Zhao, L. Zhang, M. Zhao, Z. G. Hou, and C. S. Zhang, "Gabor face recognition by multi-channel classifier fusion of supervised kernel manifold learning," *Neurocomputing*, vol. 97, pp. 398-404, 2012.
- [17] N. W. Tay, K. L. Chu, and M. Peruš, "Face recognition with quantum associative networks using overcomplete Gabor wavelet," *Cognitive Computation*, vol. 2, no.4, pp. 297-302, 2010.
- [18] Z. Sufyanu, F. S. Mohamad, and A. A. Yusuf, "A new discrete cosine transform on face recognition through histograms for an optimized compression," *Research Journal of Information Technology*, vol. 7, no.2, pp. 101-111, 2015.
- [19] Y. Zhang, D. Zhao, J. Sun, G. Zou, and W. Li, "Adaptive convolutional neural network and its application in face recognition," *Neural Processing Letters*, vol. 43, no.2, pp. 389-399, 2016.
- [20] Y. H. Kim, H. Kim, S. W. Kim, H. Y. Kim, and S. J. Ko, "Illumination normalisation using convolutional neural network with application to face recognition," *Electronics Letters*, vol. 53, no.6, pp. 399-401, 2017.
- [21] S. Bashbaghi, E. Granger, G. A. Bilodeau, and G. A. Bilodeau, "Dynamic ensembles of exemplar-SVMs for still-to-video face recognition," *Pattern Recognition*, vol. 69, no.C, pp. 61-81, 2017.
- [22] P. Kamencay, M. Zachariasova, R. Hudec, and R. Jarina, "A novel approach to face recognition using image segmentation based on SPCA-KNN method," *Radioengineering*, vol. 22, no.1, pp. 92-99, 2013.
- [23] J. Lin, J. Ming, and D. Crookes, "Robust face recognition with partial occlusion, illumination variation and limited training data by optimal feature selection," *Iet Computer Vision*, vol. 5, no.1, pp. 23-32, 2011.
- [24] J. Liu, X. Jing, S. Sun, and Z. Lian, "Local Gabor dominant direction pattern for face recognition," *Chinese Journal of Electronics*, vol. 24, no.2, pp. 245-250, 2015.
- [25] X. Liu, L. Shen, and H. Fan, "Face recognition algorithm based on Gabor wavelet and locality preserving projections," *Modern Physics Letters B*, vol. 31, no.19-21, pp. 1740041, 2017.
- [26] H. Hadizadeh, "Multi-resolution local Gabor wavelets binary patterns for gray-scale texture description," *Pattern Recognition Letters*, vol. 65, no.C, pp. 163-169, 2015.
- [27] B. W. Zheng, J. S. Wang, Y. L. Ruan, and S. Z. Gao, "Fingerprint recognition method based on Gabor wavelet transform and discrete cosine transform," *Engineering Letters*, vol. 26, no.2, pp. 228-235, 2018.
- [28] P. Parida, and N. Bhoi, "2-D Gabor filter based transition region extraction and morphological operation for image segmentation," *Computers & Electrical Engineering*, vol. 62, pp. 119-134, 2017.
- [29] R. Atta, and M. Ghanbari, "Low-memory requirement and efficient face recognition system based on DCT pyramid," *IEEE Transactions on Consumer Electronics*, vol. 56, no.3, pp. 1542-1548, 2010.

**Jie-sheng Wang** received his B. Sc. and M. Sc. degrees in control science from University of Science and Technology Liaoning, China in 1999 and 2002, respectively, and his Ph. D. degree in control science from Dalian University of Technology, China in 2006. He is currently a professor and Master's Supervisor in School of Electronic and Information Engineering, University of Science and Technology Liaoning. His main research interest is modeling of complex industry process, intelligent control and Computer integrated manufacturing.

**Yan-lang Ruan** is currently a master student in School of Electronic and Information Engineering, University of Science and Technology Liaoning, China. His main research interest is image processing and pattern recognition.

**Bo-wen Zheng** is currently a master student in School of Electronic and Information Engineering, University of Science and Technology Liaoning, China. His main research interest is image processing and pattern recognition.

**Shu-Zhi Gao** received the B. Sc. degree from Shenyang University of Chemical Technology, China in 1989, and received the M. Sc. degrees from Northeastern University, China in 2012. She is currently a professor and Ph. D. supervisor in Shenyang University of Chemical Technology. Her research interests include feedback control system, modeling of complex industry process and intelligent control.