Face recognition based on the invariant single training sample

Zhaonan Yang[1,2*], Shu Zhang [2]

[1]*College of Computer Science and Technology, Harbin Engineering University No.145, Str. Nantong, Harbin, China*

[2]*College of Computer Science and Technology, Heilongjing University of Science and Technology No.l2468 Str. Puyuan, Harbin, China*

Received 1 October 2014, www.cmnt.lv

Abstract

Despite the constantly change of human face pose, illumination, expression, and occultation, one major problem of the face recognition technique arises from the difficulties of gaining training samples. When everyone just can gain an image for face recognition, the training samples are so insufficient that the extracted feature vectors can not support the whole face sample subspace and the performance drop is expected. This problem is called the face recognition with the single training sample and has received significant attention during the past years. Researchers have proposed image-strengthen method, sample-expansion method, and generic learning framework, etc, which mostly aim to expand the number of the training samples by using computer techniques to create several combining virtual images based on the original one. Therefore, the problem simply becomes a general face recognition problem. However, these methods result in enlarging the calculation volume and requiring bigger storage space. It also needs to be retrained once a new person is put into system. These problems make it extremely difficult to popularize these methods. In this paper, we try to exclude training and to extract features directly from the hybrid Taylor-ATMT, which has constructed a set of invariants. The recognition errors caused by the change of human face expression, illumination and partial occultation could be reduced after projecting it to wavelet space to lower the dimension, and then classify categories with the use of Bayesian Decision Theory, which results in a better effect. Experiments are implemented on YALE and ORL face databases to demonstrate the efficiency of the proposed approach. The experimental results show that the average recognition accuracy rates of our proposed method which are higher than those of previous methods.

Keywords: face recognition, single training sample per person, analytical Fourier-Mellin transform (AFMT), Taylor transform, wavelet transform

1 Introduction

The research on the fact recognition started in the end of the 1960s. The face recognition with the features of non-aggressive, direct, friendly and convenient is always a research emphasis in the biological recognition field. With a further research and the extension of the application, the face recognition has an important application value and prospect in the archives management system, the safety verification system, the credit card verification, the identification of the police criminal system, the supervisory system of the bank and the airport, entry and exit barrier management, etc. Therefore, it has become a hot research issue in the computer vision, applied mathematics, automation control, virtual reality, image processing, mode recognition, information engineering and other fields in the 1990s. Many scholars have proposed all kinds of effective recognition algorithms, especially Turk and Pentland (1991) who introduced Eigenface. The Eigenface regards the image area including human face as a random vector and adopts Principal Component Analysis to gain the image feature vectors. In addition, the method can have those shapes similar to the face corresponding to the basement with a larger characteristic value. Therefore, the face image can be described, expressed and then recognized by using those linear combinations of base. Belhumeur et al. (1997) proposed Fisherface on this basis. Linear Discriminant Analysis is adopted to change the reduced Principal Component Analysis dimensionality for gaining a bigger between-class scatter and a smaller within-class scatter, which is still one of the mainstream in the current face recognition. Therefore, many different kinds of variants have been produced, such as zero space method, the discrimination model in the subspace, the strengthening discrimination model and the improving strategy based on the kernel learning. Later, Moghaddam proposed the face recognition method based on the Eigenface with the use of Bayesian probability estimation on the basis of the double subspaces. The method converts the similarity calculation problem with two face images into a problem of the double-class classification through a differential method. Then the density of the class conditional probability in the two classifications should be calculated and the face can be recognized through Bayesian Decision Theory. Another important method is Flexible Models, including Active Shape Model, ASM and Active Appearance Model. The ASM/AAM describes face as two detached parts: 2D shape and texture. The two parts adopt the statistical method to do the PCA and then do the statistical modeling for the face by combing with the two parts through PCA. In addition, Support Vector Machine, proposed by Phillips (1991), is a representative statistical learning theory being applied in the face recognition. The applications in those above kinds of algorithms are fit for

* ***Corresponding author's*** e-mail: zhaonanyanga@163.com

the condition of the multi training samples per person. However, when every one just can gain a single image in several special occasions, such as ID verification, passport verification, codes and standards, many troubles can be brought for the above algorithms by using the limited images to train face recognition system. As the face recognition with the single training sample can bring large challenges and have an important significance itself, it has become a important research in recent years and attracted widely attention. Many researchers have proposed many methods, such as Sample-strengthen Method, Sample-expansion Method, Generic Learning Framework and 3D Face Model Recognition. Those methods adopts all kinds of technologies to combine many virtual images on the basis of original ones for increasing the numbers of samples and conforming to the traditional face recognition algorithm. However, it can bring many problems, such as large storage volume, slow calculating speed. As the generated virtual images are highly correlated with the original ones, the recognition effect is worse than the multi image per person. Therefore, the paper proposes a direct comparison method without doing the samples and training the samples, in which the method directly extracts the invariant features to be applied in the single training sample. The method has the actual advantages in the face recognition with the signal image per person and provides a new research direction.

2 Summary of the spectrum face recognition

It is easily for human beings to recognize the change of the position, the direction and the dimension size of the object shapes in the two dimensional image, while most of the computers have poor recognition effect during the process of the model recognition. The recognition procedure is called as Geometric Similarity Transform which obtain various invariants, such as Translation Invariant, Rotation Invariant and Scaling Invariant. Therefore, in order to make the computers have the function of the face recognition, all kinds of different invariant features and describing factors are researched. For example, Teh and Chin (1988) adopt Moment to describe the distribution of the image intensity. Flannery and Horner (1989) adopt Fourier Descriptors to summarize the shape boundary. De Castro and Morandi (1987) prove that the Magnitude Spectrum image of the Fourier-Mellin transform has the Translation Invariant. Sheng (1989) illustrates that the Magnitude Spectrum image of the Fourier-Mellin transform is used as the Rotation Invariant and Scaling Invariant. As the references for describing the factors with all kinds of variants are still limited, how to extract the describing factors with the Translation Invariant, Rotation Invariant and Scaling Invariant is still a large challenge. In addition, how to build a Geometric Similarity Transform invariant is the discussing purpose. The invariant still remains all information, apart from its position, direction and dimension, which is called the complete invariant. The complete invariant is different from the incomplete invariant which just remains its position, direction and dimension. The complete invariant and the incomplete invariant sometimes is called the absolute invariant and the relative variant respectively. The incomplete invariant also can be applied in some aspects For example, Brandt (1992) applies it in the text recognition, in which the used feature vectors just remain a little parts of the original information.

The discussed complete invariant is the expression method of the complete invariant image, namely, the Image Normalization. The procedure is to transform the image into a standard form and all similar transformations are invariable. The normalization expression belongs to the Spatial Domain Invariants. Frequency Domain Invariants are discussed and their invariants are derived from the Fourier-Mellin transform. The Magnitude Spectrum belongs to the Translation Invariant, some related significant Phase information is removed for the Magnitude Spectrum also belongs to the incomplete invariant. In order to solve the problem, Lin and Brandt (1993) removes the linear parts of the Phase information for forming the complete invariant, which is called as Taylor Invariant. In a similar way, although the Magnitude Spectrum in the Fourier-Mellin transform is the Rotation Invariant and Scaling Invariant, the provided related Phase information can not be applied in solving the problem for it also is the incomplete invariant. Ghorbel (1994) proposes the complete Complex Spectra for analyzing the Rotation Invariant and Scaling Invariant under the Fourier-Mellin transform framework. In order to gain the Translation complete Invariant, Rotation complete Invariant and Scaling complete Invariant, Yu and Bennamoun (2007) redesign two types of Geometric Similarity Transform complete invariant to describe the Complex Spectra combining with the Taylor invariability under the framework of the TMT and the AFMT. That is, the Taylor-TMT hybrid Geometric Similarity Transform complete invariant and the Taylor-AFMT hybrid Geometric Similarity Transform complete invariant. Both of them remove the linear Phase information, the problem of the scattering invariant and the Phase scattering will be caused. The former hybrid Geometric Similarity Transform complete invariant changes the linear phase general number into the integer. It is so complicated in the actual application that Taylor invariability should be combined to form the hybrid Similarity Transform complete invariant under the ATMT framework. The method can obtain the same effect with the easiness and the increasing calculating speed in the experiment.

There is a little spectrum domain invariant applied in the references. For example, Lai et al. (2001) firstly proposed the overall Fourier invariant feature, that is, the so called Spectroface is applied in the face recognition. He Jia and Du Minghui proposed the face recognition with the single sample based on the image-strengthen method and the Fourier Spectroface. The training samples and the rebuilding figures of the interrupted singular value can be combined with a new sample, and then the image spectrum can be obtained with the Fourier-Mellin transform and used as the recognition feature. The experiment proves that the method has its effectiveness in the ORL face database. Lai Jianhuang (2002) proposed the classifier design in the complete Spectroface recognition. The Spectroface adopts the Wavelet Transform and the Fourier-Mellin transform to extract the features of the image's invariable Phase and the relative invariable expression. The paper mainly discusses the selection between the pre-processing and the similar measurement in the Spectroface method system. The two key problems influence the precision of the recognition. After the Spectroface being processed. It is compared and analyzed through the following four typical similar measurement methods: the nearest neighboring method, the average method, the Hausdroff distance method and the modified Hausdroff distance method. The experimental results show that the four kinds of methods are the effective methods to do the similar measurement for the Spectroface, and the nearest neighboring method is the most effective method. She Yongzeng and Chen Yong (2003) proposed the face recognition algorithm combining the wave figure with the decision. At first, the wavelet transform is used to decompose the facial image into the lowest sub-band at suitable levels, and the lowest sub-band is decomposed to obtain four sub-bands from the different directions. The traditional Principal Component Analysis or the Fourier-Mellin transform should be done in the four sub-bans, and then the four recognitions can be obtained. Later, the final recognition result can be obtained in terms of the decision integration scheme.

The former Spectrum invariant methods are limited in the application of the Fourier Spectroface. Aiming to the Hybrid Similarity Transform Invariant under the Fourier-Mellin transform, there are few discussions in the references. In order to analyze the recognition ability of the Hybrid Similarity Transform Invariant under the Fourier-Mellin transform framework, the Taylor-ATMT and the Taylor-TMT are proposed. The Taylor-ATMT emphasizes the Hybrid Similarity Transform Invariant so that the Taylor-ATMT can have a better effect in the face recognition.

3 The spectroface in the Hybrid Fourier-Mellin transform

The following is the illustration of all kinds of the Hybrid Similarity Transform Invariant proposed in the AFMT framework. In order to make these illustrations clear, some basic symbols are used. (x, y) represents the Cartesian space coordinate, (ρ, φ) represents the Log-Polar coordinate, (r, θ) represents the polar coordinates, (u, v) represents the Cartesian spectrum coordinate and (k, ω) represents Fourier transform within the spectrum domain. When an image is transformed into the Log-Polar coordinate or the polar coordinates, the name of the image function is still remained, and the invariant is represented by (ρ, φ) or (r, θ) respectively.

3.1 DISCRETE FOURIER TRANSFORM

If an image is the real number continuous function f(x, y) which is defined within the Cartesian grid of the integer value, 0 < x < M，0 < y < N, the discrete Fourier function is defined as follows:

$$F(u,v) = \frac{1}{MN}\sum_{x=0}^{M-1}\sum_{y=0}^{N-1} f(x,y)e^{-i2\pi(ux/M+vy/N)} . \quad (1)$$

In general, the result F(u, v) in the Fourier Transform is the complex funnction:

$$F(u,v) = R(u,v) + jI(u,v) = \left|F(u,v)\right| e^{j\phi(u,v)} .$$

Its Magnitude Spectrum is defined as follows:

$$\left|F(u,v)\right| = \sqrt{R^2(u,v) + I^2(u,v)} .$$

Its Phase Spectrum is defined as follows:

$$\phi(u,v) = \tan^{-1}\left[\frac{I(u,v)}{R(u,v)}\right].$$

Its Energy Spectrum is defined as follows:

$$E(u,v) = R^2(u,v) + I^2(u,v) ,$$

where the Magnitude Spectrum is also called Fourier Spectrum.

3.2 TAYLOR INVARIANT

The concept of the invariant is derived from the linear parts on the basis of removing the discrete Fourier phase spectrum. Therefore, the definition of Taylor Invariant is defined as follows:

$$F_T(u,v) = e^{-j(au+bv)} F(u,v) , \quad (2)$$

where a and b is obtained from the u and v differential, namely, a = φu(0, 0)，b = φv(0, 0).

3.3 TAYLOR-MELLIN INVARIANT

The Fourier-Mellin Transform is used to build up the Rotation Invariant and the Scaling Invariant. At first, the image function f(x, y) is projected in the Log-Polar coordinate, and its image is represented by f(ρ, φ) and imposes on the Fourier Transform so that the procedure is called as the Fourier-Mellin Transform. In the same way, the Taylor-Mellin Invariant imposes the Log-Polar coordinate f(ρ, φ) on the Taylor Invariant and the complete Rotation Invariant and the Scaling Invariant can be obtained. The definition is as follows:

$$M_T(k,\omega) = e^{-j(au+b\omega)} F(k,\omega) . \quad (3)$$

where $M_T(k, \omega)$ is called the Taylor-Mellin Invariant

3.4 ANALYTICAL FOURIER-MELLIN TRANSFORM

The polar coordinates changed from the horizontal coordinates can be represented as follows:

$\Pi = \left\{(r,\theta)\middle| r > 0 \ \ and \ \ 0 < \theta < 2\pi\right\}$.

The polar coordinates changed from the Fourier-Mellin Transform can be represented as follows:

$$f(u,v) = M_f(u,v) = \int_0^{+\infty}\int_0^{2\pi} f(r,\theta)e^{-iu\theta}r^{-iv}\frac{dr}{r}d\theta, \qquad (4)$$

where u belongs to the positive integer and v belongs to the real number. It represents the Fourier-Mellin Transform with the polar coordinate discrete function in the two dimensional image. The Equation (4) is usually discrete, but the convergence is very necessary. Therefore, the f(r, θ) must be convergence to Krα (α > 0 and K is a constant), and the center in near to the image original point. The Equation (4) can be modified based on the reason as follows:

$$M_f(u, s = \sigma + iv) = \int_0^{+\infty}\int_0^{2\pi} f(r,\theta)e^{-iu\theta}r^{\sigma+iv}\frac{dr}{r}d\theta, \qquad (5)$$

where u ∈ Z, v ∈ R and σ > 0, it is called Fourier-Mellin Transform. The analytical Fourier-Mellin Transform can be obtained from imposing the twisted image f^σ (r, θ) = r^σ f(r, θ) on the Fourier Transform. The formula can be changed slightly and the q is ln(r), so the Equation (5) can be rewritten as follows:

$$M_{f\sigma}(u,v) = \frac{1}{2\pi}\int_{-\infty}^{+\infty}\int_0^{2\pi} e^{q\sigma} f(e^q,\theta)e^{-i(u\theta+qv)}d\theta dq. \qquad (6)$$

The above formula can be obtained from a rapid algorithm, that is, the two dimensional rapid Fourier Transform can be calculated by the twist image $e^{q\sigma}$ f(e^q, θ) in the log-polar coordinates. The Sampling in the log-polar coordinates is built in the coordinate whose original point is image center point, and its radius is the increase of the index by n radiation lines intersects M concentric circles. The analytical Fourier-Mellin Transform is aiming to the transformation principle of the translation similarity. If f is the object in which the object f changes β angle direction and the size of α measurement factor, name ly, g(r, θ) = f(α r, θ + β). The two objeatc have the same shape and the similar objects. The relevancy of the AFMT is as follows:

$$M_{g\sigma}(k,v) = a^{-\sigma+iv}e^{ik\beta}M_{f\sigma}(k,v). \qquad (7)$$

where k belongs to the positive integer, v belongs to the real number and σ is bigger than zero. The Equation (7) is called as Shift Theorem which is fit for calculating the features of all angles, and the features include the invariability of its position, direction and size. The analytical Fourier-Mellin Transform often has two similar objects and different Phase factors, that is, the Equation (7) has no α –σ. Crimmins (1982) describes that a group of the complete invariant has nothing to do with its position, direction and size. It usually calculates the modulus of the Fourier-Mellin factors, but its invariant is incomplete so that it just can represent a Shape or Signature because of losing the Phase information. A different object with the same describing value can be mixed up during the classification because of lack of completeness. Therefore, Ghorbel (1994) advises the similarity invariant based on the AFMT to be modified into completeness and it can be represented by the positive σ value. The Equation (7) can be rewritten as follows:

$$I_{f\sigma}(k,v) = M_{f\sigma}(0,0)^{-\sigma+iv}e^{ik(M_{f\sigma})(1,0)^{-1}}M_{f\sigma}(k,v), \qquad (8)$$

where k belongs to the positive integer, v belongs to the real number and the building of the I(k, v) feature is to compensate the α -σ +iv $e^{ik\beta}$ in the Equation (7), which can be obtained by the Fourier-Mellin factors in the two objects and two normalization parameter M(0,0) and M(1,0). There are two points to support its completeness in the Equation (8):

1. Equation (8) can be reversed and the object's Fourier-Mellin Transform can be obtained;
2. The original image can be rebuilt through the AFMT reversion.

3.5 HYBRID COMPLETE INVARIANTS

The Hybrid Complete Invariants is divided into two types: the first type is based on the Taylor-Mellin and the second type is based on the analytical Fourier-Mellin Transform. When the Taylor-Mellin considers the Translation Invariant, Rotation Invariant and Scaling Invariant together, the Hybrid Complete Invariants can be built on the basis of the Taylor-Mellin Transform by combining the Translation Invariant with the Rotation Invariant and Scaling Invariant, as shown in the following:

$$S(\cdot) = M_T(F_T(\cdot)), \qquad (9)$$

where M_T is the complex spectrum in the Taylor-Mellin Invariant and F_T is the real image imposed on the Taylor-Mellin Invariant. Note that M_T is directly imposed on the complex spectrum and can not be imposed on the magnitude spectrum and the Phase spectrum, in which the two invariants are produced.

When the Taylor-Mellin Invariant and the AFMT Invariant are combined based on the analytical Fourier-Mellin Transform, the Hybrid Invariants are formed on the basis of the AFMT framework. The basic feature in the Equation (7) meets the requirement with the complex number. Therefore, the AFMT can be imposed on the Taylor Spectrum, but the Equation (7) just can be applied in the spectrum of the polar coordinate with the reciprocal measurement feature of the Fourier Transform so that the Equation (7) must be modified as follows:

$$M_{g\sigma}(k,v) = a^{\sigma-2-ik}e^{iv\beta}M_{f\sigma}(k,v).$$

The AFMT Invariant in the Equation (8) also should be modified as follows:

$$I_{f\sigma}(k,v) = \left|M_{f\sigma}(0,0)\right|^{(-\sigma+2+ik)/(\sigma-2)} \times \exp(-i\,\mathrm{var}\,g(M_{f\sigma}(0,1))) \cdot M_{f\sigma}(k,v).$$

Therefore, the Hybrid Complete Similarity Invariants based on the AFMT is as follows:

$$S(\cdot) = I_{f\sigma}(F_T(\cdot)). \tag{10}$$

3.6 ANALYTICAL FOURIER-MELLIN TRANSFORM

From the Equation (5), the Analytical Fourier-Mellin Transform given an image function f(r, θ) can be as follows:

$$M_{f\sigma}(u,v) = \frac{1}{2\pi}\int_0^1\int_0^{2\pi} f(r,\theta) r^{\sigma-iv} e^{-u\theta} \frac{dr}{r} d\theta.$$

The real time processing is very important for the extracted application based on the image contents. Therefore, the rapid algorithm can be rewritten as the Fourier-Mellin Transform model by changing the variant in which the q is equal to ln(r):

$$M_{f\sigma}(u,v) = \frac{1}{2\pi}\int_{-\infty}^{+\infty}\int_0^{2\pi} e^{q\sigma} f(e^q,\theta) e^{-i(u\theta+qv)} d\theta dq.$$

The equation above can be obtained through a rapid algorithm, that is, the two dimensional rapid Fourier Transform can be calculated with the twist image $e^{q\sigma}$ f(e^q, θ) in the log-polar coordinates.

$$M_{f\sigma}(u,v) = r^{\sigma} F(u,v).$$

The complete Rotation Invariant and the Scaling Invariant can be obtained by using the Equation (3) Fourier-Mellin Invariant to multiply a factor r^{σ}, which is as shown in the following:

$$M_{T\sigma}(k,\omega) = r^{\sigma} M_T(k,\omega), \tag{11}$$

$M_{T\sigma}(k, \omega)$ is called as the Analytical Fourier-Mellin Transform.

4 Face recognition system and database

A Face recognition system contains three phases: pre-processing phase, feature extraction phase and the classification phase. Each phase is as follows:

4.1 PRE-PROCESSING PHASE

Intensity Normalization can influence the change of the magnitude spectrum image with the change of the image intensity. In order to overcome these problems caused by the change of the illumination, the Intensity Normalization should be done in the pre-processing phase. The Gaussian Filter (namely, partially smoothing it in the 3 × 3 area) can be used to reduce the noise and then the complete normalization should be done so that the Intensity Normalization can have the average value o and the standard deviation value 1. In this way, the difference of the intensity can be reduced. As to the two dimensional image f(x, y), the complete normalization is as shown in the following:

$$f_s(x,y) = \frac{f(x,y)-m}{\sigma} \times \sigma_s + m_s, \tag{12}$$

where *m* and σ is the average value and the variance of the Intensity Normalization respectively. m_s and σ s is the expected average value 0 and the expected variance 1.

4.2 FEATURE EXTRACTION PHASE

The face recognition requires stable characteristics and it is unique in the feature space. The stable feature is invariable and these features may be the functions in th space and the spectrum. Unique features provide the correspondence one by one between the feature values and the waiting face recognition. The selected feature extracts the Spectroface in the Geometric Invariant (namely, the Hybrid Complete Invariants), which has be described in the section 3. In order to reduce the calculation, the wavelet transform is used to degrade and then the feature vector can be gained, the wavelet transform is fixed by the size of the window, but its shape, time window and frequency window can change the partial changing analysis method of the time & frequency. The wavelet multi resolution analysis proposes by Mallat (1989) was an effective tool of the image model recognition in 1989. If the signal function f(x, y)∈ L2(R2), φ(x, y) in the two dimensional discrete face image is the two dimensional wavelet function, the wavelet decomposition recursion formula in the two dimensional image is as follows:

$$\begin{aligned}
&c_{j,m_1,m_2} = (H_r \otimes H_c)(c_{j+1})_{m_1,m_2} = \sum_{k_1,k_2} c_{j+1,k_1,k_2} h_{k_1-2m_1} h_{k_2-2m_2} \\
&d^1_{j,m_1,m_2} = (H_r \otimes G_c)(c_{j+1})_{m_1,m_2} = \sum_{k_1,k_2} c_{j+1,k_1,k_2} h_{k_1-2m_1} g_{k_2-2m_2} \\
&d^2_{j,m_1,m_2} = (G_r \otimes H_c)(c_{j+1})_{m_1,m_2} = \sum_{k_1,k_2} c_{j+1,k_1,k_2} g_{k_1-2m_1} h_{k_2-2m_2} \\
&d^3_{j,m_1,m_2} = (G_r \otimes G_c)(c_{j+1})_{m_1,m_2} = \sum_{k_1,k_2} c_{j+1,k_1,k_2} g_{k_1-2m_1} g_{k_2-2m_2}
\end{aligned}, \tag{13}$$

where the infinite matrix $H_r = (H_{k_1,m_1}),\ H_c = (H_{k_2,m_2})$, $G_r = (G_{k_1,m_1}), G_c = (G_{k_2,m_2})$, $H_{k,m} = h_{k-2m}, G_{k,m} = g_{k-2m}$. The subscript "r", "c" represents the matrix's row operation and the line operation respectively, h is the low-pass filter, g is the high-pass filter, d1 is the factor in the wavelet decomposition vertical direction, d^2 is the factor in the wavelet decomposition horizontal direction, d^3 is the factor in the wavelet decomposition diagonal direction and c is the factor in the wavelet decomposition sub-band direction.

After the original image being wavelet decomposed, the resolution in the sub-band's image is reduced, and the low frequency sun-band includes a large part of the useful information in the original image. Therefore, the calculation complexity can be reduced by calculating the low resolution image. The wavelet transformation decompose the image into the sum of the closing image and the detailed image. The original image after being one class wavelet transformed can be decomposed four sub-bands and the four gained output part is LL, LH, HL, HH. The following Figure 1 is the corresponding four sub figures after being a Mallat algorithm. Therefore, the low frequency sub-band closest to the original images can be adopted rather than the proper layer wavelet transformation in terms of the performance of the multi resolution analysis. If the face image is done by n two dimensional wavelet decomposition, the size of the low frequency sub-band image is 1/22n of the original ones. Therefore, the wavelet transformation can effectively reduce the dimensional numbers.

LL	LH
HL	HH

FIGURE 1 The corresponding four sub figures in the Mallat algorithm

LL sub figure represents the low frequency part between the horizontal and vertical direction and remains the main energy of the original image. It is the result of smoothing the original image and becomes the input part of the next decomposition. LH sub figure represents the low frequency part in the horizontal and the high frequency in the vertical direction and remains the horizontal line of the original image. HL sub figure represents the high frequency part in the horizontal and the low frequency in the vertical direction and remains the vertical line of the original image. HH sub figure represents the high frequency part in the horizontal and the vertical direction and remains the intersection between the horizontal line and the vertical line of the original image.

4.3 SORTING PHASE

The main work, face recognition is to identify facial features known face data and database data matching, actually can be classified as a process, is the key to select the appropriate classifier, and classification strategy. Of face different features and classifier selection will be different, can be a traditional minimum distance method, nearest neighbor method, etc., can also be a new neural network or support vector machine (SVM), etc.

Which feature vector to obtain the image, must with the database category known image feature vector, to express the known image feature vector and to recognize the consistency of image feature vector, define the image feature vector expressions such as:

$$V^{(k)} = \left\{v_1^{(k)}, v_2^{(k)}, v_3^{(k)}, \ldots, v_n^{(k)}\right\}.$$

Their components $v_1^{(k)}$ are the characteristics of the image k vector, this set of feature vectors in the database category known image feature vector. To identify the characteristics of the image vector expression is as follows:

$$V' = \left\{v_1', v_2', v_3', \ldots, v_n'\right\}.$$

This to identify the image feature vector and image database of known category feature vector, to identify which belongs to the category model.

When a feature vector extracted on the basis of the characteristics of feature vector must choose one of the most suitable Classifier, generally the most common use of Classifier with Nearest Neighbor Classifier (on his Neighbor Classifier), Correlation Coefficient Method, the Correlation Coefficient Method). In addition, statistical decision theory is one of the basic theory of pattern classification problems. It on model analysis and the design of classifier has practical guiding significance. Among them, the Bayesian (the Bayes decision-making method is a basic method of pattern recognition.

4.4 THE DATABASE

Chose and YALE face database ORL face database identification experiment. In the YALE face database, a total of 165 (15, 11 each face images), all for the positive image, each 11 face images respectively represent different expressions, taken under cover and light conditions. From left to right, top-down respectively has the center of the light source, like (with a little cover) with a pair of glasses, happy like, like light source on the left, not like not (cover) with a pair of glasses, like a normal light sources have no facial expression, on the right side of the light source, sad like, like sleeping, surprise, blink of an eye, such as the different expression, each image resolution is 128 by 128 (see Figure 2).

In ORL face database, a total of 400 face images (each of 40 people, 10 facial image), each with 8-bit grayscale image pigment, its resolution is 92 x 112 pixels, after pre-processing cut of 92 x 92 resolution. Most people space position, the intensity of illumination direction and about the same, but with a little expression, posture (positive face, on the left side of the face, the right side of the face, looked up and down), the change of the scale, rotation (as shown in Figure 3)

4.5 FACE RECOGNITION SYSTEM PROCESS

Training, everyone were first introduced in one image, the image after normalization, to extract the feature vector, adult face image database structure, in this way, in the face each person has a sample in the database, for $v_j^{(k)} = \left\{v_j^{(k)}\right\}$, as for the rest of the image recognition portrait, recognition processing. Process as shown in Figure 4. This experiment for the proposed feature extraction method of implementation, and analysis and comparison, as the basis of the proposed method.

5 Algorithm

The proposed algorithm is based on the hybrid similarity transformation Taylor - ATMT invariant features and frequency invariant wavelet transform. The mixed Taylor - ATMT transformation is applied to similar face images to extract geometrical invariant features, and then, apply wavelet transformation to the similar geometric invariant features, the eigenvector downgrade, in order to reduce recognition system calculation.

FIGURE 2 YALE face database in the first 11 photo graph

FIGURE 3 ORL face database in the first 10 photos

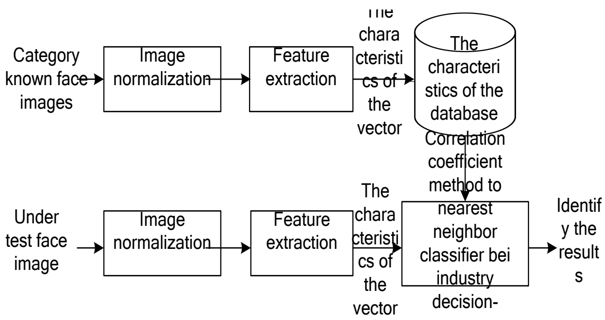

FIGURE 4 Face recognition system flow diagram

Using the nearest neighbor classifier, correlation coefficient method and minimum error rate of Baye industry decision, to research on classification results. The algorithm implementation process after the pick described such as:

Step 1: extraction of two-dimensional image grey value and normalization Equation (12).

Step 2: use Taylor conversion Equation (2) will be the image of invariant to translation, and then use ATMT Equation (11) on the translation invariant, in order to obtain scale and rotation invariant, and as a frequency signal, has the similar geometric invariants.

Step 3: using the wavelet transform Equation (13) will be the first order frequency signal decomposition downgrade, get low figure, as a feature vector.

Step 4: calculate database of known image feature vector and to recognize the characteristics of the image vector similarity between Equation (15), the largest similarity for the output category. Lu, the experiment and comparing the experiment selects two standard YALE and ORL face database, to assess proposed similarity transformation under the mixed architecture of invariant system recognition accuracy. These data face image contains different Expressions (Expressions), light (tiny Illumination) such as the change of the direction and stance (Poses) (P.I.E.), in order to assess the impact on the system.

5.1 THE LIGHTING PROBLEM OF THE SINGLE TRAINING SAMPLE

In a single training sample, a very important problem in face recognition, is the development of the algorithm is less affected by the light. To confirm that the proposed algorithm has the characteristics of Taylor – ATMT choose a group, in the first image ORL database in their personal and YALE database of each person's sixth blessing image as a known image, because they are compared with other images are not P.I.E. problem, other images as a test. In addition to choose a control group, and also is each person's fifth image of ORL database (for posture, large scale) and YALE database of each person's fourth image (for the left light source, background is black on the right) as a known image, the image as a test. And select Taylor – TMT method and Taylor – AFMT method and comparison, such as 5:

Experiments by the pre-processing stage, the first in their image through grey value standardization, in order to reduce the light changes. Then through feature extraction phase, on the basis of calculating program shown in Figure 5, extraction of image characteristic value. YALE were used for the experiment database image size is 128×128, ORL database image size is 92×92, and sigma is equal to 0.5, the above algorithm calculating program transformation expressed in the image below (Figure 6). Pay attention to the experiment combined with Taylor invariant must render the complex domain spectrum, the last stage of classification, respectively by the nearest neighbor classifier, correlation coefficient method and the minimum error rate of Baye industry decision classifier, such as matching, we became different, and the accuracy of the classifier is in the following table.

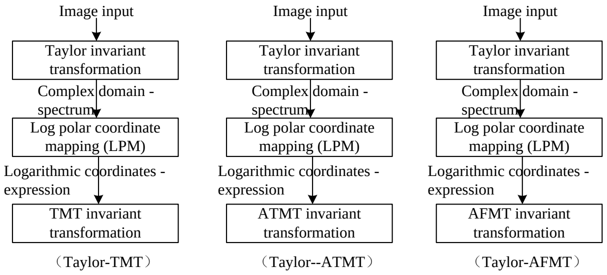

FIGURE 5 Taylor - ATMT application compared with other methods

FIGURE 6 Original image (left); into a logarithmic polar image (in); into ATMT amplitude spectrum shadow (right)

TABLE 1 Different hybrid similarity transformation method and classification accuracy of the database

1	Taylor-TMT		Taylor-ATMT		Taylor-AFMT	
2	Yale	ORL	Yale	ORL	Yale	ORL
3	Baye NNC CCM	Baye NNC CCM	Baye NNC CCM	Baye NNC CCM	Baye NNC CCM	Baye NNC CCM
4	0.69 0.69 0.70	0.65 0.65 0.67	0.69 0.69 0.70	0.67 0.67 0.69	0.71 0.71 0.70	0.45 0.45 0.69
5	0.15 0.15 0.13	0.55 0.55 0.58	0.21 0.21 0.13	0.68 0.58 0.62	0.13 0.13 0.13	0.40 0.40 0.62

Note: the above-mentioned Numbers said, the first column 1 is the way; 2 for the database; 3 for classifier; 4 for the experimental group; 5 for the control group.

The experimental results show that:

1) In the experimental group: Taylor – ATMT algorithm of classification accuracy, no matter in YALE database or ORL database are more Taylor – TMT and Taylor – AFMT slightly higher. And ATMT do not variables when combined with Taylor, its algorithm is AFMT easily.

2) In the control group: ORL database for the fifth image profile and the influence of scale, intensity of illumination Angle and consistency, so its accuracy compared with the normal image slightly, with little effect. In the fourth image illumination Angle larger figure of YALE database is deeper, its accuracy is normal image. So in terms of P.I.E. problem, and I (light) affect the classification accuracy is large. However, Taylor - ATMT algorithm accuracy than the other two is good and stable.

3) Database classification accuracy test by Yale is relatively good ORL, mainly on ORL database are consistent with the intensity of illumination Angle image is black and attitude change their poor accuracy.

4) When doing classification, because the Bayesian decision is based on distance as posterior probability estimates, therefore, the recognition rate is the same as the NNC. Although classification CCM is NNC slightly better, the difference is not big. Under illumination influence situation, however, NNC has a better performance, precision. Through the understanding of the above, to make the light affect the classification accuracy is lesser, decided to adopt mixed Taylor - ATMT algorithm and minimum error rate of bakes classifier as the tool of follow-up study.

5.2 TAYLOR – ATMT ALGORITHM EXPERIMENT

When doing the experiment, using the two methods respectively at YALE and performed on ORL database, the first method is to use a specific training group and test group, the purpose can be clearly know every image in the database because P.I.E. different effects on recognition rate. The second approach is to set up a file in the YALE and ORL database, randomly selected from the image, respectively, as training with the rest of the test, a total of ten times, take the average value is used to estimate the identification accuracy. This experiment is based on Figure 4 face recognition process based on the mixed Taylor - ATMT architecture firm do as a result, such as Table 2 to Table 4, it is worth noting is the following:

TABLE 2 random method the 95% confidence interval in the different database of recognition results

database	classifier	Taylor-ATMT
YALE	NNC	0.61 ± 0.24
ORL	NNC	0.63 ± 0.04

TABLE 3 specific methods in YALE database using NNC recognition results

YALE	YALE_a	YALE_b	YALE_c	YALE_d	YALE_e	YALE_f
Taylor-ATMT	0.42	0.59	0.70	0.21	0.67	0.69

TABLE 4 specific methods in ORL database using NNC recognition results

ORL	ORL_A	ORL_B	ORL_C	ORL_D	ORL_E	ORL_F
Taylor-ATMT	0.67	0.67	0.64	0.61	0.58	0.62

1) In the YALE database each image recognition accuracy affected by illumination, expression is very big, but in ORL database each image are greatly influenced by attitude, scale. From Table 2 shows the recognition accuracy is higher than the YALE database ORL database, this law is obviously due to the rotation, translation and scale invariant, therefore, less affected by posture, dimension problem.

2) In Table 3, YALE database within everyone's fourth picture, there are shadows, so the accuracy of the YALE_d is poorer, and the other has quite big difference. This is because the algorithm has no illumination invariant

features, so the shadow behind the (light) caused considerable conflict.

3) Within the ORL database, each image only posture and scale is different, no light, so the frequency invariant reveal recognition accuracy is relatively stable, the difference is not big.

5.3 RESULTS

Comparing existing some of the most famous algorithm is applied to face recognition of the single training sample, such as PCA and 2 dpca, (PC) 2 a, E (PC) 2 a, 2 d (PC) 2 a disturbance, etc., and SVD, according to Taylor - ATMT algorithm algorithm has higher recognition rate than other popular, and for not making a virtual image, so the instruction cycle, low storage cost.

6 Conclusion and suggestions

The important research is the face recognition problem with the signal training sample. The signal training sample means that everyone just can have a training image in some applications (the criminal identification reorganization database in the police system). It is yje extreme situation with the insufficient samples. Many traditional face recognition methods, such as Principal Component Analysis, Linear Discriminant Analysis, LDA and SVM, can not solve the problem. Therefore, the solution of the problem not only improve the availability of the face recognition system, but also can deal with the problem of the extreme small samples in the model recognition. The Taylor-ATMT is proposed to be applied in the experiments of the face recognition problem with the signal training sample, and has a better effect. Compared with other famous algorithms, the method is superior. Apart from improving the Taylor Transformation, the method is easier to combine with the Taylor Invariant compared with the AFMT and has a better recognition precision.

References

[1] Shen Y, Chen Yong 2003 Human face recognition algorithm based on wavelet subbands and decision fusion *Journal of Zhejiang University of Technology* **31**(3) 237-41
[2] Lai J 2002 Finishing the Classifier Design of Spectroface Human Face Recognition *Journal of Image and Graphics* 2002 **7**(5) 466-71
[3] Belhumeur P N, Hespanha J P, Kriegman D J 1997 *IEEE Transactions on Pattern Analysis and Machine Intelligence* **19**(7) 711-20
[4] Brandt R D 1992 Advanced Research in Machine-Printed Character Recognition *Report from the Environmental Research Institute of Michigan* Washington DC US United States Postal Service Archives
[5] Crimmins T R 1982 A *IEEE Transactions on Systems, Man and Cybernetics* **12**(6) 848-55
[6] De Castro E, Morandi C *IEEE Transactions on Pattern Analysis and Machine Intelligence* **PAMI-9**(5) 700-3
[7] Flannery D L, Horner J L 1989 *IEEE* **77**(10) 1511-27
[8] Ghorbel F 1994 A Complete Invariant Description for Gray-Level Images by the Harmonic Analysis Approach *Pattern Recognition Letters* **15**(10) 1043-51
[9] Lai J H, Yuen P C, Feng G C 2001 Face Recognition Using Holistic Fourier Invariant Features *Pattern Recognition* **34** 95-109
[10] Lin F, Brandt R D 1993 Towards Absolute Invariants of Images Under Translation, Rotation, and Dilation *Pattern Recognition Letters* **14**(5) 369-79
[11] Mallat S 1989 *IEEE Transactions on Pattern Analysis and Machine Intelligence* New York US **11**(7) 674-93
[12] Mallat S 1989 Multi-Resolution Approximations and Wavelet Orthonormal Bases of L2(R) Transactions of the American Mathematical Society **315**(1) 69-87
[13] Phillips P J 1991 Support Vector Machines Applied to Face Recognition *Proceedings of the International Conference on Advances in Neural Information Processing Systems* Cambridge MA US MIT Press 803-9
[14] Sheng Y 1989 Fourier-Mellin Spatial Filters for Invariant Pattern Recognition *Optical Engineering* **28**(5) 494-500
[15] Teh C H, Chin R T 1988 *IEEE Transaction on Pattern Analysis and Machine Intelligence* **10**(4) 496-513
[16] Turk M A, Pentland A P 1991 Face Recognition Using Eigenfaces *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* Maui HI US Institute of Electrical and Electronic Engineers 586-91
[17] Yu H, Bennamoun M 2007 Complete Invariants for Robust Face Recognition *Pattern Recognition* **40**(5) 1579-91

Authors

Zhaonan Yang, November 1974, Shangdu, China.

Current position, grades: PhD student.
University studies: Computer Science and Applied Technology
Scientific interests: video processing, data mining.
Publications: 15.

Shu Zhang, May 1973, Jixi, China.

Current position, grades: instructor.
Scientific interests: data mining technology to process database information, database security design.
Publications: 10.