# Research on Face Recognition Based on Wavelet Transform and PCA-ANN

Wang Yuanzhi[1], Luo Xiao[2]
[1]School of Computer and Information
Anqing Normal College
Anqing 246011, China
[2]School of Information and Engineering
Huzhou Teachers College
Huzhou 313000, China

*Journal of Digital Information Management*

***ABSTRACT:*** *This paper mainly studies the basic theory and technology on the identification of people's faces. The better methods are chosen and a new method is proposed. It integrates three elements (Wavelet Transform, Principal Component Analysis, Artificial Neural Networks), and the method can recognize faces more effectively. Meanwhile, this method lessens the effect of light on face appearance. Experimental result indicates that our new face recognition system is effective in the aspects of both correct ratio and speed of distinguishing.*

**Categories and Subject Descriptors:**
**G.1.2 [Approximation]:** Wavelets and fractals: **I.4.8 [Scene Analysis];** Object recognition

**General Terms:** Wavelet Transform, Face Identification

**Keywords:** Face Recognition, Feature Extraction, Principal Component Analysis, Radial Basis Function Neural Network

**Received:** 15 November 2011, Revised 17 January 2012, Accepted 22 January 2012

## 1. Introduction

Face recognition technology has developed rapidly and achieved fruitful research results, but there are still many questions: If the image background is complex, there are many similar parts of the face; face is non-rigid objects, and changes with age, it is difficult to fully describe the features; there may are many block things on face be, such as beard, glasses, hat, etc.; the images of the face are affected by the light, the angle of imaging, the distance of imaging. These problems will be solved in face recognition technology [1-5].

Face recognition step includes image preprocessing, feature extraction and face classification, etc., and for each key step in the use of different processing algorithms, and then after the combination, the recognition efficiency and speed of identification will be different, in the paper we will make the above considerations, make the key to face recognition technology select and optimize the combination of experiments, proposed a wavelet transform, principal component analysis (PCA) and artificial neural network program combines face recognition system: First, To remove the interference in the face image information, we face a series of image pre-processing, including removal of noise, grayscale normalization, scale normalization and wavelet transform; Secondly [6], KL transform based The eigenface method of face image feature extraction, image data compression dimension; Finally, the improved learning algorithm of radial basis function neural network (RBFNN) classification as a human face. Through a large number of experiments show that this system can improve the recognition efficiency and reduce light, movement, facial expressions and other conditions, the impact on the recognition rate, so as to achieve better recognition results. Use of internationally accepted test ORL face database, the correct recognition rate arrived 95.8% [7-9].

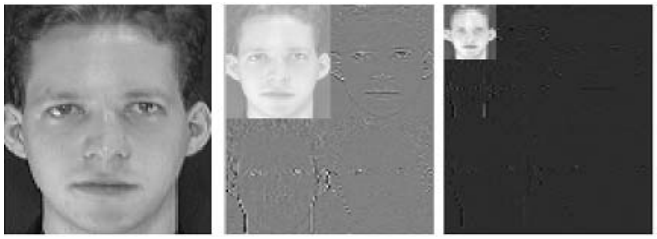
## 2. Image pre-processing based on wavelet transform

We used a single internationally accepted ORL face database shown in Figure 1. The database includes a black background at different times of the face facial expressions and details from 40 individuals, each of 1092 × 112 frontal face image composition, some of which was taken in different periods of time; facial expressions and facial details have changed, for example: do not laugh or smile, eyes open or eyes closed, with and without glasses; face attitude has changed, rotated up to 20 degrees; scale also face up to 10% of the change.

In the paper, for a face of the original image shown in Figure 2 (a) shows the results of its layer of wavelet

Figure 1. Single 10 Images in ORL Face Image Database

decomposition as shown in Figure 2 (b) below. Hominid face images through a layer of wavelet decomposition, to obtain four sub-band images. LL sub-band images which maintained a low-frequency component of the original image, as smooth as the original image; HL sub-band images of the original image to maintain the level of edge detail; HL sub-band images of the original image to keep the details of the vertical edges; sub-band images HH to maintain the original image of the ramp edge detail. For the front face recognition, since the rich people's facial expressions, and facial expression is reflected mainly in the eyes and mouth, and mouth and eyes than the vertical level of features distinctive characteristics, so the horizontal edge of the HL sub-band images portray a characteristic facial expression; outline of the face and nose than the level of vertical features distinctive characteristics, they are vulnerable to the impact of hair and the side face position; for the rigid body mode, the oblique edge of the most important information, because it represents the whole image structural features. However, non-rigid face model, oblique edge of the information by the noise, the impact of large expressions and gestures, and its stability is the worst. Appropriate level of wavelet transform low-frequency sub-band image depicts the face of constant expressions and gestures features, better stability. Moreover, an image for the n-dimensional wavelet transform, the low-frequency sub-band images of size 2n is only one original image, so the effective dimension reduction. This greatly reduces the complexity of follow-up. If you continue to make low-frequency sub-band image wavelet transform is obtained as shown in Figure 2 (c) shows the second floor of the wavelet transform.



(a) Original Image (b) First Image of Wavelet Decomposition (c) Second Image of Wavelet Decomposition

Figure 2. Multi-layer Wavelet Decomposition of Image

It can be seen from the figure; two low-frequency wavelet sub-band images after the original image is still as smooth, and the dimension of the vector image to become more low, images become more blurred. Wavelet sub band after the changes made to the face images of facial expressions and insensitivity that it obscures the faces of different expressions and different posture caused by such differences, but also greatly reduces the dimension of the vector image, but the differences between different human faces have been blurred, and with the wavelet transform of the series increases, the impact of this effect is greater. That is the wavelet decomposition of the more layers, the lower the dimension of the image vector, the smaller the subsequent complexity of the algorithm, while the same face in different expressions and the differences caused by different positions and different between human face differences will become blurred. So select the appropriate wavelet decomposition level and the effect of identifying the complexity of the algorithm is very important.

We used the wavelet low-frequency sub-image histogram graph instead of the target face image. Training histogram features for face steps include: input sample face images, Gabor for face image of two-dimensional wavelet transform to get wavelet low-frequency sub-graph as a graph similar to the original image, re-use of wavelet low-frequency sub-graph to obtain its histogram information

Experiments found that people face eyebrows, eyes, nose, mouth and other areas in the ratio of face recognition important than other areas, play a key role. While the details of changes in these areas is rich regions in the wavelet high-frequency sub-graph has a good reflection of changes in facial expression and more people from them. Therefore, we can use the low-frequency sub-graph extracted features to be identified.

## 3. Face Recognition Based on PCA

Face of the global expression has proved to be an efficient face recognition method. Principal component analysis is an optimal image compression orthogonal transformation, the aim is to find a set of data space vectors to interpret the data as much as possible the variance of the data from the R-dimensional space down to the original M-dimensional (R>> M), after the dimensionality of the data stored in the main information to make the data easier to handle. It is based on KL transform compression direction to obtain maximum information from the image information in the expression of low-dimensional space, from the information theory point of view, that is, all the orthogonal transform, KL transform corresponding to the minimum entropy, so the use of PCA method space is obtained by identifying the original image space of an optimal low-dimensional approximation. PCA for statistical feature extraction subspace form the basis for pattern recognition. It features starting from the overall image algebra, image-based information for the general classification. Service and Kirby first KL transform is used to express the optimal face image. Turk and Pentland further proposed "*Eigenface*" concept.

Discrete Karhunen and Loeve KL transform is a continuous random process as the two of them and leads the series expansion. Random image sequences is first developed by the Hotelling method of principal components, in fact, it is the KL series expansion of the discrete equivalent method. Therefore, this method has a variety of titles, such as the KL transform, Hotelling transformation, eigenvector transform, and principal component transformation. Unlike the Fourier transform of this transformation, discrete cosine transform, orthogonal transformation, which transforms the transform kernel is fixed, while the KL transform is a collection of images with different statistical properties of the nuclear matrix have different transformations that transform the nuclear matrix is a collection of images to determine the statistical properties, so the discrete KL transform is a transformation based on the demographic characteristics of the image. A non-periodic random process cannot have unrelated random Fourier coefficients of the Fourier series representation, but the relationship can be used with a number of mutually orthogonal function $\phi_n(t)$ series expansion, this expansion method is the KL expansion. The geometric point of view, PCA's basic idea is to find the best sub-space, when the multi-dimensional data x in the subspace projection, the income component has maximum variance. Also, when a new component to reconstruct the original data, in the sense of minimum mean square error of approximation of the optimal effect, that (1) the minimum value.

Set $x = (x_1, x_2, ..., x_n)^T$ is dimensional random vector of n, its mean value is $m_x = E(x) = 0$, Covariance matrix as follows:

$$C_X = E\left\{xx^T\right\} = E\left\{\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}(x_1, \cdots, x_n)\right\} \quad (1)$$

The purpose PCA of is to find an orthogonal transformation matrix $W^T = [w_1,...,w_m]$, N-dimensional vector orthogonal transformation, so the new components

$Y_i(i=1, 2,..., m)$ Between uncorrelated with each other, that is satisfied:

$$y = W * x \quad (2)$$

$$\begin{bmatrix} y_1 \\ \vdots \\ y_m \end{bmatrix} = \begin{pmatrix} w_{11} & \cdots & w_{1n} \\ \vdots & \ddots & \vdots \\ w_{m1} & \cdots & w_{mn} \end{pmatrix}\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \quad (3)$$

And y is the covariance matrix is a diagonal matrix:

$$e\left\{yy^T\right\} = E\left\{\begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix}(y_1, \cdots, y_m)\right\} = diag(\lambda_1, \lambda_2, \cdots, \lambda_m) \quad (4)$$

By (4) we can obtain

$$y_1 = W_1^T X \quad (5)$$

When all the observed data x along $w_1$ When the direction of projection, PCA will make the energy of the weight $y_1$ larger, That is, variance $E(y_1^2)$ Maximum, at this time we name $y_1$ the first principal component (PC1). The same way with $w_1$ The second vector orthogonal vector to find, $w_2$, To meet $w_1^T w_2 = 0$ Under the premise of the projector $y_2 = w_2^T x$, so the second principal component (PC2).The process is repeated in the following qualification:

(1) A new direction with the front in all directions orthogonal $w_i^T w_j = 0, \forall_j < i, \| w_i \| = 1$;

(2) Projection data with the largest variance.

Also, because each vector $w_i$. The projection direction orthogonal to each other, so get the new components in PCA $y_i = w_i^T x\ (i = 1, 2, ..., m)$ Between each other, unrelated, that is, $E\{y_i\ y_j\} = E\{(w_i^T x)(w_j^T x)\} = 0,\ i \neq j$.

PCA algorithm is achieved within the two-dimensional space of principal components to extract image features from face images of a set of source images (Source Image) generated. The source image to form the human face, a group of space-based images, and these images can also be seen as the base is a group of people related to each other face image features, while the face is obtained by the characteristics of these different combinations of coefficients of face images in the line under of the combination.

PCA done on the image, we will use every one-dimensional

image into a one-dimensional column vector stacked to form columns of the input matrix $X$ vector. You PCA output matrix $Y = WX$ column vector for the image, known as the face image source image characteristics, the characteristics of face images that face space. Shown in Figure 3 and Figure 4, features a linear combination of face images of different forms of each face image database, PCA is a linear combination with these different factors to describe each one different face images, while the mixing matrix A is unique for each row vector corresponding to the face of each piece of the linear combination coefficient, A is the same matrix $W$ can be obtained by calculating the separation.

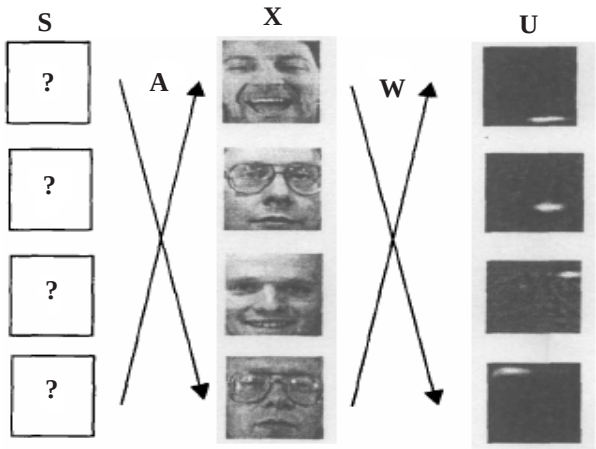



Figure 3. Image Generations and Decomposition Model

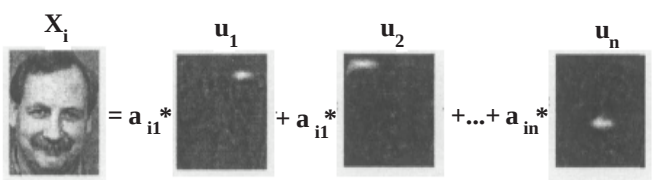



Figure 4. Linear Combination (a1, a2... an) of PCA Principal Component

Experiments show that facial images after KL transform and greatly reduce the original dimension of the space, increase the speed of face recognition.

### 4. Classifier design Based on RBF neural network

Neural networks as a new mathematical modeling approach, the study of historical data, can be found in the mapping between input and output relationships. Mapping process is a real feature extraction, pattern recognition process. Patterns in the learning process to extract and store it in the forecast period, the network can be input directly generates the next time the forecast output. How the accuracy of prediction depends not only on the adequacy of training data is typical, but also the structural characteristics of the network and algorithms have a direct relationship. When BP neural network was used in face recognition, although it can complete the identification and classification task, but the error is relatively large. This is mainly with the BP network and BP algorithm itself defects.

RBF neural network can be changed according to the specific circumstances of the network structure and network parameters, a self-learning, self-organizing, adaptive function. At the same time, learn it fast, there is no local minimum problem. RBF neural network hidden layer of nonlinear radial basis function, the input vector produces a localized response, the input space only when the input falls within a specified small, the hidden layer unit will make a significance of non-zero response.

It put the output layer, the output value of the radial basis function linear summation to arrive at the desired results. RBF network for nonlinear uniform approximation of continuous functions with, it can be a wide range of data fusion, high-speed parallel processing of data. Because more of these characteristics, can be used RBF network for face recognition, it can be a different face classification, association, integration, coordination, to achieve the learning face image potential between the law, to achieve fast, easy recognition effect.

RBF network and BP network learning process is similar to the forward and backward propagation of two parts, but its structure is rather special, it is more convenient in practical applications. It's the learning process described as follows: the forward part of the network input in the input data to calculate the output value of hidden layer node Rj (j = 1, ... u), then modify the weights W. Determine the weight W, the network continues to forward data to send, we can calculate the size of the error El. In the back-propagation in the error signal before the signal along the same route back to the input. Fixed weights, and then according to equation (5) and (6) adjust the RBF neural network hidden nodes in the center of the data and extended constants.

For each hidden node, the center distance and adjust the amount of expansion constants:

$$\begin{aligned}
\Delta C^l(i,j) &= -\xi \frac{\partial E^l}{\partial C^l(i,j)} \\
&= -\xi \frac{\partial E^l}{\partial y_k^l}\frac{\partial E^l}{\partial R_j^l}\frac{\partial R_j^l}{\partial C^l(i,j)} (i = 1,2,....r,\ j = 1,2...,u) \\
&= 2\xi \sum_{k=1}^{s}(t_k^l - y_k^l)\cdot w^l(k,j)\cdot R_j^l \cdot \frac{P(i,j) - C^l(i,j)}{(\sigma_j^i)^2} \qquad (6)
\end{aligned}$$

$$\begin{aligned}
\Delta\sigma_j^l &= -\xi \frac{\partial E^l}{\partial \sigma_j^l} \\
&= -\xi \frac{\partial E^l}{\partial_k^l}\frac{\partial y_k^l}{\partial R_j^l}\frac{\partial R_j^l}{\partial \sigma_j^l} (j = 1,2,...u) \\
&= 2\xi \sum_{k=1}^{s}(t_k^l - y_k^l)\cdot w^l(k,j)\cdot R_j^l \cdot \frac{\left\| p_l - C^j \right\|}{(\sigma_j^l)^3} \qquad (7)
\end{aligned}$$

Where, $\xi$ is the learning rate, $P(i,j)$ is *i-th* input vector of L study.

Learning process in the selection of learning rate is also

very important, the size of the network learning process greatly. If too small, the slow convergence of network training, training times increase; rapid change in the error area, too large because of the adjustment value is too large across the narrow "*potholes*", and to give training a shock, but increased iteration number of times. To accelerate the convergence process, a good idea to change the learning rate is adaptive to the number of iterations with errors and changes.

This paper is to use the ORL face library training and testing. ORL face library by 40 people each composed of 10 photos, each of the first five to take the total 200 images as training set for the RBF network training, and then take the rest of the 200 pictures on the network as a test set for testing. In the feature extraction part, the different KL threshold will be different dimension of feature vectors, making the face recognition part of the RBF network input dimensions are different, thus the effect of network structure and identification have had an impact. Experimental data show that the RBF network training in KL threshold of 0.99 when the batch of 40 have the best recognition rate of 95.8%, verify the implementation of this method is effective.

## 5. Design and Implementation of prototype system

The main subject is to develop a neural network-based face recognition experimental system. The system consists of preprocessing, feature extraction, classification based on neural network and database of four parts. Feature extraction and classification of which is the key to solve the problem of face recognition. Figure 5 shows a system block diagram:

System overall process is: first the image pre-processing to make it into the normalized image; then the resulting image further wavelet transform to extract frequency information; then use principal component analysis, feature extraction methods; final adoption RBF neural network classifier constructed to classify the output face image classification result is the result of face recognition.

Classifier which face the implementation process is: First, from the ORL face library to extract face image, and then through the image pre-processing, and then the processed image feature extraction, and finally through the RBF neural network training, the trained is the face of artificial neural network classifier.

We use the internationally accepted ORL face database to validate our proposed face recognition framework. Experiments show that the method proposed in this paper on the ORL face database can achieve good results, under different circumstances for the experimental results are listed below.

| Classification | Recognition Rate |
|---|---|
| Wavelet Transform + PCA | 76% |
| PCA+ Minimum Distance Method | 78% |
| PCA+ANN | 93.4% |
| Wavelet Transform + PCA + ANN | 95.8% |

Table 1. Identification Results under Different Circumstances

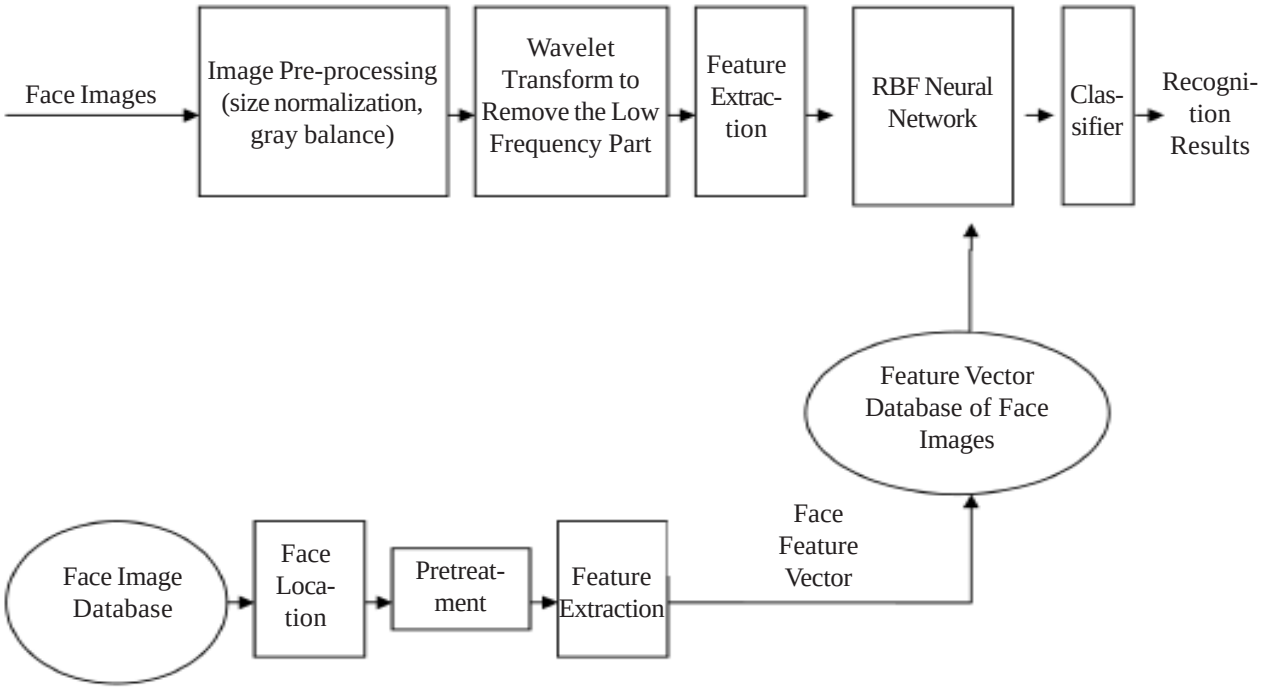



Figure 5. Face Recognition System Based on Principal Component Analysis and Neural Network

As can be seen from Table 1, the image preprocessing, feature extraction and classifier design and other key technologies in different combinations, the identification efficiency is different.

(1) Method of wavelet transform + PCA: wavelet transform is introduced for normalization and further pre-processing of images, extract low frequency information, and then use the PCA method of classification. The advantage is to identify the faster, the change of expression is not sensitive to the disadvantage of recognition efficiency is not high, only 76%.

(2) PCA + minimum distance method: refers to the normalized images using the PCA method for feature extraction, and then use the minimum distance method of classification. Recognition efficiency of this method is not high, and the human face images normalized higher and its performance are vulnerable to the effects of light and gesture.

(3) PCA + ANN method: refers to the normalized images using the PCA method for feature extraction and RBF-based neural network for classification. This identifying efficiency of the method is up to 93.4%, but the face and still very sensitive to changes in illumination.

(4) Method of wavelet + PCA + ANN: Face of the three key technologies will be optimized. First floor in the picture pre-processing stage using wavelet transform, remove the high frequency image information, whichever is low frequency information; then use PCA method for feature extraction; Finally, neural network classification. This method not only identify the fast and efficient identification, while light, less sensitive to changes in expression, with better robustness.

## 6. Acknowledgment

The work is supported by the Natural Science Key project of Anhui province Education Department of China under Grant Nos.KJ2008A18ZC,KJ2010A232 and Higher Education Research Project of Huzhou Teachers College under Grant Nos.GJC11026.

## 7. Conclusion

In this paper, the wavelet transform + PCA + ANN method has the advantages of higher identification speed, higher recognition efficiency and robustness. This system can be used to identify the face in a medium-small face database, but the feature extraction part of the system is based on KL transform, the feature space of image is extracted according to all face vector in face database. Thus the expansion of the face image database will require the re-finding of the feature space and extraction of the feature vector. There is a need to re-train the neural network.

## References

[1] Turk, M.(1991). A Pentland, Eigenfaces for Recognition[J], J*ournal of Cognitive Neuronscience*, 3 (1)102-111

[2] Wang Wei, Ma Jianguang, (2002). Development of common method of face recognition [A]. *Ordnance automated automatic detection technology*, 21(1) 121 - 132.

[3] Zhao Z. L., Liu, B., Li, W. (2012). Image Classification Based on Extreme Learning Machine, *IEIT Journal of Adaptive & Dynamic Computing*, 2012(1) 5-11. DOI=10.5813/www.ieit-web.org/IJADC/2012.1.2

[4] Zhou Delong, Zhang Xiaohua, Liu Bo, et al. (2003). JDI introduce large-scale face image database [A]. Biometric Research Progress (2) [C]. Beijing: Tsinghua University Press, 118 - 120

[5] Wang, Sheng Zhang, Fang Fang, (2006). Face detection and recognition technology review [A]. Hefei University of Technology (Natural Science), 29 (2)136 - 145, February

[6] Brunelli, R., Poggio, T. (1993). Face Recognition, Features Versus Templates, *IEEE Trans on Pattern Analysis and Machine Intelligence*, 15(10) 1042-1052

[7] Bledsoe, W. (1966). Man—Machine Facial Recognition, Technical Report, PRI, 22, Panoramic Research Inc, Palo Alto, USA, 156-167

[8] Wang, Yuanzhi , Zhang, Hui; Jiang, Shanhe, Sun, Lijuan, (2010). Fuzzy support vector machine based on improved possibilistic c-means clustering algorithm in kernel space [J]. *Journal of Information and Computational Science*, 7(2) 585-591, February.

[9] Wang, Yuanzhi; Chen, Liwei; Zhang, Fei, Jiang, Shanhe, Sun, Lijuan, (2010). Analysis and study of embedded capability to HSBH by SPA[J]. *Journal of Software*, 5 (11) 1256-1261.

[10] Gelan Yang , Yuanzhi Wang , Wanneng Shu , Huixia Jin (2009). Efficient Parallel Genetic Immune Clonal Algorithm for Information Retrieval from Textual Web Documents [J]. *Journal of Information and Computation Science.* 6 (2) 837-844

[11] Kanade, T. (1974). Computer Recognition of Human Faces, Ph, D Dissertation, Kyoto University, Japan, 130-146.

[12] Zhao Z. L., Liu, B., Li, W. (2012). Image Clustering Based on Extreme K-means Algorithm, *IEIT Journal of Adaptive & Dynamic Computing,* (1) 12-16, Jan. DOI=10.5813/www.ieit-web.org/IJADC/2012.1.