Face orientation recognition based on multiple facial feature triangles

Linlin GAO[1], Yingkai XU[2]
[1]Computer and communication engineering institute, Northeastern University at Qinhuangdao, Qinhuangdao, China
[2]Nanjing University of Aeronautics and Astronautics College of Science, Nanjing, China
452877954@qq.com, kevinhsu1008@gmail.com

Abstract: This paper proposes a new method that combining multiple feature triangles with BP neural network, to improve the efficiency and accuracy of face orientation recognition. Based on the traditional index—the inverted triangle formed by pupils and nasal tip, we find another feature triangle formed by nasal tip and corners of mouth. First we do image preprocessing which includes smoothing linear filter, edge detection and so on. Then both rough and precise detection of feature points are done. Next we extract feature triangle based on two-dimensional plane. Finally BP neural network is used for face orientation recognition. Experimental results show that an approximately 90% success rate is achieved. They also reveal that our new method improves the recognition effect.

Keywords: face orientation recognition; facial feature triangle; BP neural network; image preprocessing; feature point detection

1. INTRODUCTION

Recently, face recognition has drawn considerable attention from researchers, especially in the areas of pattern recognition and image processing. Face recognition is a computer-assisted method for biological characteristic identification. It is valuable in many applications such as security control systems, human-machine interfaces and video retrieval. Face orientation recognition is an important step to identify variable face data and simulate facade or 3-D image from non front one.

Human face contains many significant feature points such as eyes, nose and mouth. They are often selected as face location indexes in traditional method. Since more indexes lead to more neural network dimensions, longitudinal digit capacity and learning process are increased, even worse, the optimal solution can not be obtained. This is the drawback to facial feature point extraction

Multi-layer feed forward network based on BP algorithm is the main neural network in pattern recognition. It has very strong computing capability and is suitable for complicated non linear environment, so it has extensive applications, especially for classifier design. But it also has drawbacks--slow convergent speed and easily getting into local minimum. If small samples are provided, traditional BP network has many big drawbacks. Concretely, quite slow learning rate leads to long training time; learning and memory of network layers are unstable; global weight convergence can not be guaranteed and so on. Many researchers have proposed improved methods such as batching, momentum, variable learning rate (VLBP)[1] and genetic algorithm (GA)[2]. The improvement is significant, but results are still not perfect in efficiency and accuracy. Thus, this paper presents a new method that combining improved face feature triangle with BP neural network.

2. METHOD SUMMARY

Based on numerous abroad references, this paper selects some images as standard from face database, and make others automatic identification and classification. Then face orientation recognition model is built. Quoting the conclusion put by C. Lin, K. C. Fan[3], we suppose facial feature triangle reflect orientation variations comprehensively. Traditional method extracts all the feature points and reduce dimensions by abandoning some randomly, which cause losses to some degree. Choosing facial feature triangle as input layer of BP network will improve the drawback.

We select 10 people as target, each has 5 different face-angle images, that is, total 50 samples are grouped into 5 categories. So we transform face orientation recognition into matching detection between samples and features.

Recognition process is shown in Figure 1, training samples come form each group with same face-angle images.

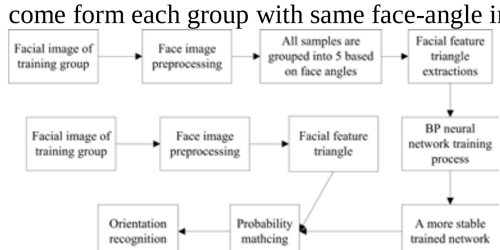

Figure 1. Overview of recognition system

3. FACE IMAGE PREPROCESSING

Many factors such as noise and defocus influence image quality and the accuracy of extracting facial features significantly. Thus, we first do the image preprocessing, which includes smoothing linear filter, median filter, edge detection and so on.

Filter technique is widely applied to noise removal and image enhancement. The results are smoothing non marginal region and preserving edge. Binary image is used in image segmentation. This paper selects method of analogy and tolerance, so that a clear continuous image is achieved, feature points are located and extracted precisely. Edge detection is a common approach to detecting discontinuities in gray level and has numerous algorithms. We choose Sobel operator because of its detection result and small noise effect.

Figure 2 shows the integrated image preprocessing.

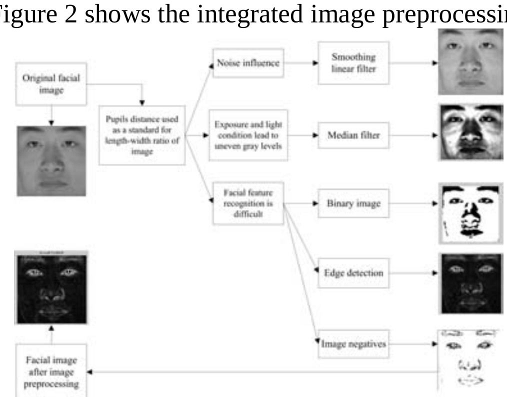

Figure 2. Overview of image preprocessing

4. FACIAL FEATURE POINT DETECTION

After image preprocessing, facial feature point detection is done. Figure 3 shows the concrete process.

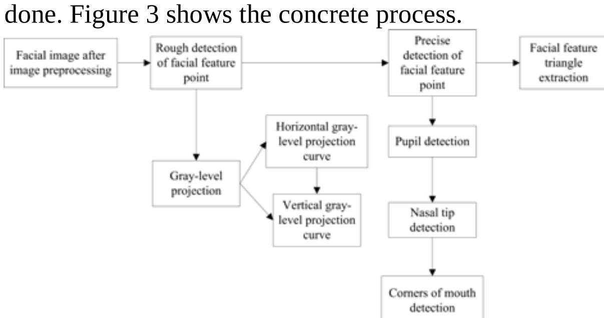

Figure 3. Overview of feature detection

4.1. Rough Detection of Facial Feature Point

We make face region's horizontal and vertical gray-level projections to obtain gray image, as shown in Figure 4.

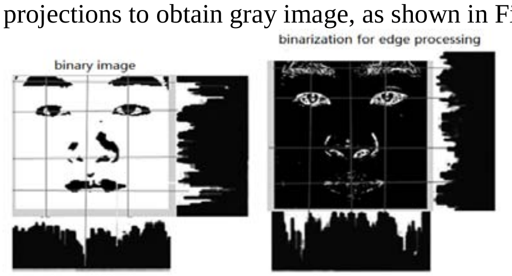

Figure 4. Rough feature detection principle

Analyzing gray-level projection curves, we find that when scanning horizontal gray-level projection from up to down, 4 wave troughs separately correspond to the positions of eyebrows, eyes, nose and mouth. While the vertical one shows that two symmetrical troughs between cheeks correspond to eyes. Thus, rough detections of facial feature points are finished.

4.2. Precise Detection of Facial Feature Point

- Pupil detection

Pupils and the surrounding have the smallest gray value. We do edge detection for these areas, choose average coordinates as pupil positions and do histogram equalization. Finally, pupils' horizontal positions are detected.

- Nasal tip detection

With the influence of natural light, in front-view image, the position of nose bridge is brighter than that of the two sides. Along nose bridge direction, the maximum gradient is found. We suppose eyes distance be 1, then search duck area at the distance of 017 to 1 ranging from the eyes distance center. Combining former rough detection, nostril position can be obtained. The position of nose bridge is a highlighted area and 015 distance above nostril.

- Corners of mouth detection

Corners of mouth have distinguished feature points. We first make horizontal gray-level projection of mouth outline, then by getting the product of vertical gradient and original image, vertical projection is computed. If the distribution of horizontal projection in local troughs presents high value, corners of mouth position can be detected.

5. FACIAL FEATURE TRIANGLE EXTRACTION

5.1. Facial Feature Triangle Analysis

In human face, pupils and nasal tip form an inverted triangle, which is the basic element of face recognition. To detect face orientation precisely, we find the second feature triangle formed by nasal tip and corners of mouth, as shown in Figure 5. The two triangles are considered as primary indexes of face orientation recognition.

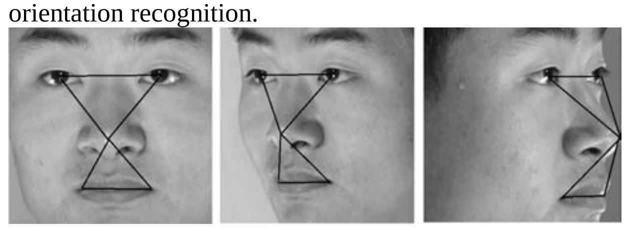

Figure 5. Frontal and slant view of feature triangle

5.2. Index Location on Two-dimensional Plane

Based on the law of head movement, we suppose head rotation be centered for media axis. In Figure 6, taking eyes for example, line *A B* is eyes distance; line *C D* is the same distance after head rotation; Line *C' G* is the projection of line *C D* on frontal plane, that is, eyes distance on two-dimensional plane after rotation.

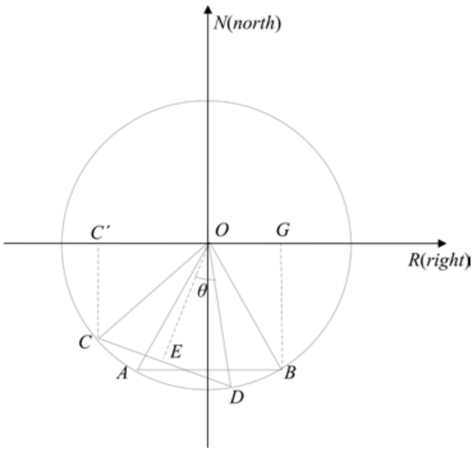

Figure 6. Top view of feature triangle with head rotation

5.3. Index Selection Based on Facial Feature Triangle

Due to multiple faces in complex backgrounds, distance comparisons are inaccurate. We select side length proportions of the two feature triangles as input layer, analyzing indexes respectively. In Figure 7, *A*, *B* represent positions of right and left eye; *C* is the middle point of line *A B*; *O* is the position of nasal tip; *D*, *E* represent corners of mouth; *F* is the the middle point of line *D E*.

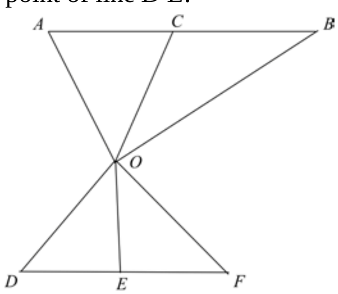

Figure 7. The distribution of facial feature triangles

The final indexes are determined as follows: there are 8 feature triangles. Group one is composed of triangle *OA/AB*, *0C/AB* and *OB/AB*; group two contains triangle *OD/DE*, *OF/DE* and *OE/DE*; triangle *AB/DE* and *OC/OF* are used to compare. Table 1 shows the index classifications.

Table 1. BP network index based on feature triangle

Index Classification	The First Feature Triangle			The Second Feature Triangle			Comparison	
Index explanation	i_0	i_1	i_2	i_3	i_4	i_5	i_6	i_7
	OA/OB	*OC/AB*	*OB/AB*	*OD/DE*	*OE/DE*	*OF/DE*	*AB/DE*	*OC/OF*

6. FACE ORIENTATION RECOGNITION BASED ON IMPROVED BP NEURAL NETWORK

6.1. BP Neural Network Theory

The basic idea of BP is that the learning process is composed of positive diffusion of signal and inverted diffusion of error. BP neural network consists of three parts: input layer, hidden layer and output layer. We suppose each be *k*-layer($k=1,2,...,n$) and the number of neurons is $N_1, N_2, \ldots, N_k$ respectively. $u_k(i)$ is the information accepted by i^{th} node in layer k. In sample P, $W^P_k(i, j)$ is the weight from j^{th} node in layer (k-1) to i^{th} node in layer k; $a^P_k(i)$ is the output of i^{th} node in layer k; $\theta^P_k(i)$ is the threshold of i^{th} node in layer k. Neurons exchange information between each layer and the information transmission is from input layer to output layer.

So the relationship between input and output is demonstrated as follows:

$$u_k(i) = \sum_{j=1}^{N_{k-1}} W_k(i, j) a_{k-1}(j) + \theta_k(i),$$

$$a_k(j) = f(u_k(i)), \quad (k = 2,3,\cdots,n;\ i = 1,2,\cdots,N_k) \tag{1}$$

The modified formula of weight coefficient is:

$$W^p{}_k(i, j) = W^{p-1}{}_k(i, j) + \eta \delta^p{}_k a^p{}_{(k-1)}(j) \tag{2}$$

where η is the learning step; $\delta^P{}_k$ is the error.

6.2. The Improved BP Algorithm

Due to the drawbacks of traditional BP network such as slow convergent speed and easily getting into local minimum, we bring up improved project.

- Connection weight updating

To restrain oscillation of iterative process, additional momentum is added in weight updating. So connection weight now is:

$$W^p{}_k(i, j) = W^{p-1}{}_k(i, j) + \eta \delta^p{}_k a^p{}_{(k-1)}(j) + \beta \Delta W^{p-1}{}_k(i, j) P - 1SY \tag{3}$$

where η is the learning step; β is momentum factor; $\Delta W^{p-1}{}_k(i, j)$ is the weight change in iteration $P-1$. From the function we know that weight change in iteration P is related to this iteration. Since strong correlation exists between samples, learning result of a sample is used for the next one, which can increase convergence rate.

- Limiting "sigmoid" activation function output

Connection weight updating relates to hidden layer output and is determined by sigmoid activation function. If the function reaches its limiting value of 0 or 1, weight change is 0, without updating. Due to saturation nonlinearity of function characteristic, values near 0 and 1(say, 0.05 and 0.95) define low and high values at output of neurons, which leads to slow updating rate and small corrected value. So limiting sigmoid function output is necessary.

6.3. Face Orientation Recognition Implementation

- Input eigenvector and output target vector

Input determination is in fact the process of eigenvector extraction. We choose characteristic quantity of facial feature triangle as input. Considering facial images have five different orientations, network output mode can be demonstrated as shown in Table 2.

Table 2. Output code and face orientation

Face Orientation	Right	Right-front	Front	Left-front	Left
Output code	5	4	3	2	1

- The number of network layers and neurons

Face orientation recognition by BP mainly determines network structure. It has been proved theoretically that the three-layer feed forward network can be trained to approximate arbitrary function well. The number of nodes in hidden layer is determined by the following function:

$$N = sqrt(m+n)+a \tag{4}$$

where m is the number of nodes in the output layer; n is the number of nodes in the input layer; a is the constant between 1 and 10.

Based on the function and gradually pruning method, the final number of hidden units is 9.

- Initial weight, learning rate and anticipation error

We select random number between 1 and 10 as initial weight, so that sigmoid activation function can be adjusted when variation is maximal. Learning rate is specified as 0.01 to guarantee convergence speed and system stability. Figure 8 shows change curve of learning rate.

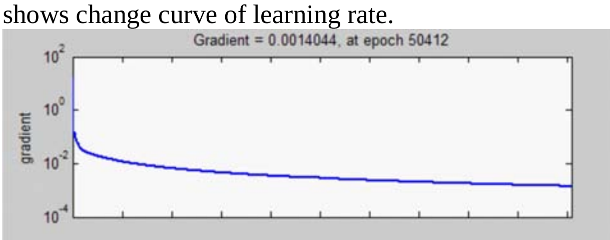

Figure 8. Learning rate change curve of 9 hidden units

During training process, anticipation error should match neurons number of hidden layer. This paper specifies anticipation error as 0.0001, training parameter is shown in Table 3.

Table 3. Training parameter

Parameter	Training Time	Training Target	Learning Rate	Displaying Frequency	Momentum Constant
Value	8000	0.001	0.01	50	0.9

The final neural network is shown in Figure 9.

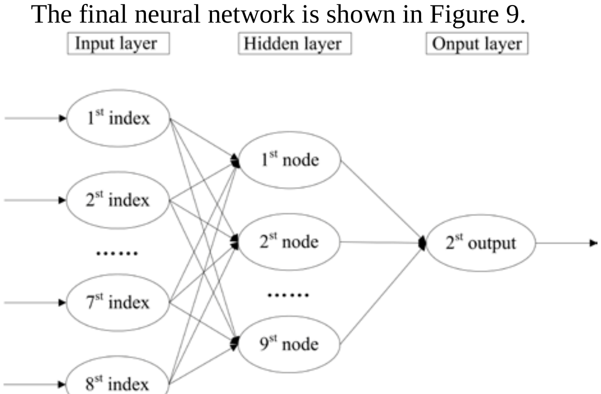

Figure 9. Final BP neural network structure

- Neural network training

Database has 50 images and we randomly select 30 of them to form a training set. Training sample parameter is shown in Table 4.

Table 4. Training sample parameter

Name	1_1	1_2	1_3	1_4	2_1	2_2
Orientation code	5	4	3	2	5	3
Name	2_4	3_1	3_4	4_2	4_3	4_4
Orientation code	2	5	2	4	3	2
Name	4_5	5_1	5_2	5_5	6_1	6_2
Orientation code	1	5	4	1	5	4
Name	6_3	6_4	7_2	7_3	8_2	8_4
Orientation code	3	1	4	3	4	1
Name	9_2	9_3	10_2	10_3	10_4	10_5
Orientation code	4	3	5	3	1	1

We use MATLAB toolbox to create BP neural network, the parameters are shown in Table 5.

Table 5. Network parameter in MATLAB toolbox

Training Function	Learning Parameter	Performance function
traingdm	learngdm	mse

Figure 10 shows training error curve, which reaches the target value of training effect.

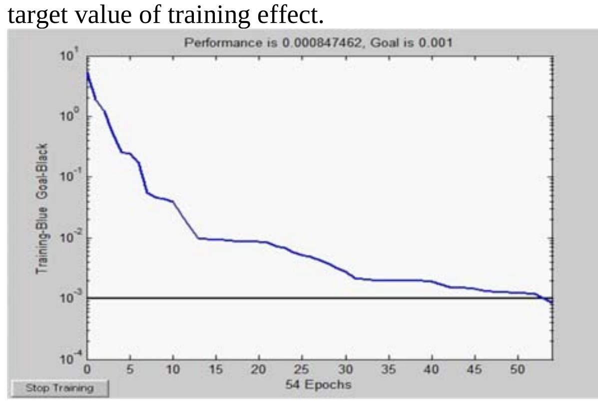

Figure 10. Training effect curve

7. TEST ANALYSIS

We use the rest of images to set up detection sample database. By substitution of the sample data into the trained BP neural network, simulation results show tested image orientation. The comparison with sample facial orientation is shown in Table 6.

Table 6. Detection sample parameter

Name	1_5	2_3	2_5	3_2	3_3
Orientation code	5	3	1	5	3
Fitting code	2	3	1	5	3
Name	3_5	4_1	5_3	5_4	6_5
Orientation code	1	5	3	1	2
Fitting code	1	5	3	1	5
Name	7_1	7_4	7_5	8_1	8_3
Orientation code	5	4	1	5	3
Fitting code	5	4	1	5	3
Name	8_5	9_1	9_4	9_5	10_1
Orientation code	100	5	4	1	5
Fitting code	1	5	4	1	5

Combining Figure 10 with Table 4, we conclude that sample size is 20 and the correct number is 18, that is, recognition rate is up to 90%.

Although limited sample size influences training and detection results, in this simulation, the output value is close to the anticipation one, which still demonstrates the feasibility of tested BP network. Since sample extractions are random, it is of great application value.

8. CONCLUSIONS

Based on the inverted triangle formed by pupils and nasal tip and the erected triangle formed by nasal tip and corners of mouth, we use smoothing filter, binary image, edge detection and facial feature geometric analysis for index location and extraction. During the process of facial orientation recognition, we select improved BP network. Experimental results show that recognition rate is up to 90%. When sample size has small change, recognition rate fluctuates in normal range.

We combine multiple feature triangles with improved BP algorithm, solving the problem of different image sizes or formats and extending angle recognition scope to three-dimensional space. But this method is also deficient, cluttered images such as partial occlusion of mouth and wearing sunglasses are not analyzed. The next step will consider further improvement in terms of altered circumstance, efficiency and accuracy.

REFERENCES

[1] T. P. Vogl, J. K. Mangis, A. K. Rigler, W. T. Zink, D. L. Alkon. Accelerating the convergence of the back-propagation method. Biological Cybernetics. 9, 59 (1998).

[2] D. E. Goldberg, J. H. Holland. Genetic algorithms and machine learning. Machine Learning. 10, 3 (1988).

[3] C. Lin, K. C. Fan. Triangle-based approach to the detection of human face. Pattern Recognition. 34 (2001) vol(34).

[4] L. H. Koh, S. Ranganath, Y. V. Venkatesh. An integrated automatic face detection and recognition system. Pattern Recognition. 35 (2002).

[5] Adnan Khashman. Application of an emotional neural network to facial recognition. Neural Comput & Applic. 18 (2009).

[6] M. H. Mahoor, M. A. Mottaleb. A multimodal approach for face modeling and recognition. In Proc. IEEE Transactions on Information Forensics and Security, 9 (2008), pp. 431-439.

[7] Yan Xu, Xiaofeng Zhang, Huaicun Gai. Quantum neural networks for face recognition classifier. Procedia Engineering. 15 (2011).

[8] Hongyuan Wang, Lei Zhao. A face recognition method based on DiaPCA and neural network. In Proc. IEEE International Conference on Electrical and Control Engineering, (2010), pp. 53-56.

[9] Ying He, Baohua Jin, Qiongshuai Lv, Shaoyu Yang. Improving BP neural network for the recognition of face direction. In Proc. IEEE International Symposium on Computer Science and Society, (2011), pp. 79-82.

[10] A. Al, Editor, Personal descriptions, English Scientific Publishers, London: Butterworth (1950).

[11] Parke F I. Computer generated animation of faces. University of Utah, Salt Lake City, UT,UTEC-CSc-72-120 (1974).