2016 12th World Congress on Intelligent Control and Automation (WCICA)
June 12-15, 2016, Guilin, China

Eye-Gaze Tracking System Based on Particle Swarm Optimization and BP Neural Network

Liling Yu, Jiangchun Xu*, and Shengwang Huang

***Abstract*—In order to enhance the practicability and accuracy of the eye-gaze tracking system, a new type low pixel eye feature point location method is adopted. This method can accurately extract the eye-gaze features, namely iris centre point and canthus points when the image pickup requirements are low. The eye-gaze tracking method based on particle swarm optimization (PSO) BP neural network is raised, to capture pictures of eyes under the same environment, and a regression model where the connection weights and threshold values are optimized by PSO algorithm is built via BP network. This method is free of the inherent defects of BP network. This method requires only a common camera and normal illumination intensity rather than high-standard hardware, which greatly cuts the restrictive requirements for the system hardware and thus enhances the system practicability. The experiment results show that PSO-BP model is of higher robustness and accuracy than BP model, and is of higher recognition rate and can effectively enhances the eye-gaze tracking accuracy.**

I. Introduction

HUMAN eye-gaze tracking plays an important role in many fields, like human-computer interaction, advertisement analysis, psychoanalysis, etc. There are many methods for realizing eye-gaze tracking, and these methods fall into 2D eye-gaze tracking ones and 3D eye-gaze tracking ones generally [1]-[3].

For 2D eye-gaze tracking, usually Pupil-Cornea Reflection technique where infrared source casts onto human eyes and generates hot spots on corneas [4]-[5], and hot spot center and pupil center together generate a vector quantity that corresponds to human eye gaze points and can indicate the location of the human eye gaze location is adopted. For 3D eye-gaze tracking method [6]-[7], which is through the establishment of eye model, where used of two or more infrared light source and form a number of bright spots in the eyeball, and then according to these bright spots through a number of complex formula to calculate the eye gaze point. These methods are inevitable to use infrared light source or multiple cameras, infrared light source not only increases the hardware cost of the system, but also produced some damage for human body.

This work was supported in part by the China National Science and Technology under Project Code KKQW201003004.

*Jiangchun Xu is with Faculty of Information Engineering and Automation, Kunming University of Science and Technology, Kunming, CO 650500 China (corresponding author, phone:+86-0871-5916643;fax: 0871-3304192; e-mail: jx19631018@163.com).

Liling Yu and Shengwang Huang are with Faculty of Information Engineering and Automation, Kunming University of Science and Technology, Kunming, CO 650500 China (e-mail: 478675082@qq.com).

Therefore, a 2D eye-gaze tracking method based on neural network is adopted in this dissertation. This method requires no infrared source, and reaches a relatively high accuracy even though the single camera's pixel is not high. Optimizing BP network by PSO and replacing the gradient modification of BP algorithm by particle swarm iteration can cut the time for network training and enhance the convergence rate of the algorithm. In addition, this improved algorithm can cut the influences of human factors, fast and accurately reflect the human eye images and fixation points, and is free of the defects of the conventional BP neural network, like slow convergence, locally optimal solution, low generalization performance, etc [8]. The improved system does not require infrared source and multiple cameras, and thus cuts the requirements and cost for the hardware and enhances the system practicability.

II. System Introduction

In this paper, the method of classification is used to detect the pupil center and the corner of the eye. It not only improves the accuracy of the detection, but also enhances the stability of the system. Among them, the key step are the eye of the iris center, eye corner detection and feature vector selection, PSO-BP neural network structure and training methods, which are directly affect the accuracy of the whole system. The flow chart of the system is shown in Fig. 1.

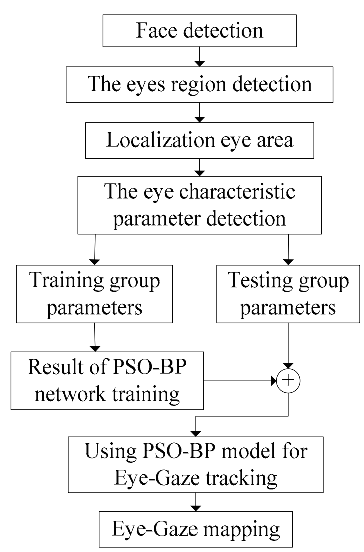

Fig. 1. System flow chart

A. Face Detection

Face detection played a very important role in computer vision. There are many methods of face detection, which are based on the method of correlation matching, geometric features and statistical theory, etc [9]. In the face detection method based on statistical theory, the AdaBoost based learning algorithm is the most widely used face detection algorithm, which is proposed by Viola in 2001. Due to the effectiveness of the method, this method is used to detect the human face in the system. Using this method, the face detection method is divided into two steps. The first step is to train the classifier based on certain features. The second step is to carry out the face detection using the trained classifier. Most of the face detection classifiers are based on Haar-like features and then trained by AdaBoost learning algorithm. After the classifier is trained, it can be detected by the region of interest in the input image. This system is based on Opencv platform, Opencv with trained classifier AdaBoost face detection data, and stored in XML files. When the system is started, the file is loaded to generate a classifier. Then the input to detect the image, the classifier can detect the face.

B. Iris Center Detection

Iris center detection is generally used in the method of threshold and ellipse fitting. However, this method requires high pixel camera, and low stability. The detection effect of the eye pupil center is not ideal in the image with low pixel. In this paper, a new method for detecting the effect of the method with a low pixel is used to detect the pupil center detection method based on the gradient vector [10]. The gradient vector of every direction in the image is obtained. The most concentrated point of the gradient vector intersection is the iris center. Specific calculation methods are as follows. As shown in Fig. 2, let c be a possible pupil center, x_i is the pixel position. $i \in \{1，…，N\}$.N is the pixel size of the image. g_i is in the x_i of this gradient vector. d_i is from c to x_i between all of the displacement vector. Then the vector g_i and the gradient vector d_i have the same direction. In fig. 2, c is not the central point in the left graph, so the d_i and g_i directions are not the same. So, we can confirm the pupil centre point by summing dot products of all displacement vectors d_i of each possible pupil centre point c and the corresponding gradient vectors g_i. The mathematical formula is as follows.

$$c^* = \max_c \{\frac{1}{N}\sum_{i=1}^{N}(d_i^T g_i)^2\} \tag{1}$$

$$d_i = \frac{X_i - c}{\|X_i - c\|}, \forall i : \|g_i\|_2 = 1 \tag{2}$$

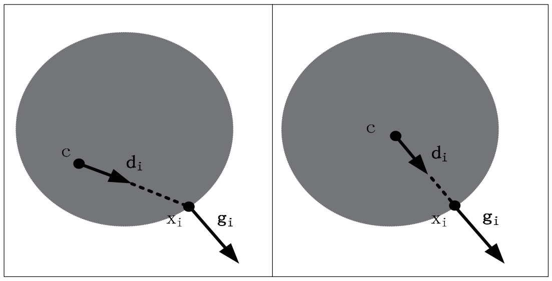

Fig. 2. Gradient vector pupil detection model

Among them, in order to improve the robustness of the light condition, the displacement vector d_i and the gradient vector g_i could be unitization.

C. Eye Corner Detection

Eye corner in image processing, simple to understand is the image on the horizontal direction and vertical direction larger changes of the point. Here the use of more classic Harris corner detection method [11]-[12]. In order to detect the corner points of the eye accurately, the face image is needed for the Gauss filtering before the corner detection to remove the effects of all kinds of noise. Then according to the detected iris center, the eye image is obtained with the size of 128×64 at this point. After removing the influence of eye iris part, the Harris corner detection is performed in this area.

III. BP Neural Network based on Particle Swarm Optimization Model

A. BP Neural Network

BP network is a multilayer feedforward network trained by the error back propagation algorithm. It can learn and store a large number of input and output mode mapping relationships without prior revealing the mathematical equations describing the mapping relationship.

Its learning rule is to use the steepest descent method, by back-propagation network to continuously adjust the weights and thresholds, so the network and the minimum sum of squared errors. The topological structure of BP Neural Network model includes input layer, hide output and output layer [13]-[14].

B. Particle Swarm Optimization

Particle Swarm Optimization, abbreviated as PSO, is a new evolutionary algorithm which has been developed in recent years [15]. The PSO algorithm, similar to genetic algorithm, is a kind of evolutionary algorithm, which is based on the random solution, and the optimal solution is obtained by iteration. And the quality of the solution is evaluated by the fitness of the degree of adaptation. But it is more simple than the genetic algorithm rules. PSO searches for the global optimum by following the optimal value of the current search, which have not the genetic algorithm "Crossover" and "Mutation" operation.

For the PSO algorithm, all particles, assuming number n, are passed through the speed of $v_i=(v_{i1},v_{i2}, \cdots, v_{id})$ to update its spatial position $x_i=(x_{i1},x_{i2}, \cdots, x_{id})$, and the optimal solution of the target optimization problem is searched in the d dimension space by the fitness F. Each iteration search will produce the ith particle's individual optimal solution $p_i=(p_{i1},p_{i2}, \cdots, p_{id})$ and the whole particle swarm optimal solution current $p_g=(p_{g1},p_{g2}, \cdots, p_{gd})$. Particles update speed and position according to the following formula.

$$v_{id}^{t+1} = w v_{id}^{t} + c_1 r_1 (p_{id}^{t} - x_{id}^{t}) + c_2 r_2 (p_{gd}^{t} - x_{id}^{t}) \tag{3}$$

$$x_{id}^{t+1} = x_{id}^{t} + v_{id}^{t+1} \tag{4}$$

In the equation, the ω is inertial factor; v_{id}^{t} and x_{id}^{t} are the d dimension of the velocity vector and the position vector of the particle i at t moment respectively; c_1 and c_2 are learning factors; r_1 and r_2 are a random number in the range of 0-1. In the equation (3), the first part is the momentum part, so that it is based on the inertia of its own speed; The second part reflects the thinking and the ability of the particles; The third part represents the mutual cooperation and information sharing among particles. After the termination of the model, the global optimal position vector pg is used as the optimal solution. Inertial factor ω adjust itself size by the following formula in the process of iteration.

$$w = w_{end} + (w_{start} - w_{end})(iter_{\max} - iter) / iter_{\max} \tag{5}$$

In the equation, $iter$ is the current algebra; $iter_{\max}$ is the maximum number of iterations; w_{start} and w_{end} are 0.9 and 0.3, respectively, the setting values of the inertia factor for the iterative start and end time.

C. Particle swarm optimization BP neural network model

As a new evolutionary algorithm, the PSO algorithm has the characteristics of fast convergence speed, high robustness and strong global search ability, which is not need to use the feature information of itself, such as gradient. With the combination of PSO and neural network, PSO algorithm is used to optimize the connection weights of neural network. It can overcome the shortcomings of BP neural network. PSO-BP can not only improve the generalization ability of neural network, but also improve the convergence speed and learning ability of neural network.

1) Structure of particles and population: The selection of training samples is used as the particle population, and the mapping between the neural network connection weights and the PSO particle dimension is established. The dimension of each particle corresponds to a connection weights in the network. Neural network input layer, hidden layer and output layer neuron number were I, H and O, the PSO particles of the dimension of space is $d = I \times H + H \times O + H + O$.

2) Fitness Function: The mean-square error of the output of the network is used as the fitness of the PSO algorithm, may be computed as shown in the following equation:

$$F = MSE = \frac{1}{2n} \sum_{p=1}^{n} \sum_{k=0}^{c} (Y_{k,p}(X_p) - t_{k,p}) \tag{6}$$

In the (5) equation, $Y_{k,p}(X_p)$ is the actual output of neural network; $t_{k,p}$ is the expected output of the training sample p at the k output.

3) Design of BP network model based on PSO algorithm: The weight optimization of BP neural network based on PSO algorithm specific steps are as follows.

Step 1: The initialization of neural network and particle swarm. The input, output and the number of hidden layer neurons, learning function and training function are designed according to the sample data. According to the size of the particle swarm, a certain number of particles are randomly generated according to the individual structure, which is the different individuals represent one groups of different weights of neural networks. Particles of *pi* and *pg* will be initialization at the same time.

Step 2: The training of neural network and the evaluation of the particles. The component of each individual in the particle swarm is mapped into the weights of the network, which constitute one neural network. Training sample will be put into the each one individual corresponding neural network for training. The optimization process of network weights is a process of iteration. In order to ensure that the training of neural network has a strong ability of generalization, in the training process of the network, that is often divided into 2 parts for given the sample space. One part as the training sample is called the training set. The other part as the test sample is referred to as the test set. In order to ensure the training sets are not the same, in the optimization process of the weight value, that it should be classified to the given sample set in the every training.

Step 3: PSO-BP model calculation. To evaluate all the individuals in the particle swarm. To find the best individual is used to determine whether need to update the particles A and B. After that, according to the PSO-BP model, the flight speed of each individual's different components is updated and a new individual particle is generated.

Step 4: Termination condition of the algorithm. When the objective function value (i.e. mean square error) is less than the given ε or the number of iterations to reach the design maximum algebra that the algorithm will be terminated.

IV. Based on PSO-BP model for eye-GAZE tracking

A. Selection of PSO-BP input feature vector

Feature vector selection is the key to the whole system. The feature vector of BP neural network needs to have some differences, otherwise the forecast output of neural network is not accurate. According to the method of Baluja and Pom-erleau, the image is normalized to size of 32×16 in the first time. Then the entire of all pixels of the size of 32×16 eye image is used as a feature vector of a size of 512. The experiment proved that this method is not ideal. Therefore, in this system, the position information of the pupil center and the eye corner of the two eye images are used as the feature vector of the neural network [18]. This feature vector has good difference. It can greatly reduce the dimension of the feature vector and improve the speed of the system. The position information of the pupil center and the eye corner is composed of a set of 16 feature vectors.

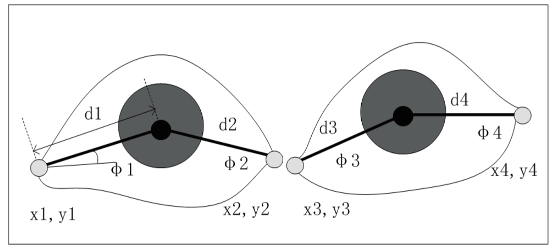

Fig. 3. Geometrical characteristics of the eye diagram

As shown in Fig. 3, the 4 values are composed of the angle between the pupil center and the angle of the eye. Other value consists of the corner of the eye and the center of the iris.

B. Training Neural Network

Using the PSO-BP model, by back-propagation network to continuously adjust the weights and thresholds, and use the PSO algorithm to optimize the weights, so the network and the minimum sum of squared errors [19]. In this system, two separate neural network bp_x and bp_y are created to be used for calculating the coordinates of the X and Y directions, respectively. The method of training is used to choose the method of back-propagation. bp_x uses three layers of structure, where the input layer contains 16 nodes, the hidden layer contains 12 nodes and the output layer contains 5 nodes. The structure of the bp_y input layer and hidden layer is the same as the bp_x, and the output layer contains 3 nodes. The X direction of PSO particles dimension is 181, and the Y direction is 163. The detailed structure of the two networks is shown in Fig. 4.

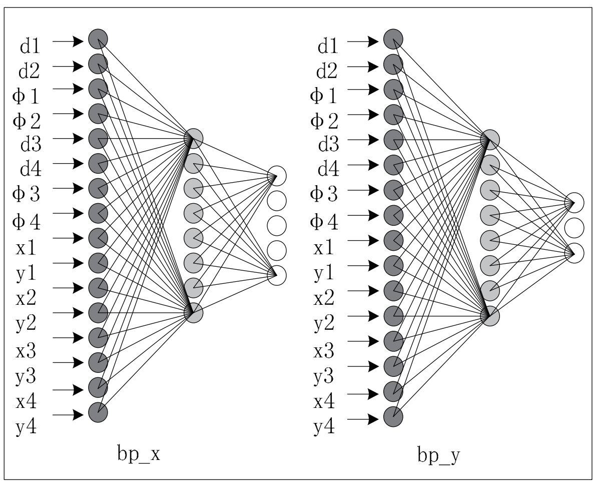

Fig. 4. Training structure diagram of bp_x and bp_y

In this paper, the computer screen is divided into 15 regions in the form of 5x3, and each region is selected one point for testing. When the eye is fixed on a certain area point of the screen, the data of the eye characteristics is acquired, and the data of 20 samples are acquired, which is obtained by the 15×20 group sample data. Finally, the eye characteristic parameters data are trained by PSO-BP model. After the training, the neural network data is stored in the file, and the parameters of the neural network can be called directly when start to confirmatory experiment.

In order to compare the merits of the PSO-BP model and the standard BP neural network, the training samples are trained by PSO-BP and BP neural network, respectively. And the BP model iteration number is taken 400 times, training target is 0.001, hidden layer is 8 layers, and learning rate is 0.1.

V. Experimental Verification

This experiment is carried out under normal illumination condition, face distance computer screen 50cm or so, 2 million pixels industrial camera, the computer screen pixel size is 1280×1024. The 50 experiments were performed on the 15 regional training points in the case of the head remained unchanged or slightly changed, and the mapping points in the corresponding range that is identified successfully. Each time the training time is about 150ms and the recognition time is about 220ms.And the recognition rate of the BP model and the PSO-BP model are compared, because the PSO is used to optimize the connection weights and threshold values of BP neural network. This improved method has high global optimization capability, high accuracy, fast convergence, etc. So, the PSO-BP method has fluctuations of the accuracy, as shown in Fig. 5.

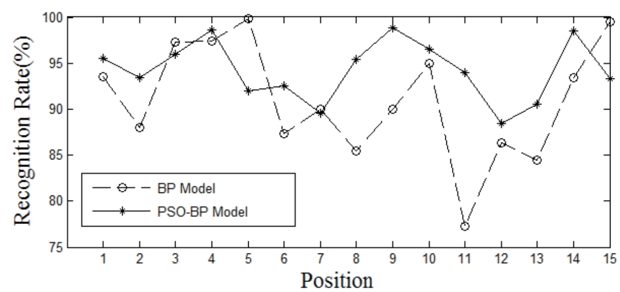

Fig. 5. BP and PSO-BP recognition rate comparison chart

To clear the results of the experiment, the results of the 15 groups of data in the 50 groups were selected for mapping. Among them, the eyes of the real eye-gaze reference point are fixed. The eye-gaze tracking is shown in Fig. 6.

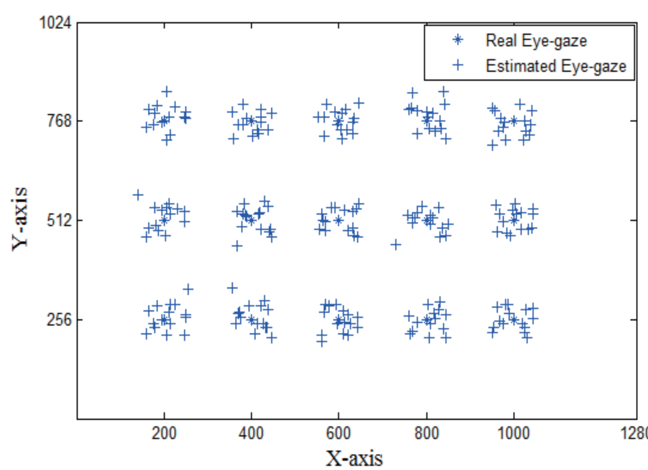

Fig. 6. Eye-gaze tracking mapping

Experimental results show that the system has good effect on the tracking of the eye gaze. Compared with BP neural network, PSO-BP has higher accuracy and robustness. Although the average recognition rate of BP neural network has reached 91.1%. But the volatility is large, and the lack of robustness. The average recognition rate of PSO-BP neural network is 94.9% and the recognition rate of the 15 regions is stable and the PSO-BP system has strong robustness, and can achieve the expected target.

VI. CONCLUSION

In this dissertation, the eye-gaze tracking method based on the particle swarm optimized BP neural network is presented, and the particle swarm optimization is used to optimize the connection weights and threshold values of BP neural network. This improved method is of high global optimization capability, high accuracy, fast convergence, etc., free of the defects of the conventional BP network, such as slow convergence, locally optimal solution, and so on, and reaches the optimal robust and accurate eye feature detection algorithm. The PSO-BP model is built and the algorithm is verified, and the verification results show that the model is featured by simple algorithm, high robustness and high recognition rate. This system requires only an ordinary camera, and thus cuts the restrictive requirements and cost for the hardware and enhances the system practicability. The following study will focus on the eye-gaze tracking method under the condition of free head movement within a larger scope, in order to realize real high accuracy omnibearing eye-gaze detection and use the eye-gaze tracking system in more fields.

REFERENCES

[1] Y. Zhang, R.C. Zhao, X.B Zhao et al, "Eye-gaze tracking technology of 2D and 3D methods are reviewed," *Signal Processing*, 2007, 23(4A): pp. 218-223.

[2] X.B Zhao, X.C Zou, "The principle and realization of the key technology of eye-gaze tracking," *Measurement and Control Technology*.2007, 26 (S): pp. 327-329.

[3] J.N. Chi, C. Zhang, T. Hu et al, "Stady on eye characteristic detection and gaze direction calculation in gaze tracking system," *Control and Decision Making*.2009, 24(9): pp. 1345-1350.

[4] C. Zhang, J.N. Chi, Z.H. Zhang et al, "A new eye-gaze tracking method based on Pupil-Cornea Reflection technique," *Chinese Journal of Computers*.2010, 33(7): pp. 1272-1285.

[5] Y. Ebisawa, K. Fukumoto,"Head-Free, Remote Eye- Gaze Detection System Based on Pupil-Corneal Reflection Method With Easy Calibration Using Two Stereo-Calibrated Video Cameras," *IEEE Transactions on Biomedical Engineer*.2013,60(10): pp. 2952-2960.

[6] R.A. Liu, S.J. Jin, W. Song et al, "Eye-Gaze Tracking Using One Camera," *Computer Application*, 2006, 26(9): pp. 2101-2104.

[7] Z.Y. Lu, X.C. Zhao, "Eye-gaze tracking method for free movement of head," *Journal of Nanjing University of Aeronautics and Astronautics*, 2010, 42(4): pp. 435-439.

[8] B.R. Hong, F.H. Jin, Q. Guo, "Hopfield neural network based on ant system," *Journal of Harbin Institute of Technology*，2004，11(3): pp. 267-269.

[9] M. Liu, H. Li, "Face detection and recognition based on AdaBoost classifier and feature face," *Journal of zhangzhou teachers college*：Natural Science Edition, 2010, 2:43-17.

[10] F. Timm, E. Barth. "Accurate Eye Centre Localisation by Means of Gradients," *Institute for Neuro and Bioinformatics*. 2012, 30(8): pp. 164-169.

[11] J. Zhang,*Opencv2 computer vision programming manual*. Beijing: Thomson Learning Press, 2013.

[12] F. Gao, X.M. Zhang, "Precise positioning method of eye corners in eye gaze tracking system," *Computer Engineering*, 2007, 33 (18) : pp. 199-204.

[13] P.F. Yan, C.S. Zhang, *Artificial neural networks and evolutionary computing*. Beijing: Tsinghua UniversityPress, 2006.

[14] L.X. Yao, L.F. Song, Q.Y. Li, et al, "Power systemshort-term load forecasting based on fuzzy clustering analysisand BP neural network," *Power System Technology*, 2005,29(1): pp. 20–23.

[15] J. KENNEDY, R. EBERHART, "Particle swarm optimi-zation," *IEEE International Conference on Neural Networks*, 1995, 27(4): pp. 1942-1948.

[16] J.X. Zhang, "Application of particle-swarm-optimiza-tion-trained artificial neural network in high speed mil-ling force modeling," *Computer Integrated Manufac-turing Systems*, 2008, 14(9): pp. 1710-1714.

[17] C.Z. Feng, M.W. Shen, "Application of gaze tracking in human computer interaction," *Journal of Zhejiang University*, 2002, 29 (2): pp. 225-232.

[18] L. Sesma, A. Villanueva, R. Cabeza, "Evaluation of Pup-il Center-Eye Corner Vector for Gaze Estimation Using a Webca-m," *Sant-aBarbara*,CA.2012,28(3): pp. 217-220.

[19] Y. Hou, "Research on BP neural network and its application based on steepest descent algorithm," *Modern Business Trade Industry*, 2011, 8: pp. 245-248.