Individualized Detection for Distant Learning Behaviors

Kuo-An Hwang, Hsu-Liang Tseng , Chia-Hao Yang

Department of Computer Science and Information Engineering and Department of Information and Communication Engineering, Chao Yang University of Technology
Graduate student, Department of Computer Science and Information Engineering, Chaoyang University of Technology
Ph.D. of Graduate Institute of Informatics, Chaoyang University of Technology

kahwang@cyut.edu.tw

s9727638@cyut.edu.tw

s9433901@cyut.edu.tw

168 Gifeng E. Rd., Wufeng
Taichung County, Taiwan, R.O.C.

Abstract - **Distant learning has been extensively studied in the recent years. It is not only an educational approach that can be employed in remote areas, but also can serve as a supplementary tool when students and teachers are distantly located and unable to carry out face to face learning. However, due to insufficient self-control, students, when learning at home, are likely to be influenced by external objects, and thus, are not concentrated on learning. This study utilized the detection algorithm of Webcam and computer, and installed regular webcam on the desks of students for facial detection. The facial features of the students were first captured by the principle component analysis (PCA), and analyzed by artificial neural network. The learning status of the students was then determined based on the real-time facial features, so that teachers and parents can know the students' learning status, and help them to overcome the distraction, and reach the goal of individualized learning.**

Keywords— **Image processing, Distance learning, Artificial neural network**

1. INTRODUCTION

Distant learning has been extensively studied in the recent years. It is not only an educational approach that can be employed in remote areas, but also can serve as a supplementary tool when students and teachers are distantly located and unable to carry out face to face learning. Moreover, for students who are unable to attend school due to infectious diseases, they could listen to lecture in front of their computers. However, the shortcoming of distant learning is that students are likely to be distracted, feel sleepy, drowse, or be absent-minded when learning at home. It indicates that in-home learning may be affected by many weary conditions, thus, the students' learning behaviors require close observation and concern. Computer is the most useful tool to monitor learning behaviors.

In biometrics, human facial recognition plays an important role. The primary requirement for most facial recognition systems is locating the facial areas and collecting data on facial features for recognition. Yang used the changes of facial features, such as eyes and mouths, along with a Bayesian network, to evaluate the students' mental status [17]. However, due to different detection factors, such as environment, shielding, posture changes, low lightness, and plane revolution, which may due to the low accuracy of sensors, environmental conditions, or the characteristics of the subjects [6], biometric systems often require learning samples based on multiple subjects and features, in order to maintain the applicability.

Monitoring students' learning behaviors is a personalized work; it is to monitor a specific subject over a long period of time. If the individual differences of every face can be considered during analysis, a more efficient result can be achieved. This study aimed to design a system that focuses on the facial recognition and comparison of children. The system processes multiple images of the children when they carry out distant learning at home to recognize their levels of concentration. Based on the facial features collected from students through the artificial neural network, the system establishes individual parameter modules for the students.

2. LITERATURE REVIEW

The main physiological features of humans when experiencing fatigue include: eyes-closing, lack of concentration, yawning, head turn and wrinkles-formation on the top of the nose. As fatigue may cause traffic accidents [9], fatigue detection has been employed on vehicles. When the driver's concentration decreases, the system sends out a warning [10],[4],[15]. The possible indicators of the fatigue behaviors of drivers include: 1) percentage of eyes-closing, 2) duration of eye-closing, 3) frequency of winking, 4) frequency of nodding, 5) relevant positions of faces, 6) dazing [5]. Fatigue detection is based on image recognition.

These technologies can be applied to distant learning to detect learners' learning status and bad learning behaviors, such as dazing, sleeping, and leaving. The interaction and observation records of distant learning can provide teachers the basis for evaluation, and achieve the instructional objectives in affective domain. The students' frequency of interactive discussion can serve as the basis for participatory evaluation, in order to assess their participation and interaction [3]. According to the attendance and the responses in the affective domain, the final concentration level can be calculated based on the scores of the two stages [12].

Principal Component Analysis (PCA) is an analysis that has been widely applied. Its principle is to transform complicated multidimensional data into a data set with fewer dimensions so as to facilitate analysis [18]. Pentland used PCA for facial recognition. Their system was to capture local features, and the positions of the features, such as eyes and mouth, were manually marked, for recognition by PCA. PCA has also been applied to pre-processing [14]. Yang and Sun used PCA to capture the data of facial features, and then used the data for training to achieve the function of facial recognition [2],[7].

There are many types of artificial neural networks, among which backpropagation (BP) neural network is the most commonly used one. Its structure contains the input layer, hidden layer and output layer. It is trained with input vector and the corresponding target vector, until the network approaches a function [16]. Han proposed a morphology-based system to perform the division of eye similarity, used BP neural network to complete the facial validation [11]. Sun used PCA+LAC to capture the facial features, and the facial classification was carried out by trained BP neural network [7].

3. SYSTEM STRUCTURE AND FLOW

This system is designed for detection and determination of facial images. A webcam (see Figure 1) was installed at the student end to capture his facial features. The continuous facial images were processed and analyzed by artificial neural network, in order to determine whether the students showed lack of concentration, fatigue, or doze, thus give out a timely warning and produce records. Teachers and parents can check the record to find out about the students' concentration status, and handle the situation accordingly.

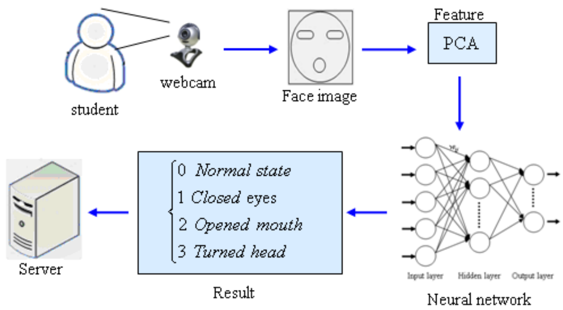

Fig. 1 Systematic structure of students' self-learning

The system is designed to extract one facial image from the webcam per second. The image output rate can be adjusted based on the system processing speed. Higher output rate can produce more accurate result on the detection of the changes in subjects' facial features. After capturing the images, the face features are selected by PCA, and analyzed with trained artificial neural network. During the acquisition of facial features, it is important to reduce the overall computational complexity, such as by reducing the image resolution, the computational complexity during feature recognition, or the detection times with good feature-tracing algorithm. We employed PCA for images processing, to reduce the computational complexity, increase the image output rate, and minimize the errors in status determination.

3.1. Finding the Facial Area

In this step, YcbCr is used for complexion partition to find the facial area. In YCbCr, Y is brightness, Cb is blue minus brightness (B-Y), and Cr is red minus brightness (R-Y). The advantage of YCbCr is its little impact on the elimination of colors, as human eyes have higher sensitivity to brightness than to colors [1].

Complexion can be used to screen various color spaces. The equation proposed by Lin is used for computation [1]. The computed Y, Cb and Cr can realize the input images of every pixel in the skin test by the detector, based on the pixel complexion detection, and further compute the Cb and Cr module values. If the pixel of Cb and Cr are within a certain range, this area is regarded as a complexion area, by which the facial area can be found.

3.2. Extracting Facial Features

M input facial images are taken as training samples in our study. Each sample has pixel matrix of $N \times N$, and is arranged into Γ vectors of $N^2 \times 1$ by serial connection, as shown in Figure 2.

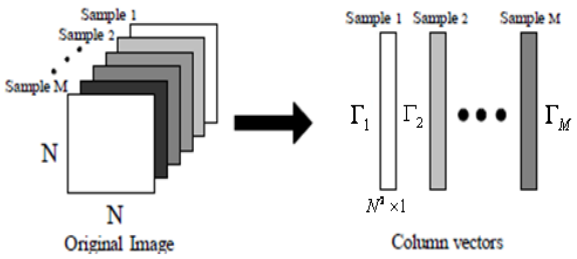

Fig. 2 Transform facial images into one-dimensional vectors

Each facial image corresponds to a vector Γ. In Eq. (1), the average vector Ψ is the $N^2 \times 1$ vector of each image connected in series. The values are summed and divided by the number of image samples. Therefore, this average vector is also called the average image [13].

$$\Psi = \frac{1}{M}\sum_{i=1}^{M}\Gamma_i \tag{1}$$

Average vector Ψ is the mean face, which indicates the common part of these images. The common parts are deleted in M pieces of facial images to highlight the different parts, and obtain the difference vector of each face. This image is called the difference image.

$$\Phi_i = \Gamma_i - \Psi \tag{2}$$

M difference vectors Φ are arranged into A matrix of $N^2 \times M$ dimension, and the internal product of A matrix and its transposed matrix is used to obtain the covariance matrix (C) as shown in Eq.(4):

$$A = [\Phi_1, \Phi_2, \ldots, \Phi_M] \tag{3}$$

$$C = \frac{1}{M}\sum_{i=1}^{M}\Phi_i\Phi_i^T = AA^T \tag{4}$$

Since the computed dimension of C matrix by Eq.(4) is $N^2 \times N^2$, the evaluation of eigenvectors and eigenvalues of C matrix is very difficult and complicated. The method proposed by Turk to compute L matrix [8] is applied to simplify the computational complexity. The dimension of L matrix is $M \times M$, and eigenvector v_i (i = 1,2,…, M) can be obtained from L matrix. As in Eq.(5), v is multiplied by A to obtain the main component υ.

$$L = A^T A = \frac{1}{M}\sum_{i=1}^{M}\Gamma_i^T\Gamma_i \tag{5}$$

$$\upsilon_i = Av_i \tag{6}$$

υ_i is the training and test sample predicted by PCA. The sample vectors of each facial sample is arranged in descending order according to the computed eigenvalues, and are combined with the corresponding eigenvectors into eigenspaces, in order to find their weighing vector Ω in eigenspaces.

$$\omega_k = \upsilon_k^T \cdot \Phi \tag{7}$$

$$\Omega = [\omega_1, \omega_2, \ldots, \omega_M] \tag{8}$$

Using this method, the computational time is greatly reduced.

3.3 Detection of Mental Status

A BP network, as shown in Fig. 3, is constructed for the detection of metal status. The connection between two layers has a weight value of W_{ij} . In different cases, the weight value W_{ij} makes automatic adjustment to minimize the difference.

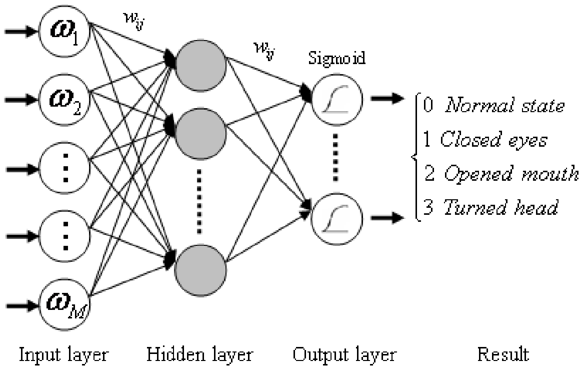

Fig. 3 Back-Propagation neural networks

As people's faces differ, the features vary as well. To allow the modules of artificial nerve to train different parameter values according to different users so as to form individualized artificial neural modules, the subsequent detection can use individually formed modules to achieve differential recognition. By using this individualized parameter for comparison analysis could achieve more accurate result than that without individualized modules.

3.3.1 Artificial Neural Network Training

When new users setup the system, the system asks them to take samples required for training data.

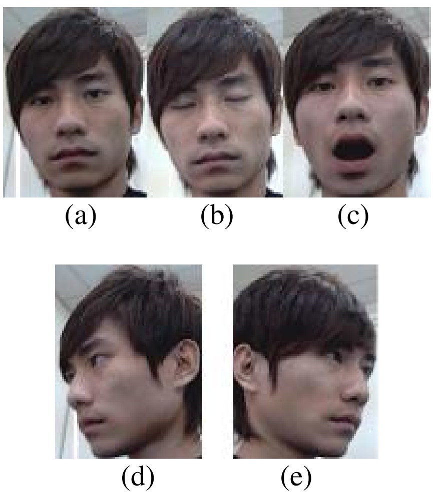

Fig. 4 Trained facial expressions (a)normal state, (b) closed eyes, (c) opened mouth, (d), (e)turned head.

The required samples include normal state, closed eyes, opened mouth, and turned head, as shown in Fig. 4. The users are required to follow the sample collection procedures to establish their individual sample data.

After training, the different parameters of individual faces are recorded in the individual artificial neural modules. Each user has a different set of parameters, thus, the comparison of individualized data can achieve higher accuracy, individualized analysis, and recognition of each facial status.

3.3.2 Individualized Analysis

The processing flow of individualized analysis is shown in Figure 5.

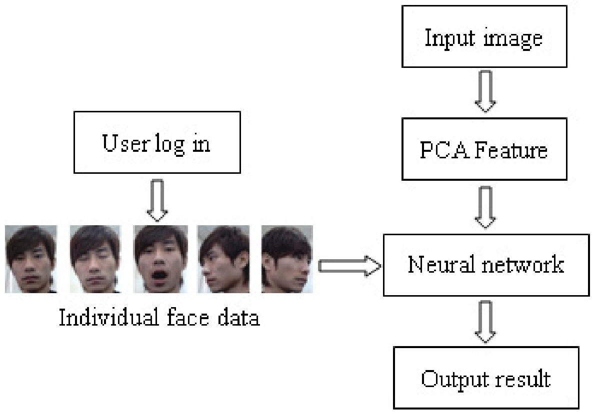

Fig. 5 Procedures of individualized analysis

Part 1: Identify the user to acquire his individual artificial neural parameter database.

Part 2: Process the images that are inputted by Webcam to process the extracted features, and use the artificial neural parameter database to analyze the facial expressions to obtain the final output results.

3.3.3 Detection of Mental Status

The processing flow of mental status is shown in Figure 3. The values of input nodes in the input layer are from the facial features processed by PCA, and the linear transfer function $f(x)=x$. Simple single layer is used to process the hidden layer. The input processing units have interactive effect. The node number of processing units cannot be determined based on standard, and the optimal number is determined by tests and non-linear transfer function. The output variables of the output layer employ non-linear transfer function. This text applies the most commonly used non-linear transfer function in BP artificial neural networks - Sigmoid function:

$$f(x) = \frac{1}{1+e^{-x}} \tag{9}$$

Based on the real-time images inputted by webcam, the facial location can be found based on complexion. PCA is then used to extract the individual features, which are compared with those in the corresponding facial feature data through artificial neural network, in order to produce the final result. The output results include: 0: normal state, 1: closed eyes, 2: opened mouth, and 3: turned head.

3.3.4 Analysis of Recognition Result

The continuous images captured by webcam at a rate of 1 image per second are processed. When the time of eyes-closing is longer than the threshold, it is regarded as a feature of eyes-closing; when the times of mouth-opened widely is more than the threshold, it is regarded as a feature of yawning or talking; if the are of facial search is larger than threshold, it is determined that the subject has left the seat. Finally, these data are analyzed to determine the mental status of the subjects, and the result is shown on the teachers' computers.

4. CONCLUSIONS

With the aim to prevent the lack of concentration of students during self-learning at home, this study designs a self-learning monitoring system. The system uses webcam to capture real-time images and employs image processing technology to identify the facial locations of the students. It then uses artificial neural network to determine the concentration level of the students. To realize individualized processing, individualized artificial neural modules of students are developed in the artificial neural network, in order to determine the changes of facial features of the students, analyze their mental status. With the monitorial report, it allows teachers or parents to know the mental status of their students from remote sites.

REFERENCES

[1] Chiunhsiun Lin," Face detection in complicated backgrounds and different illumination conditions by using YCbCr color space and neural network", *Pattern Recognition Letters*, Volume 28, Issue 16, Pages 2190-2200, 2007.

[2] Jie Yang , Xufeng Ling, Yitan Zhu, Zhonglong Zheng ," A face detection and recognition system in color image series," *Mathematics and Computers in Simulation*, Volume 77, Issues 5-6, Pages 531-539, 2008.

[3] K.-A. Hwang and C.-H. Yang, "A Synchronous Distance Discussion Procedure with Reinforcement Mechanism: Designed for Elementary School Students to Achieve the Attending and Responding Stages of the Affective Domain Teaching Goals within a Class Period", *Computers & Education*, vol. 51, pp. 1538-1552, 2008.

[4] P. Smith, M. Shah and N. d. V. Lobo, "Determining Driver Visual Attention With One Camera," *IEEE Transactions on Intelligent Transportation Systems*, vol. 4, pp. 205-218, 2004.

[5] Q. Ji , Z. Zhu and P. Lan , "Real-Time Nonintrusive Monitoring and Prediction of Driver Fatigue," *IEEE Transactions on Vehicular Technology*,, vol.53, pp.1053-1059, 2004.

[6] Raja Tanveer Iqbal, Costin Barbu, Fred Petry," Fuzzy component based object detection", *International Journal of Approximate Reasoning*, Volume 45, Issue 3, Pages 546-563, 2007.

[7] Te-Hsiu Sun, Fang-Chih Tien," Using backpropagation neural network for face recognition with 2D + 3D hybrid information," *Expert Systems with Applications*, Volume 35, Issues 1-2, Pages 361-372, 2008.

[8] Turk, M. and A. Pentland. "Eigenfaces for recognition," *Journal of Cognitive Neuroscience*, 3(1): p. 71-86, 1991.

[9] H. Gu and Q. Ji, "An Automated Face Reader for Fatigue Detection, " *IEEE International Conference on Automatic Face and Gesture Recognition*, pp.111-116, 2004.

[10] H. Ueno, M. Kaneda and M. Tsukino, "Development of Drowsiness Detection System," *IEEE International Conference on Vehicle Navigation and Information Systems*, pp. 15-20, 1994.

[11] Han, C.C., Mark Liao, H.Y., Yu, G.J., Chen, L.H.,"Fast face detection via morphology-based pre-processing,"*Pattern Recognition* , 33 (10), 1701–1712, 2000.

[12] Kuo-An Hwang, Chia-Hao Yang and Chuan-Ren Wang," Assessment for Attending and Responding stages of Affective Teaching Goal in Distance Learning Based on Fuzzy

Fusion," *Information Technology and Applications, 2009.* International Forum on Volume 3, Page(s):275 – 277, 2009.

[13] Lih-Heng Chan; Salleh, S.-H.; Chee-Ming Ting," PCA, LDA and neural network for face identification," *Industrial Electronics and Applications*, 4th IEEE Conference on 25-27, Page(s):1256 – 1259, 2009.

[14] Pentland, A. Moghaddam, B. Starner, T. ," View-Based and Modular Eigenspaces for Face Recognition," *Computer Vision and Pattern Recognition*, Proceedings CVPR '94., IEEE Computer Society Conference on 1994.

[15] T. Hamada, T. Ito, K. Adachi, T. Nakano, and S. Yamamoto, "Detecting Method for Drivers' Drowsiness Applicable to Individual Features," *IEEE Proceedings on Intelligent Transportation Systems*, vol. 2, pp. 1405-1410, 2002.

[16] Kuo-Yuan Chen," The Automatic Measuring Strategy Design for CMM Based on Image Processing," Thesis for Master of Science Department of Mechanical Engineering Tatung University, 2005.

[17] Ming-Ru Yang," Children Self-Learning States Monitor at Home Using Image Processing" ,Graduate Institute of Networking and Communication Engineering Chaoyang University of Technology, 2008.

[18] Meng-Yi Chen," A Study of Image Fusion in Dual Mode Imaging Systems," A Thesis for the Degree of Master of Science Graduate Institute of Graphic Communications and Publishing Shih Hsin University, 2008.