Children Self-Learning States Monitor at Home Using Image Processing

Kuo-An Hwang
Department of Computer Science and Information Engineering & Graduate Institute of Networking and Communication Engineering, Chaoyang University of Technology
kahwang@cyut.edu.tw

Ming-Ru Yang
Graduate Institute of Networking and Communication Engineering, Chaoyang University of Technology
s9530620@cyut.edu.tw

Chia-Hao Yang
Graduate Institute of Informatics, Doctoral program, Chaoyang University of Technology
s9433901@cyut.edu.tw

摘要

由於時代的進步，許多父母都非常注重學童的學習。但是社會的發展，許多父母因為工作的繁忙，而無法陪伴學童旁學習。當學童在家自主學習時，由於自制能力的不足，很容易受到其他外物的影響，產生學習不認真或不專心，因此本研究利用 IPcam 和電腦的偵測運算，在學童書桌上裝置簡單的 IPcam，透過學童的臉部偵測，再利用影像處理的方式擷取學童的眼睛、嘴巴等臉部特徵，用貝氏網路判斷，根據這些臉部特徵產生的數據判斷學童學習狀況，自動偵測學童學習狀態，讓父母親可以知道學童自主學習的情況。

關鍵詞：影像偵測、學童自主學習、貝氏網路

Abstract

If the children are Self-Learning at home, they are easily distracted by outside influences, due to a lack of Self-Control and become inattentive or indolent. Usually, parents are busy and unable to continually monitor their children in details all time.

For what research utilizes the IPcam and detection algorithms on the computer to do monitor children Self-Learning States. The IPcam is installed to the desk of the child at home. The face image of the child is then captured. The eyes, lips, and other facial traits of the child are found by image processing techniques. The states of Self-Learning of child are determined based on Bayesian Networks Model with the data generated from the facial traits. These facial traits automatic detection of the children can enable parents to easily determine the learning progress of their children.

Keywords: Image Detection, Children Self-Learning, Bayesian Networks Model.

1. Introduction

It is very important that parents have to pay a great attention to the learning process of their children to guide them good learning habits when they are young.[12] However, economic pressures have forced many parents to maintain a busy schedule and they are not able to accompany the children in their studies. When a young child is self-learning at home, he is easily distracted by outside influences, and due to a lack of Self-Control, resulting in inattention, or even tiring situations.

In order to monitor the learning conditions of a child, various companies have invented monitoring devices suitable for home usage. A camera can be placed in the room of the children. The captured images from the camera are transmitted to the parents' computer through the home network to monitor the status of the children in the room. However, this type of manual monitoring is not practical as both of the parents are busy and unable to continually monitor the children all time. Moreover, some behaviors of the children are relatively subtle and must be immediately discovered by the parents for adjustments. Otherwise, precious learning periods will be wasted and the self-learning habit of the children will be lost.

There are many researches those aim to detect the state of the automobile driver while driving. When the human body is fatigued, the frequency of eyelid blinking will be increased, the increased frequency of yawning due to increased fatigue, the state of fatigue will be displayed on the face, etc. However, Self-Learning children at home will not only express fatigue for inattention behaviors, but also include other habits such as looking all around. Other behaviors like leaving the seat to do other things and eye gaze are currently not presented in the automobile fatigue detection technologies.

In our research, we will propose a system, which can analyze facial trait from the captured images to determine if the children are fatigued, head turned, nodding off, or even not at their seat. When the inattention of the children occurs, appropriate audio or visual warning can be used to notify the parents. The parents will be freed from paying full attention to the monitoring device, and can be fully notified from the computer notifications. This will alleviate the pressure on the parents as well as enable the self-learning habits of the children.

2. Fatigue Detection

When a person is not well rested or sleep deprived, his eyes will become shut, concentration will reduce, heads will unconsciously nod, and vertical muscle wrinkles above the nose. Most common traffic accidents [2] also result from the fatigue of the driver. In prior researches, physical traits of the driver from fatigue will trigger the on board system to alert the driver [6][8][11] with auditory warnings. Possible signs of driver fatigue include: 1. Level of eye closure; 2. Duration of eyelid closuring; 3. Frequency of eyelid closuring; 4. Frequency of head nodding; 5. Frequency of yawning; 6. Face away from the driving position; 7. Eye gazing [7]. Many researches reported that when the heart rate, pulse rate, brainwave reflex of the driver speeds up, the warning system will emit the auditory warning [7].

Various researches also presented a method of least deviation error through Bayesian model for probability and statistics, where the frequency of eyes closuring or mouth stretch can be analyzed. The facial fatigue states can be applied towards the image processing detection [10].

3. Children Self-Learning System and Infrastructure

For monitoring the learning process of the children, the system is proposed and shown in Fig. 1. In which, an IPcam is installed to the desk of the child. It captures and transmits images to the parents' computer through the home network. The PC in the parents’ room is used to analyze images to determine the state of the child’s Self-Learning through the image processing technique and Bayesian Networks Model. The analyzed fatigue results can be relayed to the parents for fulltime monitoring.

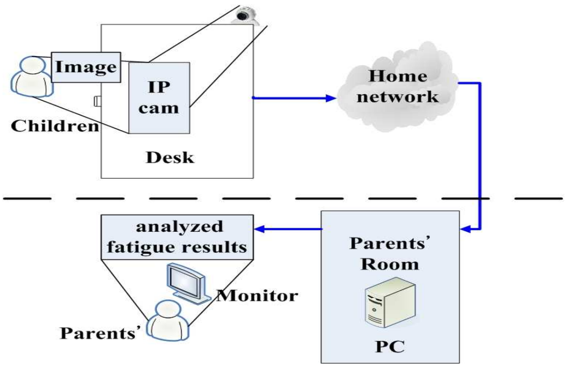

Fig. 1 Children Self-Learning System structure

3.1 Bayesian Networks Model

The learning habit of child at the desk is defined as below three states: Inattention, Drowsiness, and Falling Asleep. Inattention is the state when the child leaves the desk or the child gazes at other objects, then the studies are interrupted. Drowsiness is defined by three behaviors, the head of the child may lean forward due to nodding off, the head may be shifting from side-to-side from stretching, or the eyelids may narrow or shut from yawning. When the children are in a near sleep state, the eyelids may be shut and the head may lean to a direction, Therefore, we defined the possible factors for head turning and eyelids closure conditions. Fatigue states will differ from the children for various probabilities, wherein the entire network model [10] is as shown in Fig. 2, and Table. 1 is the code lookup table for various behaviors.

Human beings are easily affected by personal issues or physical conditions, thereby resulting in the gazing state. The use of image processing to detect gazing is to eliminate the possibility of a dummy in place of the child. When performing image detection of the human face, the skin color is used to detect the eyes and lips on the face. If there is a period of inactivity of the eyes or lips, the detected face may possibly be a dummy or the child is eye gaze.

3.2 Creating the Probability Table

Before using the Bayesian Networks Model to determine the physical fatigue of the child, we need to define parameters for the network. Usually, these conditional probability data must be obtained from analysis and statistics of large amounts of experimental data. We obtain the optimal data [10] from defined probabilities in prior references or with continuous adjustments.

During fatigue recognition, linking between different nodes will cause different results, where the probabilities of these nodes will be the key to arrive at the different results. From reference [5] it can be seen that in addition to the fatigue characteristics from blinking or yawning, the surrounding influences such as the temperature in the room or the light will result in different physical deviations. The probability of these surrounding events will also have a defined distribution. Qiang Ji [7] represented the possible surrounding events with a specific probability. From physical fatigue with multiple deciding factors, such as when the facial wrinkles and lips opening may result in a high probability of yawning, we can obtain an even more accurate result from linking between different nodes.

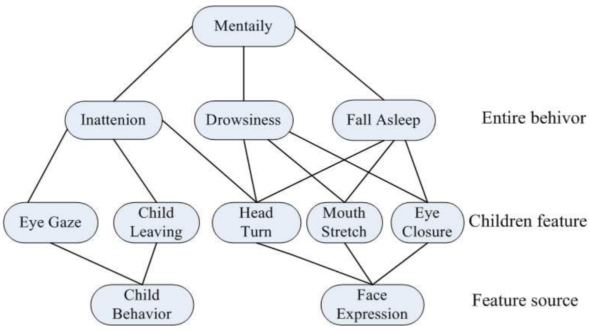

Fig. 2 Bayesian Networks Model

Table. 1 Lookup Table for Behavior Codes

Behavior	Code	Behavior	Code
Eye Gaze	G	Mouth Stretch	M
Child leaving	C	Eye Closure	E
Head Turn	H		

4. Detection of Facial Traits and Behaviors

The monitored targets include Eye gazing, Children leaving, Head turning, Mouth Stretch, etc. Probability of trait behavior occurrence is computed according to facial traits detections that rely on action detection, skin area detection, eye detection and lips detection. Template images are required for open eye, close eye, open mouth and closed mouth actions, for the purpose of detecting these actions.

4.1 Search for the eyes area

When a facial image of a child is captured by the IPcam, the eye area is searched by the following steps, as shown in Fig.3.

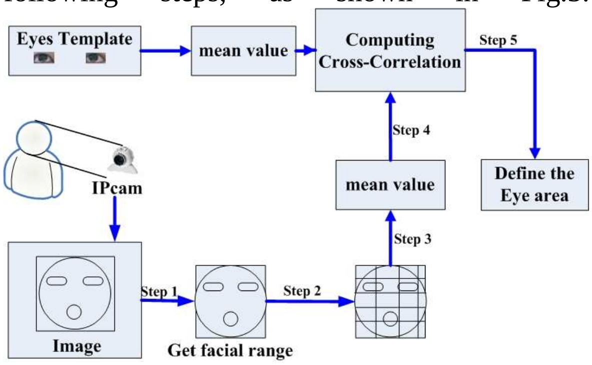

Fig. 3 Eye-area Detection Flowchart

$$S(x,y)=\begin{cases}1 & \text{if}\,(g<Q_+)\,\&\,(g>Q_-)\,\&\,(W>0.004)\\ 0 & \text{otherwise}\end{cases},\quad g=\frac{G}{R+G+B} \tag{1}$$

Where

Q_+: up boundary,
Q_-: down boundary,
g: image green pixel,
W: white noise,
R: red color,
G: green color,
B: blue color
$S(x, y)$: skin gray array

Step 1: Normalize the captured RGB color image, and then compute the skin gray array $S(x, y)$, by Eq. (1)[9]. The facial range is defined by skin region, that is the area with $S(x, y) = 1$.

Step 2: Divide the face image into n blocks, where n is selected according to the template size.

Step 3: Transform RGB color image into the gray values, and compute mean values of all the n blocks.

$$\gamma(u,v)=\frac{\sum_{x,y}[f(x,y)-\bar{f}_{u,p}][t(x-u,y-v)-\bar{t}]}{\left\{\sum_{x,y}[f(x,y)-\bar{f}_{u,p}]^2\sum_{x,y}[t(x-u,y-v)-\bar{t}]^2\right\}^{0.5}} \tag{2}$$

Where

$r(u, v)$: Normalized Cross-Correlation,

$\bar{t}$: The mean of the feature,

t: template, $f(x,y)$: Image pixel value,

$\bar{f}_{u,p}$: The mean of $f(x,y)$ in the region under the template

Step 4: Compute the Normalized Cross-Correlation by using the Normalized Cross-Correlation method[4], as shown in Eq. (2), In which, mean values of the blocks are compared with the eye-template. That is to define the similarity coefficient of the facial image in respect to the eye-template.

Step 5: Select the area that has high Cross-Correlation to be the eye area. In which a threshold is used to make the decision.

If the Cross-Correlation of block has a value lower than the threshold, the area is less likely to be the eye area. If more than two eye areas are positioned, a misjudgment may have been incurred and a resetting of the Template or an adjustment of the threshold must be carried out.

4.2 Searching for the lips area

By further comparing the template of closed lips with the closed lips in the child's facial image captured by the camera, it is possible to verify whether the child is in a closed lips status.

For detecting the lips area, it also utilizes the color characteristics of the skin and the lips. Compared with the skin, lips have a color with a more red bias. Using the previously mentioned complexion area as defined in Eq. (1), with the additional R-G>15 condition, results in the Eq. (3).

$$L(x,y)=\begin{cases}1 & if(g<Q_{-})\&(W>0.001)\&(R-G\geq 15)\\ 0 & otherwise\end{cases},\ g=\frac{G}{R+G+B} \quad (3)$$

Where	*R*: red color,
Q-: down boundary,	*G*: green color,
g: image green pixel,	*B*: blue color,
W: white noise	*L(x, y)*: lip gray array

4.3 Verification of Self-Learning traits

When the child is learning alone in the room, there will be two sources of traits: child's behavior and face expressions. Three different face expressions are identified: 1.Eye Closure; 2.Mouth Stretch; 3.Head Turn. Two types of behavior are identified: 1.Child Leaving; 2.Eye Gaze.

By continuous detection of images captured by the camera, if the eye closuring time is greater than a preset threshold value, the state is determined as having an excessive eye closuring duration.

If the lips open count value is greater than a preset threshold value of a mouth open probability, the state is determined as having a stretching trait.

If the range of a missing eye search is greater than a preset threshold value of the eye search probability that can determine the child leaving.

If the child's eye area stays in a same position more than the time probability threshold value of a preset eye area, the state is determined to be having a trait of the eye gazing while sitting on the seat.

While the child is not turning his head, the positions of the eyes maintain a certain distance with the edge of the face; while he turns his head, the position of the eyes tends to approach the face boundary, and even overlaps with the latter. Therefore computing the distance between the eyes position and the face boundary enables to determining whether the child turned his head.

The verification of the above self-learning traits is imported into a Bayesian Networks Model for deciding whether the child is acting his behaviors out of tiredness.

4.4 Discussions

While detecting the facial images of the child, some external factors may cause incorrect detection of the camera mechanism. For example, if child wears sunglasses while reading, the system may determine it as an eye close. If child wears mask while reading, the system may determine it as an open mouth.

However, according to the Bayesian Networks Model, there are three tiredness factors: 1.Eye closure; 2.Mouth stretch; 3.Head turn. If child has not stretched or turned his head, high probability value in the time of eye-closure is obtained, but the probabilities for the other two items will be substantially low; summing up all the three tiredness factors may not result in a tiredness state. The same applies for the rest of the items.

5. Conclusion

Our research uses IPcam to capture real-time facial images, and combines with image processing technologies, as well as the dedication in using the computer, to automatically detect the states of the child self-learning. With evaluations made with the Bayesian Networks Model, the study also assists parents in controlling learning conditions of their children By remote monitoring of the children's learning state, parents have the ability in controlling learning conditions of their children. They are freed from problems that may arise from self-learning of the children at home.

Moreover, by using the continuous image analysis in our study, low-cost cameras can be used to reduce the equipment cost.

The algorithm for monitoring children's self-leaning states at home is via behavior detection that mostly detects behavior traits. However some of the environmental influences, such as whether the camera is aligned with the facial area, change of light sources, background complexity, or different computer processing speeds, may have somehow different effects on the results. If, in the future, the system can be made capable of automatically adjusting light and brightness of the image, the overall detection mechanism may become more humanly natured.

Reference

[1] Berqasa, L. M. et al., "Real-Time System for Monitoring Driver Vigilance," ***IEEE Transactions on Intelligent Transportation Systems***, vol.7, pp.63-77, 2006.

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

[3] Gan, L., Cui, B., and Wang, W., "Driver Fatigue Detection Based on Eye Tracking," ***IEEE Proceedings on Intelligent Control***, vol.4, pp.21-12, 2006.

[4] Haralick, R. M., and Shapiro, L. G., ***Computer and Robot Vision*** vol. 2nd: Addison-Wesley, 1992.

[5] Horng, W. B., Chen, C. Y., Chang, Y., and Fan, C. H., "Driver Fatigue Detection Based on Eye Traacking and Dynamic Template Matching, "***IEEE International Conference on Networking, Sensing and Control***, vol. 1, pp. 7-12, 2004.

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

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

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

[9] Soriano, M., Martinkauppi, B., Huovinen, S., and Laaksonen, M., "Using the Skin Locus to Cope with Changing Illumination Conditions in Color-Based Face Tracking," ***IEEE Nordic Signal Processing Symposium***, pp. 383-386, 2000.

[10] Tsai, P. X., "Learner Attending Auto-Monitor in Distance Learning Using Image Recognition and Bayesian Networks," ***Graduate Institute of Networking and Communication Engineering of Chao yang University of Technology***, 2006.

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

[12] Wang, Y. C., "Analysis of strategies for a kindergarten teacher in treating issues of learning behaviors of kids," ***Graduate Institute of Department of Early Childhood Education of Chia Yi University***, 2002.