An Adaptive Gaze Tracking System in the Diverse Environment

Chiao-Wen Kao [1], Hui-Hui Chen [2,*], Shyi-Huey Wu [3], Bor-Jiunn Hwang [4], Hung-Sheng Lai [5]

[1] National Central University, Taiwan
[2,4] Ming Chuan University, Taiwan
[3] Vanung University, Taiwan
[5] Fu Jen Catholic University, Taiwan
[1] chiaowenk@gmail.com; [2] huichen@mail.mcu.edu.tw*; [3] shyihuey@mail.vnu.edu.tw; [4] bjhwang@mail.mcu.edu.tw; [5] 083878@mail.fju.edu.tw
* corresponding author

Abstract

For understanding the learning effect, the gaze tracking is a more objective method to collect the data on eye movement and used to analyze the learning behaviors. In order to provide an analysis tool to apply in diverse environments, an adaptive gaze tracking system is proposed in this paper. This system can offer helpful suggestions for educators when improving the instruction. Additionally, the experimental results show that the estimation of gaze points was successful even when the distance and head rotation alter.

Keywords: Learning Behavior; Learning Effect; Adaptive Gaze Tracker

1. Introduction

The gaze tracking is popular and widely used in various fields in recent years. The interactive devices, like Smart TV, kiosk or robot can give the diverse feedback based on eye movements. From the perspectives of the advertiser, how to attract the pedestrian's attention is a key factor to make the advertisement effective. Hence, the gaze tracking system is a reliable useful tool for researchers to explore people's visual behaviors. Base on cognitive psychology, human beings usually rely on the visual to receive the information. Therefore, gaze tracking is the most efficient way to obtain the firsthand observations regarding to human visual attention. As the aforementioned 2 reason, many researchers [1-2] are into proposing the gaze mapping method or developing the equipment for gaze tracking.

The techniques for gazing tracking on the methods for gaze estimation have been proposed numerously. According to the variation types of devices, the system can be divided into intrusive type and non-intrusive type. In the intrusive type, the former method uses the election potentials measurement with contact electrodes placed near the eyes. Another method let the user wearing the special devices to catch the user's gaze in very close to the eyes. Even though these methods and systems can provide high accuracy, but it is really not comfortable and convenient for the general user.

Hence, the non-intrusive gaze tracker has been developing recently. In this type, the proposed method is often using camera or tie in with infrared light sources to capture the eye images and record the eye movement. After obtaining the features, the kernel of non-intrusive gaze tracker is to establish a mapping function for estimating the gaze point or gaze direction. The mapping function [3-8] may base on neural network, 3D spatial model or linear regression model is according to the image capture equipment and usage scenario. Compare with the intrusive gaze tracker, user needless wear special equipment and more easily apply in general applications are the benefits of contactless system

Although these researches reported on the various eye gaze tracker integrations, but still have some limitation on the available distance range and multiple users' condition. Hence, some researchers combine Kinect and PTZ (Pan-tilt-zoom) camera to gain the distance and angular deviation between users to the monitor screen. The Face

Tracking SDK of Kinect computes the 3D positions of facial feature points in real time. It is helpful to obtain the position of head, position of the eye and the head posture including the angle of roll, pitch and yaw. Based on this characteristic, the 3D multi-model is proposed to estimate the gaze point with applying on the small screen device by [9]. Additionally, [10] estimates the gaze point by using PTZ and Kinect. The kernel mapping function is using Bayesian multinomial logistic regression, in addition, the user's head direction is computed by Kinect and the eye image is simultaneously captured from PTZ camera. This method can effectively improve the accuracy cause from the distance and multi-users though the computational complexity is higher.

In order to improve the accuracy and usability, we propose an adaptive module try to optimize the tracking and mapping procedures and apply to different environments. The remainder of the paper is organized as follows. In section II, the proposed method is presented. The experimental results are carried out in section III. Finally, the paper ends with our conclusions with discussion and recommendations for future work in section IV.

2. The Proposed Method/Algorithm

In the proposed method, we use the PTZ camera to extract the high resolution eye image, and Kinect is used to extract 3D facial feature, angle of rotation and depth value. Figure 1 is a schematic drawing that illustrates the experimental environment. To avoid covering the pupil with eyelid, two image capture devices are set under the monitor.
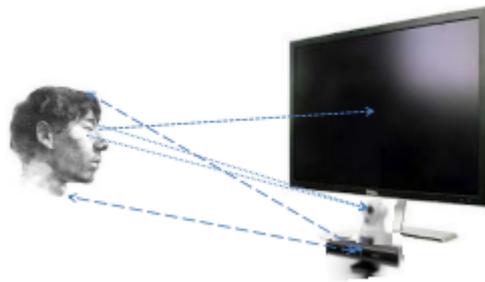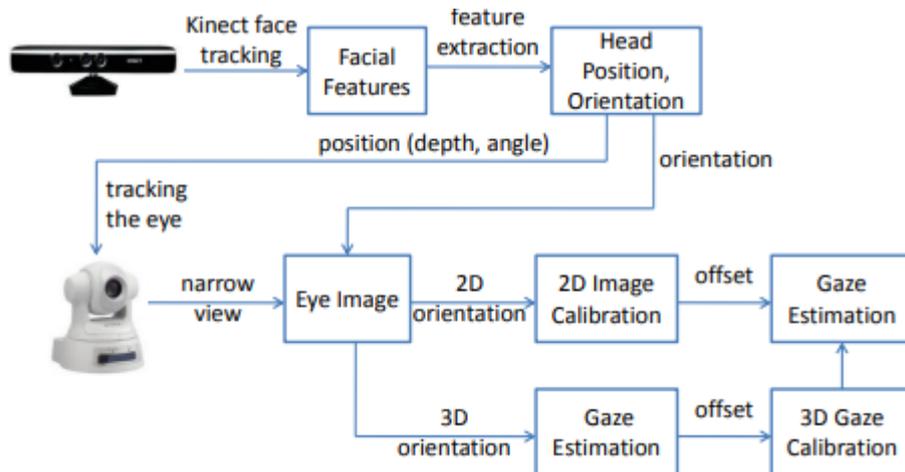

Figure. 1. Experimental environment

Figure. 2. Block diagram of proposed method

The several factors that affect accuracy of eye gaze tracker including: image resolution, ambient light and head movement problem, hence the calibration process is necessary to improve the accuracy of gaze estimation. Figure 2 shows the block diagram of an adaptive gaze tracker system. After obtaining the 3D facial feature from Kinect, the head position, facial rotation angle and depth value also can be computed, as shown in Figure 3. These information will used to control the PTZ for capture the complete and high resolution eye image. The definition of the detail derivation as follows.
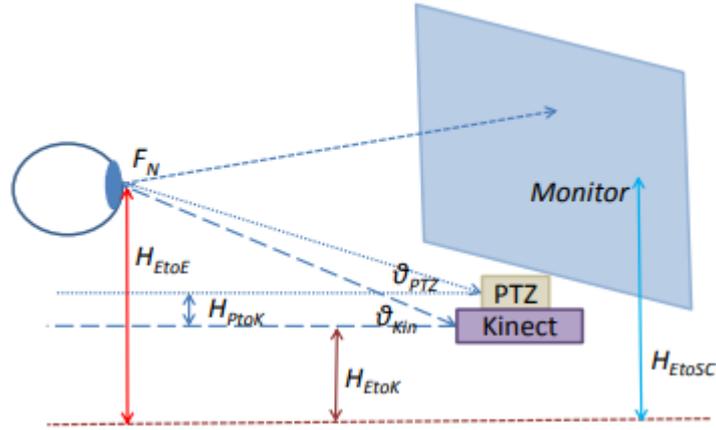
Figure. 3. A diagram between eyeball, monitor, PTZ and Kinect

Firstly, the field of view is divided into MxN blocks, as shown in Figure 4 illustrated. The divided block size is depended on the distance between the user position and the monitor. Hence, the ratio of block of target object and others is calculated as horizontal and vertical plane respectively.

Figure. 4. The screen divided into MxN blocks

Figure 5 shows the diagram of horizontal field. The ratio between left area (W_{vl}), right area (W_{hl}) and width of main target object (W_{hr}) according as equation (1)-(5).

$$D_{phs} = D_{es} \times tan\vartheta_{sh} \tag{1}$$

$$tan(\vartheta_{hl} - \vartheta_{sh}) = \frac{(W_{sh} - D_{phs} + W_{hl})}{D_{es}} \tag{2}$$

$$W_{hl} = D_{es} \times tan(\vartheta_{hl} - \vartheta_{sh}) - W_{sh} + D_{phs} \tag{3}$$

$$tan(\vartheta_{hr} + \vartheta_{sh}) = \frac{(W_{sh} + D_{phs} + W_{hr})}{D_{es}} \tag{4}$$

$$W_{hr} = D_{es} \times tan(\vartheta_{hr} + \vartheta_{sh}) - W_{sh} - D_{phs} \tag{5}$$

where D_{phs} is horizontal offset, and ϑ_{sh} is the facial horizontal rotation angle. After gain the ratio of horizontal, the vertical plane calculate the ratio between top (W_{vl}), bottom (W_{vr}) and the main object (W_{vr}) by the equation (6)-(10), as sown in Figure 6.

$$D_{pvs} = D_{es} \times tan\vartheta_{sv} \tag{6}$$

$$tan\left(\vartheta_{vb} - \vartheta_{sv}\right) = \frac{\left(W_{sv} - D_{pvs} + W_{vb}\right)}{D_{es}} \tag{7}$$

$$W_{vb} = D_{es} \times tan\left(\vartheta_{vb} - \vartheta_{sv}\right) - W_{sv} + D_{pvs} \tag{8}$$

$$tan\left(\vartheta_{vu} + \vartheta_{sv}\right) = \frac{\left(W_{sv} + D_{pvs} + W_{vu}\right)}{D_{es}} \tag{9}$$

$$W_{vu} = D_{es} \times tan\left(\vartheta_{vu} + \vartheta_{sv}\right) - W_{sv} - D_{pvs} \tag{10}$$

where D_{pvs} is vertical offset, and ϑ_{sv} is the facial vertical rotation angle. According to 5 equations (1)-(10). The diagram of visual range and target objects is as shown in Figure 7. In this case, the user is sitting on position 1 as shown in Figure 8(a). When block 5 is appointed the main target area, the ratio of horizontal can be normalizing to $\left(W_{hl}: W_{hr}: W_{sh}\right)$.
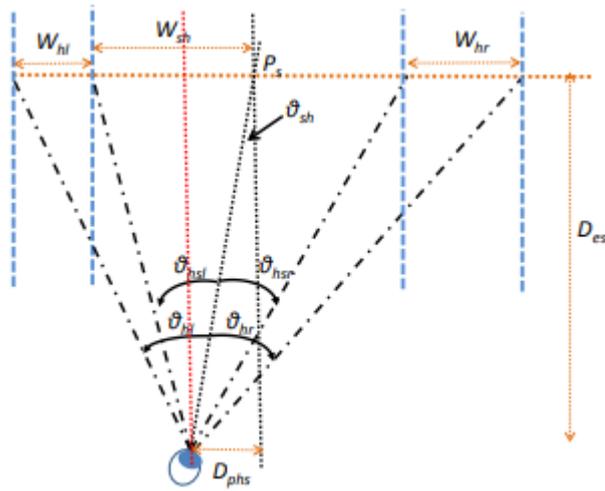

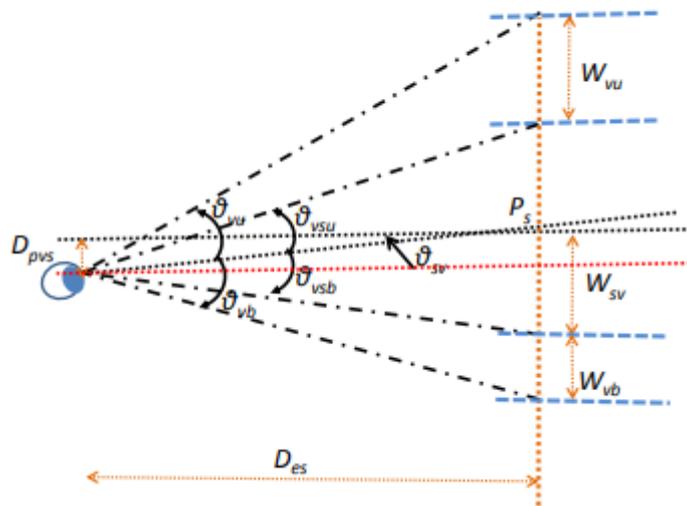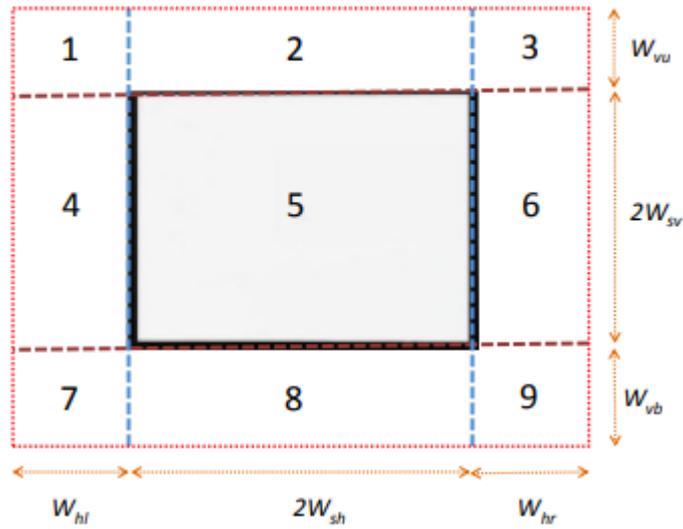**Figure. 5.** The diagram of horizontal field

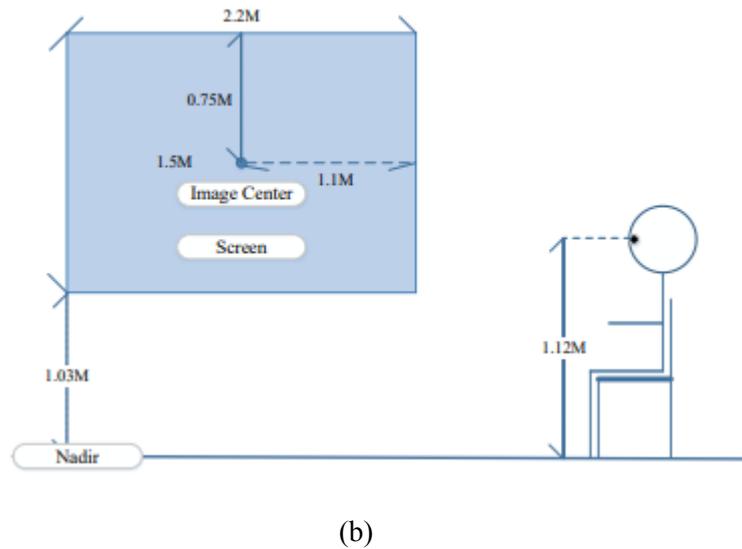**Figure. 6.** The diagram of vertical field

Figure. 7. The diagram of visual range and target objects

3. Experimental Results

In order to provide a comfortable testing environment, the visual range is 30 degree. Figure 8 showed the different testing positions and the corresponding parameters for the experiment. Table 1 shows the parameters according to Figure 7.

(a)

(b)
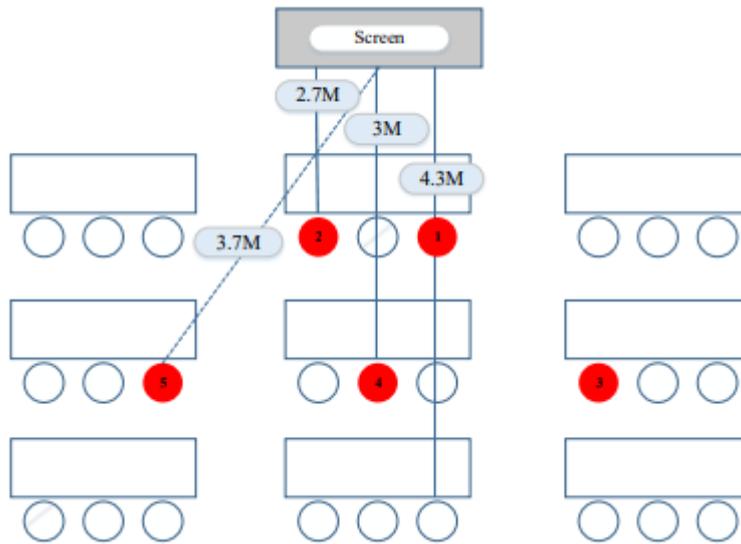
Figure. 8. (a) The different testing positions (b) Corresponding parameter

Refer to Figure 2, modify the PTZ camera by using the features of depth and rotation angle. After that, the Voting scheme [12] is used to estimate gaze points. In Figure 7, the target object is divided into 9 blocks, labeled as 1 to 9 respectively. The relationship between the user and each block size is described in Table 1. Every position will obtain the eye image of the different angle, for example, the Figure 9 illustrates the captured eye image. In this case, the user is sitting in the position 1, a little left side of the screen, so the eye image is much clearer and the rotation angle near zero.

Table. 1. The parameters according to Figure 7

Position / Feature	1	2	3	4	5
W_{hl}(cm)	70.54	58.41	64.22	173.4	92.38
W_{hr} (cm)	52.44	58.41	137.6	15.48	92.38
W_{vu} (cm)	85.97	181.5	153.2	104.5	153.2
W_{vb} (cm)	85.97	38.67	118.6	154.4	118.6
Yaw (degree)	15	-10	10	5	0
Pitch(degree)	10	20	10	5	10
Roll(degree)	0	-5	5	0	5

After extraction the feature, the Voting scheme is used to estimate gaze point. Figure 10 and Figure 11 show the histogram of the horizontal and vertical domain respectively. For example, the captured eye image is like number 5 in Figure 9. Then the image will divide into MxN blocks, 3x3 is adopted in here, and convert color to grayscale for calculating the histogram of each eye block. The block 5 in Figure 10 present the value of eye block 2, 5, 8 is lower than other eye blocks, that means the user maybe focus on the monitor, block 2, 5, 8, refer to Figure 7. After calculating the horizontal domain, the voting scheme will be used again to calculate the vertical domain to decide which the monitor block is the user's real foci.
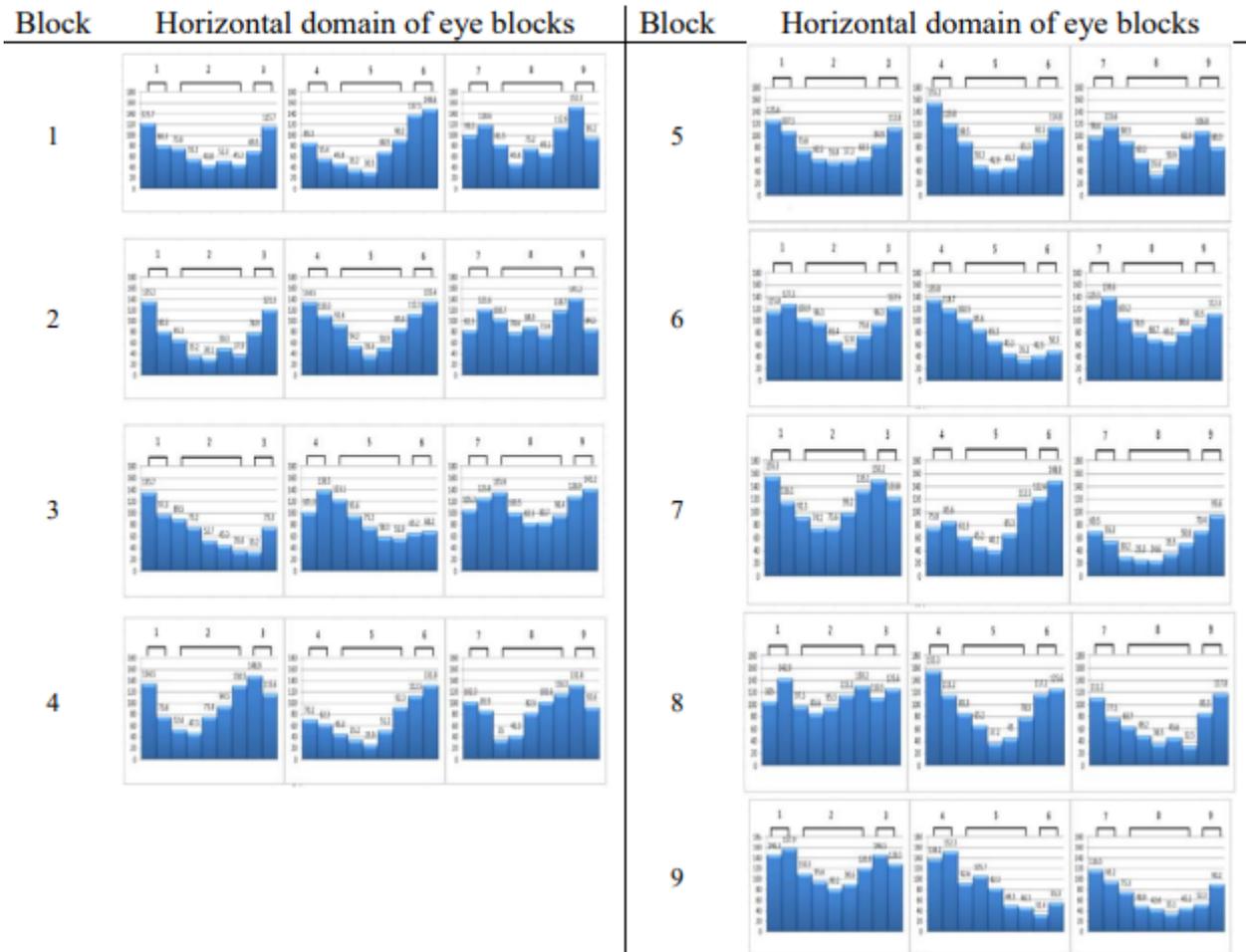
Block	Horizontal domain of eye blocks	Block	Horizontal domain of eye blocks
1		5	
2		6	
3		7	
4		8	
		9	

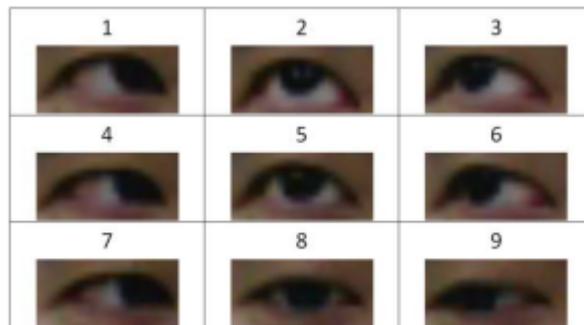**Figure. 10.** The histogram of horizontal domain

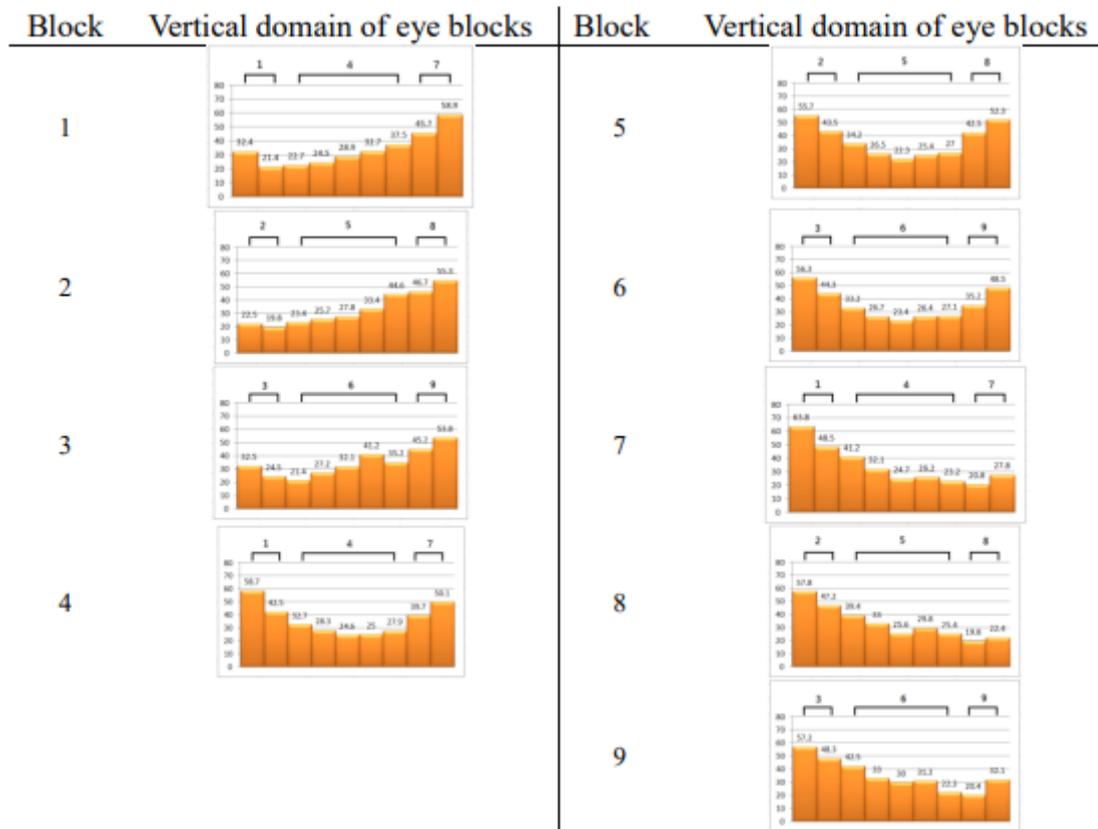**Figure. 9.** The captured eye image of the 9 blocks

53

Figure. 11. The histogram of vertical domain

4. Conclusion

The gaze tracking is an objective way to gather the eye movements of people. The collected data can help an educator understand their learning behaviors. In order to record the learning behaviors in the diverse environments, an adaptive gaze tracking method is proposed in this paper. This system not only allows the multiple users in a normal environment, and also is applicable with different types of devices. The accuracy of this system is about 82% influenced by ambient light. Even though the accuracy is not excellent, but the provided information is enough for analyzing the learning behaviors. From this point of view, our results show that the estimation of gaze points successful even when the distance and head rotation altered. In the future work, we will survey new PTZ camera with better zoom lens to increase the accuracy and will try to use multi-monitors to create a more complex scenario.

References

[1] H. C. Chen, H. D. Lai, and F. C. Chiu. "Eye Tracking Technology for Learning and Education," Journal of Research in Education Sciences, vol. 55, no. 4, pp. 39-68, Dec. 2010.

[2] D. L. Tang, T. R. Lee, and C. M. Tsai. "An Exploratory Study on Relationship between Preference and Scanpath-Evidence from Color Preference Sorting Task," Chinese Journal of Psychology, vol. 47, no. 4, pp. 339-351, Dec. 2005.

[3] Hansen D. and Q. Ji, " In the eye of the beholder: A survey of models for eyes and gaze," IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol.32, Issue 3, 2010, pp. 478–500.

[4] Sigut J. and Sidha S., "Iris Center Corneal Reflection Method for Gaze Tracking Using Visible Light," IEEE Transaction on Biomedical Engineering, Vol. 58, Issue 2, 2011, pp. 411-419.

[5] Juan J. Cerrolaza, Arantxa Villanueva, and Rafael Cabeza, "Study of Polynomial Mapping Functions in Video-Oculography Eye Trackers," ACM Transactions on Computer-Human Interaction, Vol. 19, Issue 2, 2012, Article 10

[6] Xiao-Hui Yang, Jian-De Sun, Ju Liu, Xin-Chao Li, Cai-Xia Yang, and Wei Liu, "A Remote Gaze Tracking System Using Gray-Distribution-Based Video Processing," Biomedical Engineering: Applications, Basis and Communications, Vol. 24, Issue 3, 2012, pp.217-227

[7] Y. Sugano, Y. Matsushita, Y. Sato, "Appearance-Based Gaze Estimation Using Visual Saliency," IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 35, No. 2, Feb. 2013, pp. 329-341.

[8]   H.R. Chennamma and Xiaohui Yuan, "A survey on eye-gaze tracking techniques," Indian Journal of Computer Science and Engineering, Vol. 4, Issue 5, October 2013, pp. 388 - 393.

[9] Kenneth Alberto Funes Mora, Jean-Marc Odobez, "Gaze Estimation from Multimodal Kinect Data," IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), pp. 25-30, 2012.

[10] Reza Jafari and Djemel Ziou, "Gaze estimation using Kinect/PTZ camera," IEEE International Symposium on Robotic and Sensors Environments (ROSE), 2012, pp. 13- 18.

[11] C.W. Kao, B.J. Hwang, C.W. Yang, K.C. Fan, C.P. Huang, "A Novel with Low Complexity Gaze Point Estimation Algorithm," First International Conference on Instrumentation, Measurement, Computer, Communication and Control (IMECS 2012), pp. 204-208, March 14-16, Hong Kong, 2012.