# A Novel with Low Complexity Gaze Point Estimation Algorithm

Chiao-Wen Kao, Bor-Jiunn Hwang, Che-Wei Yang, Kuo-Chin Fan, Chin-Pan Huang

*Abstract*—**In this paper, a novel with low complexity gaze point estimation algorithm in unaware gaze tracker is proposed which is suitable in normal environment. The experimental results demonstrate our proposed method is feasible and has acceptable accuracy. Besides, the proposed method has less complexity in terms of camera calibration process than traditional method.**

*Index Terms*—**Gaze Point Estimation; Unaware Gaze Tracker; Voting Scheme**

## I. INTRODUCTION

Interactive Installation is the most popular issue in recent years. Such as using Hand Gesture, Human Posture, Eye Detection, Gaze Tracking, Speech recognition to control the computer, device or play games. The Gaze tracking can be used in many applications such as web usability, advertising, sponsorship or in communication systems for disable people. Numerous techniques of eye gaze trackers have been developed [1-13]. These eye gaze tracker found in literature can be divided into two groups, intrusive techniques and non-intrusive techniques, respectively. Intrusive methods usually use special devices to attach the eye skin or wear head-mounted to catch the user's gaze in very close to the eyes [1].

The most widely used current designs for eye trackers are using a non-contacting video camera to focus on eyes and records their movement. Compared with intrusive methods, the non-intrusive methods have the advantage of being comfortable during the process of gaze estimation [13]. Video-based eye trackers typically use the corneal reflection and the iris center as feature to track over time [2-12].

The gaze calibration procedure that identifies the mapping from pupil parameters to screen coordinates using neural network has become more popular for eye gaze tracker. Baluja and Pomerleau proposed a neural network method without explicit features [3]. Each pixel of the image is considered as an input parameter of the mapping function. Once the eye is detected, the image of the eyes is cropped and then used as the input of ANN (Artificial Neural Network). In [9], authors proposed remote eye gaze tracker based on eye feature extraction and tracking by combining neural mapping (GRNN) to improve robustness, accuracy and usability under natural conditions.

For 3D model-based approaches, gaze directions are estimated as a vector from the eyeball center to the iris centers [8]. A stereo camera system is constructed for 3D eye localization and the 3D center of the corneal curvature in world coordinates. Points on the visual axis are not directly measurable from the image. By showing at least a single point on the screen, the offset to the visual can be estimated. The intersection of the screen and the visual axis yield the point of regard.

The purpose of this paper is to propose a novel with low complexity gaze point estimation algorithm in unaware gaze tracker and which is suitable in normal environment.

The remainder of the paper is organized as follows. In section II, the proposed Voting scheme algorithm is presented. The gaze evaluation model and results are carried out in section III. Finally, the paper ends with our conclusions with discussion and recommendations for future work in section IV.

## II. PROPOSED VOTING SCHEME ALGORITHM

A gaze tracker is used to acquire eye movements. A general overview of the gaze tracker is shown in Fig. 1, comprising Face Detection, Eyes Detection, Eyes Tracking and Gage Estimation. Eyes Detection and Gaze Estimation are important functionality for many applications including driver's physical condition analysis, helping disabled people operate computer, auto stereoscopic displays, facial expression recognition, and more. The eye positions should be calculated first to estimate the person's gaze coordinates. This section describes an algorithm for tracking gaze direction on the screen.

### A. Preprocessing

Several preprocessing steps must be done before performing gaze tracking, as shown in Fig. 1. Firstly, detecting face in image is a fundamental task for surveillance system. This paper use Haar-like Features which firstly proposed by Paul Viola and Jones to detect the face [14] [15]. Haar-like features are digital image features used in object detection and recognition. . Each classifier uses K rectangular areas to make decision which the region of the image likes predefined image or not. Fig. 2 exhibits the Haar-like shape features sets including Line features, Edge features and

Manuscript received December 30, 2011; revised January 17, 2012.
Chiao-Wen Kao is with the Department of Computer Science and Information Engineering, National Central University, Taoyuan, Taiwan. (e-mail: chiaowenk@gmail.com).
Bor-Jiunn Hwang is with the Department of Computer and Communication Engineering, Ming Chuan University, Taoyuan, Taiwan. (e-mail: bjhwang@ mail.mcu.edu.tw).
Kuo-chin Fan is with the Department of Computer Science and Information Engineering, National Central University, Taoyuan, Taiwan. (e-mail: kcfan@csie.ncu.edu.tw).
Che-Wei Yang is with the Department of Computer and Communication Engineering, Ming Chuan University, Taoyuan, Taiwan. (e-mail: yang.bng86@gmail.com).
Chin-Pan Huang is with the Department of Computer and Communication Engineering, Ming Chuan University, Taoyuan, Taiwan. (e-mail: hcptw@mail.mcu.edu.tw).

Center features.

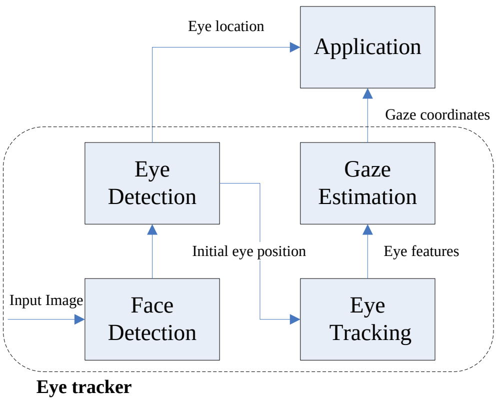


Fig. 1. General overview of the components of eye and gaze tracker

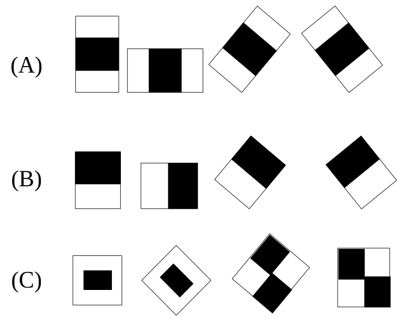


Fig. 2. Haar-like shape features sets
(A) Line features (B) Edge features (C) Center features

The eye features are similar to the facial structure, so we also used Haar-like features to detect eyes. The face and eyes detection results are shown in Fig. 3.

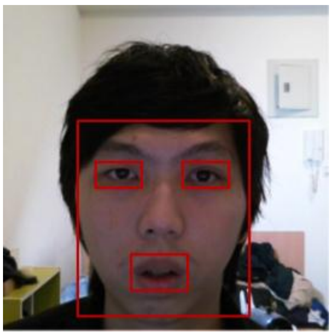

Fig. 3. Face and eyes detection

Nevertheless, the eyes detection results may have missed caused by marking mouth. The eye candidates' positions is satisfied the facial structure. Therefore, the follow processes, namely correction process, are proposed to determine the eye candidates accurately.

*(i)* According to the facial structure, the eyes are located at region of 2/3 along the vertical dimension usually.

*(ii)* The procedures is converting eye candidates to YCbCr color space and then using the skin color filter to remove skin pixels. In other words, skin color threshold, RCr=[133,173] and RCb=[77,127], is used to redefine the region of eye candidates. Fig. 4 shows the correction process.

*(iii)* Finally, two more fitting regions of eye candidate are founded. Fig.5 shows the rectified result of eye detection.

### *B. Voting scheme*

After locating the positions of eyes, Voting scheme is executed to estimation a gaze position on the monitor. Fig. 6 exhibits the flowchart of estimating gaze position comprising three macro function blocks, Initial Stage, Predict horizontal position of iris center and Predict vertical position of iris center. And which are described as following.

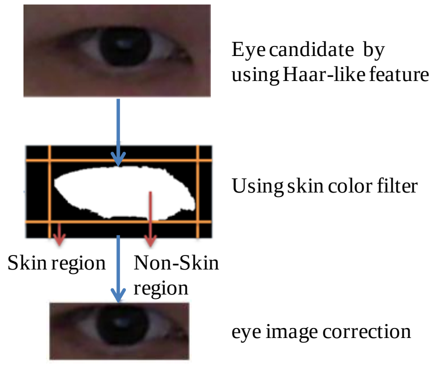


Fig. 4. Correction process

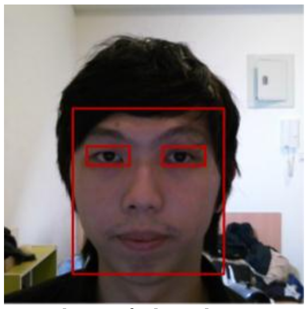

Fig. 5. The rectified eye detection

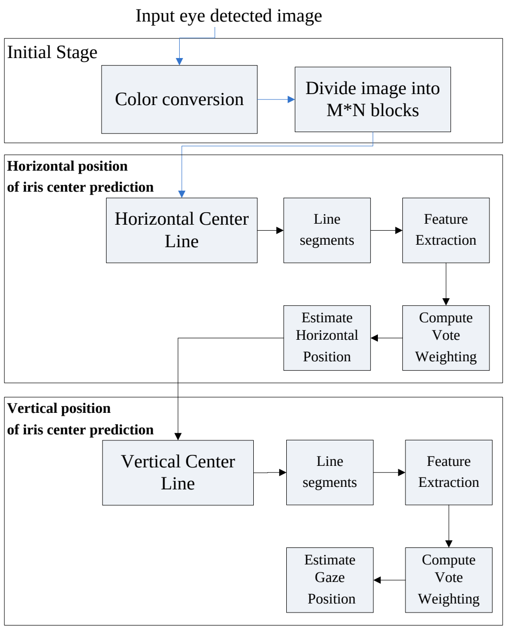


Fig. 6. Flowchart of estimating gaze

**Initial Stage:**

Step 1. From the biological point of view, it will be feasible to distinguish between iris and sclera by using grayscale. Therefore, the detected eye color image is converted to grayscale to estimate iris center position.

Step 2. The object in full-screen is divided into M*N

ISBN: 978-988-19251-1-4
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

blocks, where N along the horizontal dimension and M along the vertical dimension.

Step 3. Divide detected eye images into the same number of blocks.

**Predict horizontal position of iris center:**

Step 1. To get the center line of vertical dimension in each block, HBLij, for i=1,…N, j=1,…M.

Step 2. Divide $HBL_{ij}$ into N equal line segments, $HBL_{ij\text{-}k}$, $k$=1,…$N$.

Step 3. Compute the vertical projection and mean of the $HBL_{ij\text{-}k}$, respectively.

Step 4. Adaptive thresholds ($Th$) are obtained to quantify the mean values according to the method in [11-12].The quantified mean value $Q_{ij\text{-}k}$ of each line segment is computed by (1).

$$Q_{ij\text{-}k} = \lfloor (y - y_{base}) / Th \rfloor + 1 \qquad (1)$$

Where $\lfloor x \rfloor$ denotes the nearest integers less than or equal to x. $y$ and $y_{base}$ represent maximum and minimum mean values of $HBL_{ij\text{-}k}$, respectively.

Step 5. Sum of the quantified mean value, $S_{ik}$, is computed by (2)

$$S_{ik} = \sum_{j=1}^{M} Q_{ij-k} \qquad (2)$$

$$S=\{S_{ik} \text{ for } i=1\ldots N,\ k=1\ldots N\}$$

Step 6. Initial voting weight $Wt_{ik}$. The set $S_N$ is composed by the lowest of $N$ values in S, where

$$\begin{cases} Wt_{ik} = 1, & if\ \ S_{ik} \in S_N \\ Wt_{ik} = 0, & otherwise \end{cases} \qquad (3)$$

The block weights $Wt_i$ are obtained by summing of the voting weight as (4).

$$Wt_i = \sum_{k=1}^{N} Wt_{ik} \quad (4)$$

Step 7. Finally, to find maximum value of $Wt_i$ to determine the iris center of horizontal.

Therefore, the candidate of horizontal position of iris center can be found by using the Voting scheme.

**Predict vertical position of iris center:**

Step 1. It's a great similarity between getting the vertical and horizontal position. To get the center line of vertical dimension in $W_{ti}$ which computed by (3), $VL_j$, for $j$=1,…$M$.

Step 2. Divide $VL_j$ into $N$+2 line segments on average, $VL_{j\text{-}k}$, $k$=1,…$N$+2. From the biological point of view, vertical eye movement is smaller. Therefore we divided segment into more detail in order to improve the accuracy.

Step 3. Compute the horizontal projection and mean of the $VL_{j\text{-}k}$, respectively.

Step 4. Repeat the step 5~step 8 in Horizontal position of iris center prediction procedure.

Step 5. Finally, select maximum value of $VL_j$ to represent the iris center of vertical in this block.

Based on these procedures of Voting scheme, we can estimate the gaze position on the screen facilely. For example, assume the test object in full-screen is divided into 3*3 blocks as shown in Fig.7. And the gray scale eye image is also divided into 3*3 blocks. Thus, we can get 9 center line segments in the blocks, as shown in Fig. 8.

The results of computing the vertical projection and mean of each line segment of Fig. 8 are shown in Fig. 9 and Fig. 10, respectively. Based on Fig. 10, the quantified mean value and sum of the quantified mean value are performed by (1) and (2), respectively, the results are shown in Fig. 11. Initial voting weight is performed by (3) and then summing of the weights by (4), the results are shown in Fig. 12. The candidate of horizontal position of iris center is determined by selecting the lowest of three values as shown in Fig. 13.

Fig. 7. Divide full-screen advertisement into 3*3 blocks

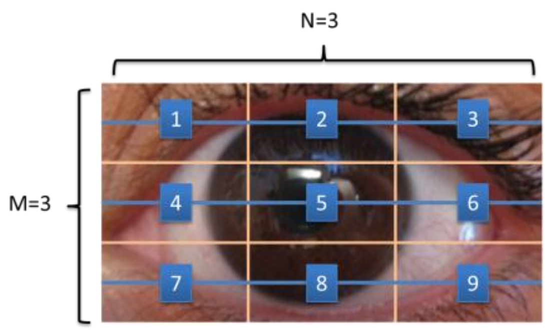

Fig. 8. Example of divide grayscale eye image into 3*3 blocks

The purpose of vertical position estimation is to determine the horizontal candidate. As experimenting, brightness spots on the iris that maybe influence the vote result. Hence, in the Voting scheme, more divided segments in vertical are performed to improve accuracy. Fig. 14 shows the estimation result, and the vertical position of iris center is determined in the block 5.

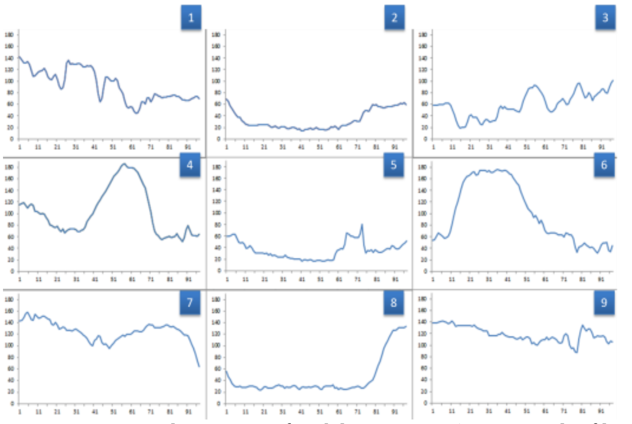

Fig. 9. Vertical projection of each line segment (x-axis: pixels of line segment; y-axis: gray scale, range of values is [0,255])

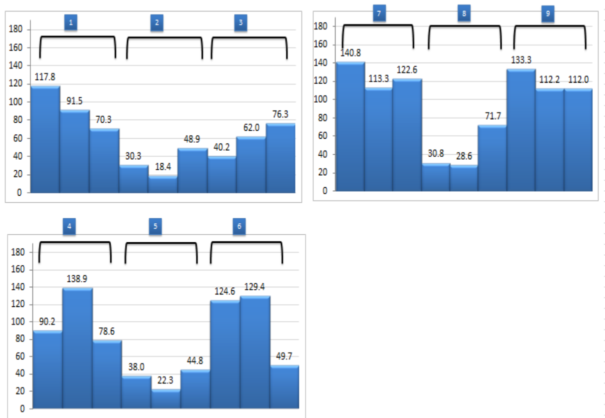

Fig. 10. Mean of each line segment pixel values

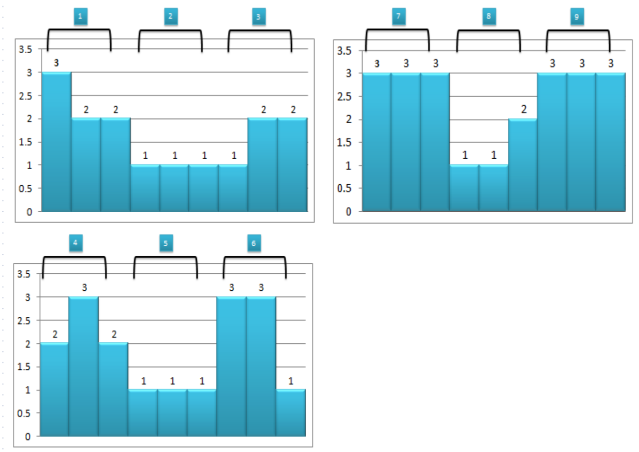

Fig. 11. Sum of voting weight of each line segment

Fig. 12. Candidate of horizontal position of iris center

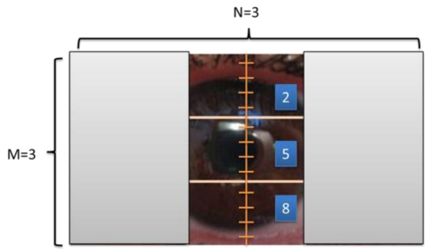


Fig. 13. Estimation horizontal position of iris center

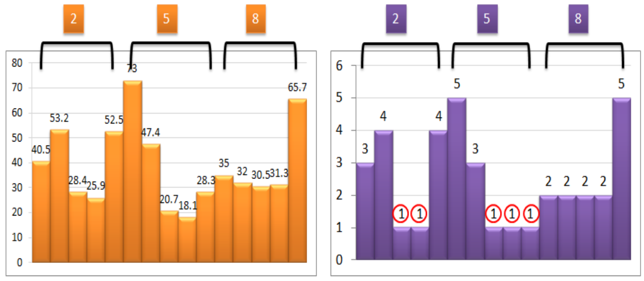

Fig. 14. Compute vote weight of each line segment

## III. Experimental Results

In this section, the experimental tests are given to evaluate the performance of proposed Voting scheme. The functionalities of tests are implemented by OpenCv on a 3.4GHz 4-GB PC environment.

We have evaluated the proposed method by three cases, as shown in Fig. 15. Case 1: White background, black target object. Case 2: Black background, white target object Case 3: White background, random target object color. The distance between participant and camera is about 50~80 cm. And the test block is emerged randomly with using red cross in the block center to attract the subject.

The experiment results are obtained by 15 participants to test each case in 3 times, and summarized in TABLE I. Based on TABLE I, the average accuracy is higher than 80% in case of 3*3 blocks. But when full-screen is divided into more than 3*3 blocks the accuracy is reduced.

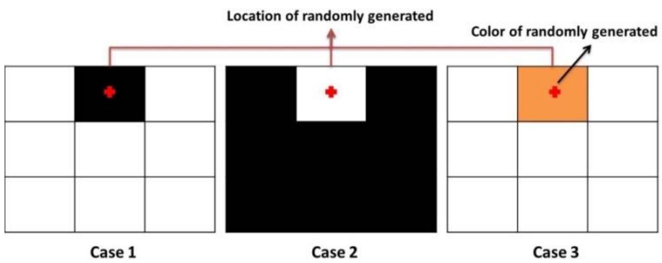


Fig. 15. Three cases in evaluated proposed method

TABLE I
THE PERFORMANCE OF THE PROPOSED APPROACH IN EACH CASE

| MxN | Case 1 | Case 2 | Case 3 | Average |
|---|---|---|---|---|
| 3x3 | 85.33% | 84.11% | 83.33% | 84.25% |
| 5x5 | 66.66% | 61.33% | 64.75% | 64.25% |

ISBN: 978-988-19251-1-4
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

## IV. Conclusion

We have surveyed several categories of eye tracking systems from the different methods of detecting and tracking eye images to computational models of eyes for gaze estimation and gaze-based applications. However, most of systems setup increases have higher both the complexity and cost. Stated thus, we propose a novel unaware method, namely Voting scheme, to estimate gaze tracking based on appearance-manifolds. In this system, the user only sits in front of a computer and use the webcam on the monitor to capture the user's image sequences. This method first calculates the histogram of grayscale eye image and use dynamic thresholds to quantify the pixel values. Then gaze direction on the screen can be predicted by using voting scheme. The experimental results demonstrate the effectiveness of proposed gaze tracking approach. Based on this, we have tried to find out how people look at content of website or advertisement. However, some problems still need to be solved. Firstly, the proposed method cannot deal with low resolution image sequences. In addition, the blurred or bad illuminated image sequences could affect the tracking result. Future work will be to deal with those problems and achieve more robust algorithm.

## Acknowledgment

This work is supported by the National Science Council in Taiwan. The project contract number is NSC 100-2221-E-130-024-.

## References

[1] Craig A. Chin, Armando Barreto et al, "Integrated electromyogram and eye-gaze tracking cursor control system for computer users with motor disabilities," Journal of Rehabilitation Research & Development, Vol.45, Num.1 2008.

[2] T. Cornsweet, H. Crane, "Accurate two-dimensional eye tracker using first and fourth Purkinje images", *Journal of the Optical Society of America*, vol. 63, pp.921-928, 1973.

[3] S. Baluja and D. Pomerleau, "Non-intrusive gaze tracking using artificial neural networks", *School of Computer Science, CMU, CMU Pittsburgh, Pennsylvania 15213*, Jan. 1994.

[4] Shota Miyazaki1, Hironobu Takano1 and Kiyomi Nakamura, "Suitable Checkpoints of Features Surrounding the Eye for Eye Tracking Using Template Matching," *SICE Annual Conference*, Step. 2007

[5] Shinji Yamamoto and V.G. Moshnyaga, "Algorithm Optimizations for Low-Complexity Eye Tracking," *IEEE International Conference on Systems, Man, and Cybernetics*, October 2009

[6] Yali Li, Shengjin Wang, Xiaoqing Ding, "Eye/eyes tracking based on a unified deformable template and particle filtering," *Pattern Recognition Letters* ,January 2010

[7] Diego Torricelli, Michela Goffredo, Silvia Conforto, Maurizio Schmi,"An adaptive blink detector to initialize and update a view-based remote eye gaze tracking system in a natural scenario,"*Pattern Recognition Letters*,June 2009

[8] Hirotake Yamazoe, Akira Utsumi, Tomoko Yonezawa, Shinji Abe, "Remote gaze estimation with a single camera based on facial-feature tracking without special calibration actions," *Proceedings of the 2008 symposium on Eye tracking research & applications*, pp. 26-28,March. 2008

[9] D. Torricelli, S. Conforto, M. Schmid and T. D'Alessio,"A neural-based remote eye gaze tracker under natural head motion",*Comput. Meth. Prog. Bio*, pp. 66–78, 2008.

[10] Hu-Chuan Lu, Guo-Liang Fang, Chao Wang and Yen-Wei Chen, "A novel method for gaze tracking by local pattern model and support vector regressor," *Signal Processing*, vol. 90, issue 4, April 2010, pp.1290-1299.

[11] Chiao-Wen Kao, Che-Wei Yang , Kuo-Chin Fan, Bor-Jiunn Hwang, Chin-Pan Huang," AN Adaptive Eye Gaze Tracker System in the integrated Cloud Computing and Mobile Device,"*ICMLC*,July 2011

[12] Chiao-Wen Kao, Che-Wei Yang, Kuo-Chin Fan, Bor-Jiunn Hwang, Chin-Pan Huang," Eye gaze tracking based on pattern voting scheme for mobile device," *Instrumentation, Measurement, Computer, Communication and Control*, Oct. 2011

[13] Dan Witzner Hansen, Qiang Ji, "In the Eye of the Beholder: A Survey of Models for Eyes andGaze," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 32, No. 3, pp. 478-500, Jan. 2010.

[14] Viola and Jones, "Rapid object detection using boosted cascade of simple features", *Computer Vision and Pattern Recognition*, 2001

[15] Takeshi Mita, Toshimitsu Kaneko,Osamu Hori,"Joint Haar-like Features for Face Detection,"*Proceedings of the Tenth IEEE International Conference on Computer Vision*,2005.


ISBN: 978-988-19251-1-4
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)