Real Time System Realization for Binocular Eyeball Tracking Screen Cursor

Kwang Ryol Ryu*, Duck Hyun Chai** and Robert J. Sclabassi*

*University of Pittsburgh Medical Center, **Mokwon University,
conan@neuronet.pitt.edu

Abstract

A real time system realization for binocular eyeball tracking cursor on the computer monitor screen is presented in the paper. The processing for searching iris and tracking the cursor are that a facial is acquired by the small CCD camera, convert it into binary image, search for the eye two using the five region mask method in the eye surroundings and the side four points diagonal positioning method is searched the each iris. The tracking cursor is matched by measuring the iris central moving position, The cursor controlling is achieved by comparing two related distances between the iris maximum moving and the cursor moving to calculate the moving distance from gazing position and screen. The experimental result are obtained by examining some adults person on the system.

Keywords

Binocular Eyeballs Tracking, Irises Tracking, FRM, FPDP, SFPDP, Screen Cursor Matching

Ⅰ. Introduction

Human interface system has been studying and researching for computer and automatical system tracking a moving eye instead of handy mouse. Recently a mouse and keyboard are trying to be removed from the main frame for the handicapped and complicate controlling system. This cutting edge technique is that a cursor on screen is tracked and moved with the pupil or iris to click a icon and letter with blink. That should be available for the handicapped computer, a sleeping alarm system for a car driver, a armed weapon controlling system in tracking gaze, a rescuing robot, industrial operation and so on.[1] They are required a recognizing face image from camera, searching for eye and tracking a moving iris. There are a few method to find eye and iris, template matching[2], geometrical model and linear interpolation method[3,4] multiple-nervous network, using perspective features[5,6], statistical model on eye and face, using reflecting lights at eye and an infrared.[7] They are focusing on a monocular and image sensor close with the eye to be uncomfortable. Now the proposed system has binocular, keep distance 30-40cm from camera to eyes, convert into binary image, search for eyes using FRM(Five Region Mask)[8,9], find out the eyeball center by SFPDP(Side Four Points Diagonal Positioning)[10,11], also find the nose is to be eliminated an error and track the horizontal and vertical gazing direction of irises designating the center point between irises and match the screen cursor on vertical and horizontal plane coordinate at real time and the same time. Finally the system is tested by samples.

Ⅱ. System Realization

2-1 Searching Eyes and Tracking Eyeballs

A facial image is acquisitioned with restricted condition, convert into binary image and track the eyes and irises. The image should be shown the eyes and irises well to adjust the critical value. A image with hair on the face and hat, and dark lighting makes it impossible to discriminate the eyes. A decent eyes and eyeballs image converts into binary image with adjusting critical value by lighting circumstance.

The image is white when the value is low at light surroundings and black when the value is high at the dark area conversely. The mixing area makes difficult to discriminate the eye and irises.

A binary image leads to speed up the processing time than gray image and RGB color image. That is composed of 2 bits making the littlest data of image. In addition, Irises are black that it is lower affected by the critical value to convert to binary image from gray and color. Thus, that leads to be easy to find and demarcate an eyes and nose from face.

Eyes and Eyeball is tracked by using FRM and SFPDP method. A most of the binary image has eyes, eyebrow, nose, hair and ear. There are five blank region on neighboring eye. That is, between eye and eye, eye and ear, eye and eyebrow, eye and nose.

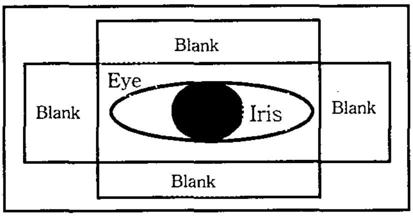
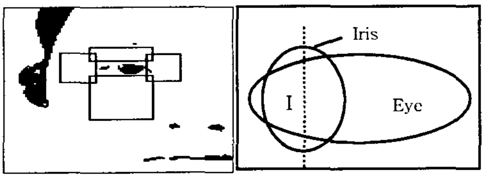


그림. 2-1 Five Region Masks

As shown in fig. 2-1, making a mask builds the FRM that the masks has 4 blanks white area(up, down, left and right) and black area of center. The fig. 2-2 (a) is shown the applied FRM for image. The eyes are recognized by only FRM.

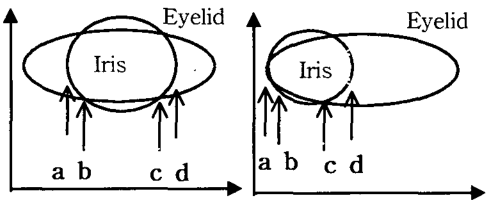


(a) Applied FRM (b) Iris Moving

Fig. 2-2 Detected eye and Iris position

The demarcating eye and iris is used a moving iris for the tracking gaze. an eye and iris is closer and a part of iris is a few covered with pupil, and eyebrow makes a shadow. In these conditions the correct iris position is detected by finding a part of bold and thick area gazing the center. However, an iris is cover with eyebrow when an iris is moving for the left and right causes error. An iris part of eye appears on the eye area with bold and thick than another area relatively. Thus, to reduce an error, like shown fig. 2-2(b), choose the center I line that the pixel distribution is wide being a pupil shades the iris.

FPDP and SFPDP methods tracks iris center position using four coordinates on the iris boundary. When the coordinate get designated, an error comes out in the particular state. As shown in fig. 2-3 the FPDP is designated the two coordinates at the boundary part of the left and right down of iris respectively. The four pixels at the left and right of eyes like the arrows is moving to the brim of iris. The coordinates are designated at the close iris brim. The irregular brim of iris at the coordinate designating section is excluded from a shading the pupil.



(a) Correct Coordinate (b) Coordinating Error

그림2-3 FPDP Coordinate Designation

However, the iris is positioned at the center is designated normally. Meanwhile, the case that the iris is moved to the left or right like the fig. 2-3 (b) is covered with the pupil brim, and a part of points for selecting the iris brim is indicated a pupil to be a incorrect position. This error is solved by SFPDP method that differs from the four points designation in the position which iris is located on the center of eye and moved to the left and right. The processes are that the left and right boundary of eye and iris is selected and measured a slope of face. The coordinate is designated by depending on the iris location. As shown in fig. 2-4 the iris center puts the cross point (Xh,Yh) on the perpendicular bisector passing f1(Xf,Yf) and f2(Xg,Yg). The coordinate for iris center expression is found by equation (1). Where Xf and Yf are mean between a and c, and Xg and Yg are median between b and d.

2-2 The measurement of irises moving

(1) Horizontal moving

The two irises and eyes distance are measured for calculating the horizontal moving distance.

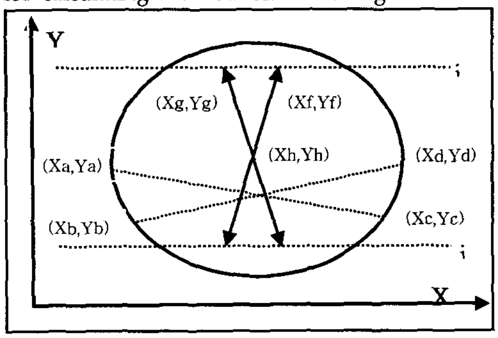


Fig. 2-4 Irises Center Position Tracking Coordination

$$f_i(X_i,Y_i) = ((-Y_f - \frac{Xc-Xa}{Yc-Ya}Xf + Yg + \frac{Xd-Xb}{Yd-Yb}Xg),\ (-\frac{Xc-Xa}{Yc-Ya}(-Yf - \frac{Xc-Xa}{Yc-Ya}Xf + Yg + \frac{Xd-Xb}{Yd-Yb}Xg) + Xf + \frac{Xc-Xa}{Yc-Ya}Xf)) \quad --(1)$$

And the left and right moving distance of iris is calculated by the differences between eyes and irises distances. The processing for horizontal moving distance measurement for binocular irises are the designating location, selecting the center, binocular location and finding the distance by the center position between binocular and irises. The fig. 2-5 is shown the calculating distance between irises and eyes. IC means the median value of eyes between the left iris position L and right iris position R. EC points out the center value between the end of eyes. And distance d is difference of them.

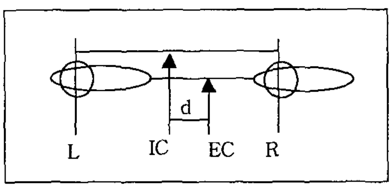


Fig. 2-5 Measurement of horizontal moving

The initial value of d is that irises is gazing at the center of screen. d is calculated by depending on irises moving. The left moving has minus and the right has plus. That result leads to the horizontal moving direction by the measurement d value.

(2)Vertical moving

The first process searches for the nose location for vertical moving distance. The vertical moving distance is calculated by the differences between the coordinate of irises center and nose in gazing at the center and the another sight on the screen. The processing are the designating location of binocular, searching for nose position to below direction, selecting the basis point for irises moving horizontal, and the measurement of distance between vertical moving position of irises. Nose pixels location is found by scanning to below direction between eyes along the perpendicular bisector as shown in fig. 2-6.

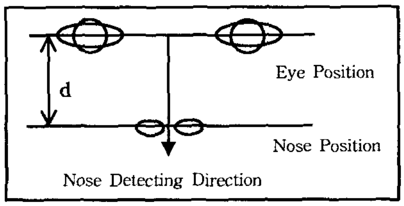


Fig. 2-6 Tracking nose position

The shadow of nostrils appears to be throwing light over the face from up to down direction. That is distributed from the horizontal lines long. This feature enable to designate the nose location. d is the distance between irises and nose. The initial value d is set by gazing at the center on screen to be fixed the face. When irises is moving to up and down, the distance m results from irises distance and d as shown in fig. 2-7. If m is plus value subtracting m from the irises position, then irises is moving to up direction, and if minus, then down. Thus the vertical moving distance is measured by depending on m.

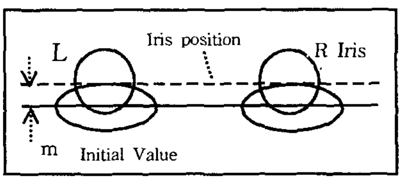
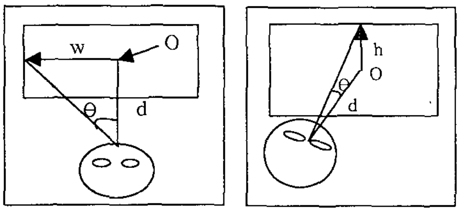


Fig. 2-7 Measurement of vertical moving

2-3. Screen cursor matching

The moving irises should correspond with screen cursor. The processing are the designating the initial value, the measuring the maximum moving distance of irises, the measuring the maximum visual angle and calculating the angle between the maximum and the real value relatively, and tracking cursor on the screen. The relation for the visual direction and visual point on the screen enable to find out in fig. 2-8. O is indicated the center of screen. The basis point gazing at the center on the screen is the center point between irises. D is distance between the center of irises and screen in fig. 2-8. w means horizontal gazing distance, and Θ is the angle between O and the gazing the end of the screen. h and w indicates the vertical and horizontal staring distances.



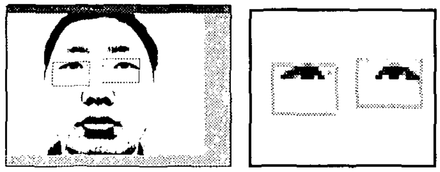
(a)Hor. moving (b)Vert. moving
Fig. 2-8 Gazing position on screen

The intial value for eyes and irises coordinates is designated from the irises distance 0 when is gazing at the center of screen in fixing the face in fig. 2-5 and 2-7. the pixels unit of the Irises distance is converted into Θ. The processing is related with the maximum variation for the irises maximum moving distances and angles *Dmax* and *Θmax* in gazing at the end of the left and right, top and bottom on the screen, and the real horizontal *Dh*, *Θh* and vertical value *Dv*, *Θv* in fig. 2-5 and 2-7 respectively. That enables to be expressed with $D_{h_{max}} : \theta_{h_{max}} = D_h : \theta_h$ for the horizontal, and $D_{v_{max}} : \theta_{v_{max}} = D_v : \theta_v$ for the vertical. The horizontal and vertical distances are calculated with $w_h = d \times \tan\theta_h$ and $w_v = d \times \tan\theta_v$. The relation expression are matched an eyes and screen.

III. Experiment and Results

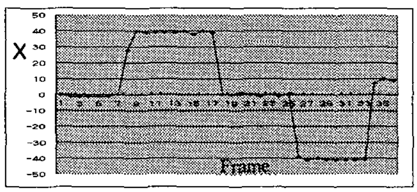
3-1 Experimental Arrangements

The experimental arrangement are composed of screen, camera and face. The lighting is illuminated from up to down to be shown irises clearly. The camera puts below the face. The distance between face and screen keeps 43cm. The camera is CCD color with 3.0M pixels, the height placing is 13cm, and the distance to eyes keeps 33cm. The screen is used the LCD 18.1 inch, resolution has 1280x1024 pixels, and the effective screen size is 26x28.5cm. That contains about 0.028cm per pixel. The color image is inputted as 680x480, convert into binary image, applied to FRM, and track the eyes as shown that result in fig. 3-1.

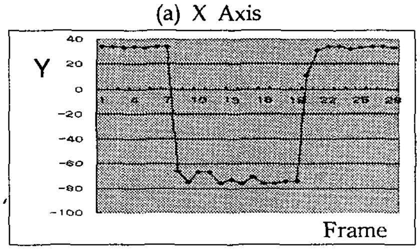


(a) FRM Binocular (b) Detected Eyes
Fig. 3-1 Binocular Detected Images

The irises position are tracked at inside eye by designating the eyes boundary in the irises position and the end edge of eyes as shown in fig. 3-1 (b). The tracked binocular and irises is shown with square and plus symbol. The iris center coordinate is applied to FPDP and SFPDP. The four points for tracking the iris center position are indicated with the cross symbol. The screen position coordinate (x,y) is shown at the top of the window. The horizontal and vertical X, Y are measured for the calculating the maximum moving distances of the irises center coordinates. Fig. 3-2 is shown the variations to X, Y coordinates respectively.



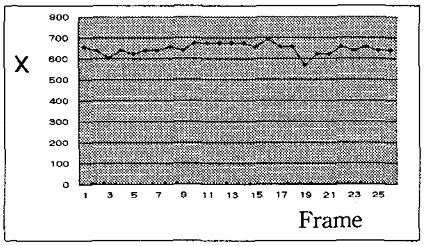
(a) X Axis



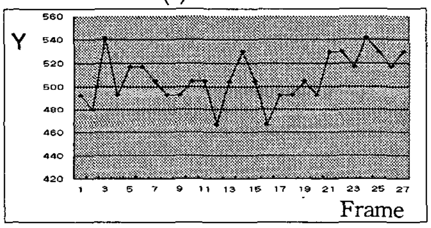
(b) Y Axis
Fig. 3-2 Iris's maximum moving range

3-2 Experimental results

X axis is moved by 40 pixels in gazing at the end of the right on the screen, and -40 pixels, at the end of the left. Y axis is moved by 34 pixels in gazing the top on the screen, and -77 to -68 pixels, at the bottom of the screen. The gazing of the center on the screen points out by 0 pixel at the X axis and -8 to -3 pixels at the Y axis. That results in X axis to be steady and Y axis to be variable error relatively. Fig 3-3 is shown the horizontal and vertical movement of screen cursor in gazing at the horizontal and vertical direction based on the center of screen, and the gazing position is moved the end of screen to the other for 2 seconds. That is tested by 10 persons and repeated 6 times.



(a) X Axis



(b) Y Axis

Fig. 3-3 Screen cursor moving

The mean error pixel for the moving screen cursor has 75 that is the maximum moving distance of irises at the X axis, 166 and 60 is the distance between irises and nose in gazing at the top and bottom of the center of screen at the Y axis. The center coordinate on the screen (X,Y) is positioned (0,132). X coordinate for the vertical movement puts on (642,3), and Y coordinate for the horizontal movement has (510,17). That results in the error designating after calculating the distance difference between the visual point and screen position. The mean error for accuracy of the system is compared with the conventional methods as shown in table 3-1. The system is equal to multiple nervous network method relatively. The frame variations per second are examined two case which is found the eyes and not found it as shown in fig. 3-4. The higher number is faster. The case of Not-found-eyes is shown 0.5 fps on the average, and the found-eyes, about 3.75 fps speed. The case of found-eyes is faster 0.6 frame(2.4 times) than multiple nervous network, 0.1(1.5 times) than linear interpolation.

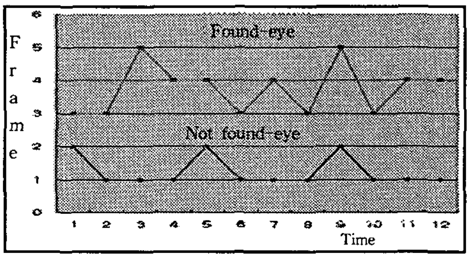


Fig. 3-4 Processing speed measurement

Table. 3-1 Comparison with methods

methods/comparison	multiple-nervous network	Template matching	Linear interpolation	The system
speed (fps)	1.11	0.91	2.5	3.75
mean error (cm)	0.47	0.42	0.47	0.47

IV. Conclusion

The system realization at real time for binocular eyeballs tracking by FRM, FPDP and SFPDP methods and screen cursor matching on the pc monitor is proposed in this paper. The screen cursor is controlled by corresponding with the irises maximum moving distance and the screen cursor maximum moving distance. The cursor moving distance is calculated in proportion to irises movement in gazing at the monitor screen. The system test is processed with 10 person iteratively, and compared with the conventional methods. the mean error for the horizontal and vertical movements has 17.4 pixels and 7 pixels. The mean error between practical gazing point and screen cursor position is 0.47cm. The processing speed is faster in comparison with the referenced methods by about 2 times. In conclusion the system is faster and simpler. This system enables to apply to the handicapped to use a computer, and is required the study not to be affected with illumination of surroundings and not to be reflected in the lighting of eyeballs.

References

[1] Hideo Kawai, Shinichi Tamura, "Eye Movement Analysis System Using Fundus

Images." Pattern Recognition. Vol. 19, No.1, pp. 77-84, 1986

[2] Hori, Y. Shimizu, K. Nakamura, Y. Kuroda, T."A real-time multi face detection technique using positive-negative lines-of-face template" Pattern Recognition, vol.1, pp. 765-768, 2004

[3] Jae Hee Kim, "A study on the Development of Data Base of Human Sensibility and Its Support for the Commercial Product", Yonsei University, 1999

[4] S.-H.Jeng, et al., "Facial feature detection using geometrical face model: an efficient approach", Pattern Recognition, vol.30, pp.273-281, 1997

[5] S. Baluja & Pomerleau "Non-intsusive Gaze Tracking Using Artificial Neural Networks" Technical Report CMU-CS-94-102, Carnegie Mellon University.

[6] Jae Hee Kim, "A Visual Computer Interface for Ergonomics", Yonsei University, 2000

[7] A. Haro, M. Flickner and I. Essa, "Detection and tracking eyes by using their physiological properties, dynamics and appearance", Proceedings of IEEE CVPR 2002

[8] D. Maio, D. Maltoni, "Real-time face location on gray-level static images", Pattern Recognition, Vol.33, pp.1525-1539, 2000

[9] Huck hyun Chai and Kwang Ryol Ryu, "An Enhancement of Tracking Capacity for the Centering Position of Eye and Iris on the Facial Image", J. of KIMICS. Vol.8, No.2, pp.323-326, 2004

[10] Duck Hyun, Chai and Kwang Ryol, Ryu, "A Study on the Enhancement of Tracking Capability for Iris Image", *Proceedings of ISMICS'2004*, pp. 24-27, 2004

[11] Duck Hyun Chai and Kwang Ryol Ryu, "Iris Position Tracking Improvement using S-FPDP on the Moving Iris", J. of KIMICS. Vol.9, No.5 pp.249-253, 2005.