Eye Gaze System to Operate Virtual Keyboard

Vidya Itsna Saraswati, Riyanto Sigit, Tri Harsono
Department of Computer Engineering
Electronics Engineering Polytechnic Institute of Surabaya
E-mail : vidyaitsna@yahoo.com, riyanto@pens.ac.id, trison@pens.ac.id

***Abstract*—For many people who has disability in communication, giving some information to other people is a hard thing to do. To make the communication easier, some people who is not able to speak or has disability in motoric function espescially hand, need a system that can help them to communicate. Based on that condition, in this research a system that can help them to communicate was built. This system used eye gaze to write some information with virtual keyboard. The process are user's face is captured by a camera, then the camera do some feature detection of face to get position of eye gaze, and this position will become a reference of pointer position to choose some word in virtual keyboard. HaarCascade method was used to detect some feature of face, and Integral Projection method was used to get position of the eye movement. Based on the results of the experiment, the comparation ratio between duration of normal writing and typing using system for two words is 1:13 second. With those two words that was typed in this system, user can give an information quite clearly enough.**

***Keyword— eye movement, virtual keyboard, Haar Cascade, Integral Projection*.**

I. INTRODUCTION

Communication is very important things in human life to connect and make an interaction with others. But, some people could not do communication very well because they have some disability to do that. For example, someone who could not speak and their motoric function could not work normally will make them could not do communication very well. Because some people who has dissablity in their motoric function especially in their hands could not do some movement to make a sign language.

Eye based HCI have been developed by many researchers. But, some system that have been developed is using some hardware, such as infrared (IR) [5][9][10] or electroculographs (EOG) [4]. The IR camera is used for allow user's movement. To robust against various type of users with different feature of eye, effect of illumination, user's movement, etc, knowledge about feature has been used to eliminated the influences [8][9]. But, when user use the eye based HCI system for a long period of time, they may feel tired after several minutes. So, the system should consider that problem. A moving keyboard has been developed to eliminate the fatigue effect [10]. There are all keys selection from the keys in the existing conventional keyboard by only looking at right or left with small angle changes.

In this paper, we developed a virtual keyboard using eye gaze system that is displayed on monitor, so user could type some words or simple sentence. For the disable people or handicap person, they have some dissability in their motoric function, so using this system they could type the words using eye gaze. Eye gaze is a natural interaction for every person that is so easily to do. The system consist of a virtual keyboard application that is operated using eye gaze of the user. So, with this system they can give some information to the others easily. In this system, there is a pointer that is moved using eye gaze of the users. So, the users should moves their eye gaze to choose the letter that they want. With this system, it is expected that some people who has some disability could do communication well. So, the other people can understand their desire or the information from them.

II. IMAGE CAPTURE AND FACE DETECTION

In this research, image capture process of the user is using a webcam with pixel size is 640x480. The type of this webcam is Logitech C525. The position of the user's head should in straighten up position and parallel in vertical axis. The captured image is converted to be grayscale format. It will make the next process is faster, because the computation process only occurs in one layer, that is gray layer. The example of image capture process is shown in Fig.1.

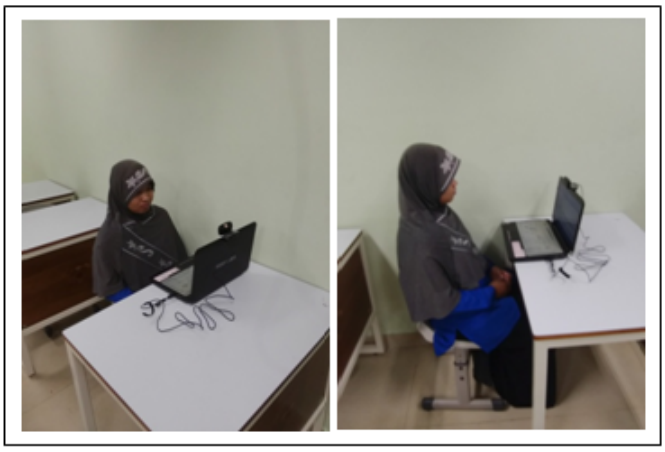

Fig.1. User's position to the camera

After the user's image is captured, the next process is detect the face of the users. This process using Haar Cascade method that is an object detection method from Paul Viola and Michael Jones. This method is a machine learning based approach which is cascade function is trained from some positive image and negative image. Positive image contain of face image, and negative image contain of non face image. These sample is used to train the classifier. Then, feature extraction from those classifier should be done. The feature that is used by Viola and Jones is based on Wavelet Haar. Wavelet Haar is a rectangle single wave. The

computation of feature value use a media integral image. The data training process in this method will generate a cascade which is contain of stage-stage as decision tree to determine which object that is a face or not. The example of the result from face detection process is shown in Fig.2.

Fig.2. Face Detection Result

III. FACE AREA CROPPING AND EYE DETECTION

From the previous process, the user's face is detected. Then, a face area cropping process will be done to the image of detected face to makes eye detection area to be smaller. In this research, the user's eye that is used is the right eye, so after the right face area has been cropped, the right eye of the user will be detected. Face area cropping process use cvSetImageROI() function that is used to take the area as Region of Interest (ROI) by make a rectangular that is determine center interaction area. The result of face area cropping of the users is shown in Fig.3.

Fig.3. Face Area Cropping Result

After the right area of the user's face is cropped, then the next process is eye detection. The method of this process is HaarCascade method that is a same method of face detection process. Fig.4. shows a eye detection result.

Fig.4. Eye Detection Result

IV. IRIS SEGMENTATION

To get the eye gaze position, the first step is thresholding process of detected eye. The type of this thresholding process that is used is binary threshold. Which is if pixel intensity of the image more than threshold value, the value of that pixel is 255 (white). But, if the intensity of the image less than threshold value, the value of that pixel is 0 (black). This thresholding process is used to remove all color, except the black color. So, the black eye area or iris will be showed.

The result of iris segmentation process is crucial and affect the position of eye gaze that will be taken. So, the morphology open-close should be used too to remove the small area and fill the hole in the image. Morphology open-close process will make the segmentation result to be better. The result is shown in the Fig.5.

(a) (b)

Fig.5. Iris Segmentation Result. (a) without morphology. (b) using morphology

V. EYE GAZE POSITION

The process to get the eye gaze position is find the center point of iris. This center point will be used as a reference for pointer movement. To find the center point of iris, integral projection equation that has been modified is used. This modification aims to make the result of integral projection equation is a position value of iris center point.

In this process, there is an addition process that is should be done. It is resize the image of iris segmentation to be bigger. That will make the change of position value is easily to observ.

This is the equation to find the center point of iris :

$$wx = \frac{\sum_{x=0}^{k} x.px}{jp} \quad (1)$$

$$wy = \frac{\sum_{y=0}^{k} y.py}{jp} \quad (2)$$

Where :
wx,wy : center point (x,y) of iris
x,y : x and y position that is calculated
px, py : sum of pixel in x and y
jp : sum of pixel in the image

VI. POINTER MOVEMENT

When the application of virtual keyboard is started, there is an initialization process which is the user should see the right side and left side. This initialization process has two result, there are minimal value of x position for right movement (wx_right), and minimal value of x for left movement (wx_left). The pointer movement system will use those value as a reference to move the pointer. When the user see the right side, the pointer will be move to the right until users see the normal position again.

The key layout that is used in this system is mobile phone key layout. And some people might be familiar with that key layout. The virtual keyboard for this system is shown in Fig. 5.

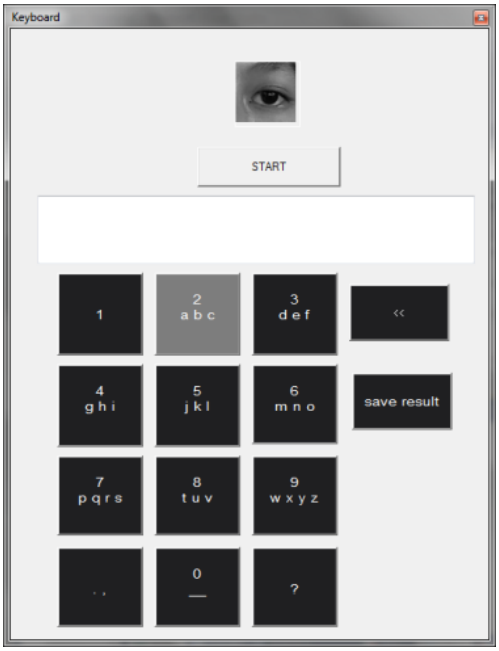

Fig.5. Proposed Virtual Keyboard

In the virtual keyboard that has been developed, there is a space to show the detected eye of user. The button that has different and brighter color is the pointer of this system. For example, in Fig.5., the pointer is pointing the "2abc" button.

In this pointer movement system, there are two movement mode, vertical movement and horizontal movement. At first, the pointer movement is using horizontal movement which is the location of the pointer is in the all of button in one column. The illustration if this movement is showed in Fig.6. At first, the location of pointer is in the first column of button. Then user see at right, so the pointer will be moved to one column at right until user see the normal position to stop the pointer movement.

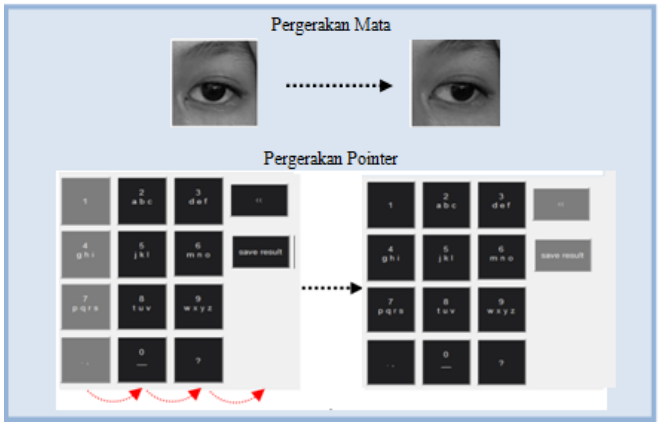

Fig.6. Eye Movement and Horizontal Pointer Movement

If user want to changes the mode of pointer movement to be vertical movement, users should do eye blink within minimal 0.5 second. After that, right-left eye movement will be detected as upward-downward direction of pointer movement. The illustration of the vertical pointer movement is showed in Fig.7.

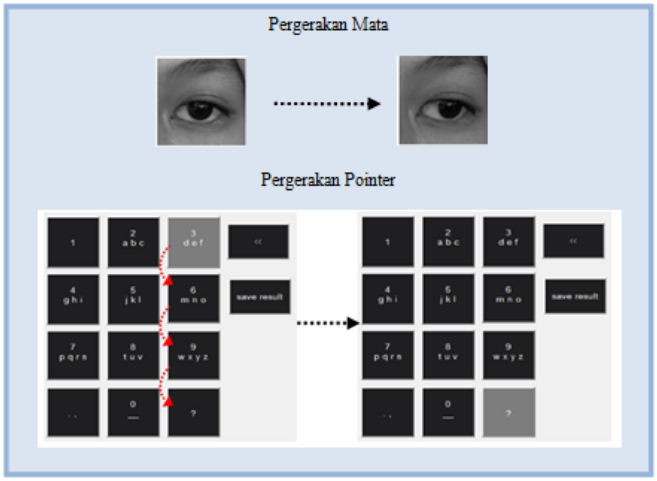

Fig.7. Eye Movement and Vertical Pointer Movement

In the vertical movement mode, the location of pointer is just in the one button, not in one column of button anymore. So, when user see at right, the pointer will be moved to downward and pointed in a button below the previous button. After the user see the normal position, the pointer movement will be stopped. If the pointer is pointing a button, user can choose a letter in that button by a quick blinking.

If user want to choose a button that is in the other column, user should changes the mode first by blinking within minimal 0.5 second to changes mode to be horizontal movement.

VII. EYE BLINK DETECTION AND CLASSIFICATION

The process of eye blink detection based on the whether the black area (iris) is detected or not. Because, when the eye is blinking, the iris is disappear. So, eye image when blink or not can be distinguished easily. Based in the threshold value, the contour whether non blinked eye and blinked eye is different. The contour of non blinked eye is almost circle, but the contour of blinked eye is various.The image of blinked eye contour is shown in Fig.8.


Fig.8. Not blinked eye contour


(a) (b) (c)

Fig.9. Blinked eye contour. (a) no contour. (b) small dots. (c) long contour

Based on those contour result, to differentiate between blinked eye or not is using comparison of large area and long area of the contour.

The detected eye blink is classified to be two, there are spontaneous eye blink and non spontaneous eye blink. Non spontaneous eye blink has duration more than 250 ms. This type of eye blink will be classified to be two, there are eye blink to change the mode of pointer movement, and click action to choose the letter in button. This eye blink classification is shown in the Table 1.

TABLE 1 Classification of Eye Blink

Eye Blink Classification	Duration (second)	Action
Spontaneous eye blink	≤ 0.25 second	-
Non spontaneous eye blink	0.26second – 0.5second	Choose the letter in button (click)
	≥ 0.5 second	Choose the mode of pointer movement

VIII. EXPERIMENTAL SETUP

In this system, the eye detection process is very important step that can affect the result of the next process. Table 2 will shows the optimum distance of eye detection process.

TABLE 2 Examination of eye detection process

User	Distance (cm)				
	40	45	50	60	70
1	√	√	√	X	X
2	√	√	√	X	X
3	√	√	√	X	X
4	√	√	√	X	X
5	√	√	√	X	X

Based on the table 8.1, the optimum distance between user's face and webcam is 40 cm- 50 cm. When the distance is 60 cm – 70 cm, the user's eye is not detected.

The result of iris segmentation process can affect the stability of eye movement position. So, the iris should segmented very well. The right adjustment of threshold value and morphology size can generate a good result of iris segmentation. Table 3 will shows the experiment of adjustment threshold value and morphology size in various lightning condition. Normal lightning has luminosity 30 lx, bright lightning has 81 lx, and dim lightning has 8 lx.

TABLE 3
Examination of iris segmentation with different threshold value

Lightning condition	Threshold	Morphsize	Segmented Iris
Normal	5	3	Not success
Normal	5	7	Not success
Normal	7	3	Not success
Normal	7	7	Success
Normal	11	3	Not success
Normal	11	7	Success
Bright	5	3	Not success
Bright	11	3	Not success
Bright	28	3	Not success
Bright	43	3	Not success
Bright	43	7	Success
Dim	3	4	Success
Dim	5	1	Not success
Dim	5	7	Not success
Dim	11	7	Not success

Based on the result of table 3, the conclusion is the threshold value that is too low for normal lightning and bright lightning can remove all the black eye area or iris, so the iris is not segmented clearly. And for the dim lightning, low threshold value is needed to get the black eye area.

The examination process of this entire system started from the first step, there are running program, do initialization, until typing process. This examination is to compute the average duration of typing process. The duration that is needed for each person to type some word is different. In the Table 4 and 5 will be shows how much the duration that is needed by user typing a same word, that is "cuci baju". After that, the average duration will compared with duration of writing normally. Based on examination process, the duration of writing normally for one word is 4 second, and for two words is 10 second. This is the equation to compute the comparation ratio :

$$Rasio = Xa : Xb \quad \text{.....(3)}$$

Where :
Xa = duration of writing normally
Xb = duration of typing using system

TABLE 4 Typing one word

User	Duration (second)
1	113 second
2	100 second
3	49 second
4	62 second
5	67 second
6	44 second
7	58 second
8	44 second
Average	**67 second**
Ratio	**1 : 13**

TABLE 5 Typing two words

User	Duration (second)
1	160 second
2	145 second
3	125 second
4	130 second
5	140 second
6	126 second
7	128 second
8	110 second
Average	**133 second**
Ratio	**1 :13**

Based on the result that is showed in table 4 and 5, the conclusion is the average duration that is needed by user to type one word in less than one minute, and the duration to type two word is less than two minute. But, when those value is compared with duration of writing normally, the ratio is 1:13 second. So, the duration of typing using system is longer. But in that duration, user and especially for someone who can't write normally, can typing a simple sentence that is quite clear.

IX. CONCLUSION

Based on the result of this research, the conclusion are :

1. The process to get eye gaze position is find the center point of iris from aye image that has been processed using threshold.
2. Eye gaze position of user is used as a reference for pointer movement, which is user should minimize movements to make the position always stable and does not fluctuate

3. Spontaneous eye-blink and non spontaneous eye-blink has different duration. The duration of spontaneous eye-blink is less than 0.25 second
4. The comparation ratio between normal writing duration and typing using this system is 1:13 second. And during that duration, user can typing a simple sentence that is quite clear.

REFERENCE

[1] Rafael, S., Nuno S., Pedro M., Arnaldo Abrantes. 2013. *Eye Gaze as a Human-Computer Interface*. Elsevier, Portugal.

[2] Krolak, A., Strumillo, Pawel. 2011. *Eye-blink Detection System for Human-Computer Interaction*. Springer, Poland.

[3] Ciesla, Michal.,Koziol, Przemyslaw. *Eye Pupil Location Using Webcam.*Krakow, Poland

[4] Simini, Franco.,Touya, Andres., Senatore, Agustin., Pereira, Jose. 2011. *Gaze Tracker by Electrooculography (EOG) on a Head-Band.* Uruguay.

[5] Keil, Andrea., Albuquerque, Georgia., Berger, Kai., Magnor, Marcus. 2010. *Real Time Gaze Tracking with a Consumer-Grade Video Camera*. Germany.

[6] Alnita, Nikmatuz Zahroh. 2010. Pengendali Senjata dengan Head Gaze menggunakan Metode Eigenface dan Iris Gaze menggunakan Metode Hough Circle Transform.PoliteknikElektronikaNegeri Surabaya, Surabaya.

[7] Risa Indah. 2009. Pengendali Pointer dengan Gaze Tracking menggunakan Metode *Haar cascade* sebagai Alat Bantu Presentasi. Politeknik Elektronika Negeri Surabaya, Surabaya

[8] Arai, Kohei., Mardiyanto, Ronny. 2011. Eye-Based HCI with Full Specification of Mouse and Keyboard Using Pupil Knowledge in The Gaze Estimation. Springer, Berlin.

[9] Arai, Kohei., Mardiyanto, Ronny. 2010. Improvement of Gaze Estimation Robustness Using Pupil Knowledge

[10] Arai, Kohei., Mardiyanto, Ronny. 2011. Eye-Based HCI with Moving Keyboard for Reducing Fatigue Effects.