

# Eye-based Human Computer Interaction Allowing Phoning, Reading E-Book/E-Comic/E-Learning, Internet Browsing, and TV Information Extraction

Kohei Arai

Dept. of Information Science  
Saga University  
Saga City, Japan

Ronny Mardiyanto

Dept. of Information Science  
Saga University  
Saga City, Japan

**Abstract**—Eye-based Human-Computer Interaction: HCI system which allows phoning, reading e-book/e-comic/e-learning, internet browsing, and TV information extraction is proposed for handicap student in E-Learning Application. The conventional eye-based HCI applications are facing problems on accuracy and process speed. We develop new interfaces for improving key-in accuracy and process speed of eye-based key-in for E-Learning application, in particular. We propose eye-based HCI by utilizing camera mounted glasses for gaze estimation. We use the sight for controlling the user interface such as navigation of e-comic/e-book/e-learning contents, phoning, internet browsing, and TV information extraction. We develop interfaces including standard interface navigator with five keys, single line of moving keyboard, and multi line of moving keyboard in order to allow the aforementioned functions without burdening the accuracy. The experimental results show the proposed system does work the aforementioned functions in a real time basis.

**Keywords**-Eye-based HCI; E-Learning; Interface; Keyboard.

## I. INTRODUCTION

Recently, the development of eye-based human computer interaction is growing rapidly. This grow is influenced by the growing of the number paraplegics. The number of paraplegics extremely increased (It was reported that in 2009 the number of paraplegics in U.S.A. has gained up 40% from 2007) caused by accident working (28%), motor vehicle accident (24%), sporting accident (16%), fall (9%), victim of violence (4%), birth defect (3%), natural disaster (1%), and others [1].

Nowadays, the eye-based Human-Computer Interaction: HCI has been widely used to assist not only handicap person but also for normal person. In handicap person, especially paraplegic, they use eye-based HCI for helping them to self-sufficient in the daily life such as input text to computer [2], communication aids [3], controlling wheelchair [4] [5], having meal on table using robot arm [6], etc. The eye key-in system has been developed by many researchers [2]. The commercial available system provided by Tobii Tracker Company<sup>1</sup> has been used by many researchers for developing text input, customer interest estimator on business market, etc [7].

Technology has been successful at rehabilitating paraplegics' personal lives. Prof. Stephen Hawking<sup>2</sup>, who was diagnosed with Amyotrophic lateral sclerosis<sup>3</sup> (ALS), uses an electronic voice synthesizer to help him communicate with others [8]. By typing the text through aid of a predictive text entry system, approximating his voice, he is able to make coherent speech and present at conferences. To give another example, a paraplegic patient wearing a head-mounted camera is able to draw figures, lines, and play computer games [2]. Clearly, through use of assistive technology, handicapped people are able to do feats on par with non-handicapped people.

The published papers discussing eye-based HCI system are categorized into: (1) vision-based and (2) bio-potential-based. The vision-based method utilized camera to capture image and estimate the user sight. The key issue here is how the method/system could be deal with environment changing. Lighting changing, user movement, various types of user, etc have to cooperate with the system. The vision-based system could be explained as follows,

- 1) Ref. [9] developed eye mouse based on user's gaze. After face is found and tracked, eye location is searched by projection of difference between left and right eye images. Output of this system is only left and right direction which used to control mouse pointer. No upward and downward directions are used. It has been implemented to control application such "BlockEscape"<sup>4</sup> game and spelling program.
- 2) Ref. [10] developed eye mouse which user's gaze is obtained from pupil location by using Haar Classifier (OpenCv function<sup>5</sup>). Also, blinking is used as left click mouse event.
- 3) Ref. [11] developed camera mouse using face detection and eye blink. Center position of face is detected by using Adaboost<sup>6</sup> face detection method and tracked by using

<sup>2</sup> [http://en.wikipedia.org/wiki/Stephen\\_Hawking](http://en.wikipedia.org/wiki/Stephen_Hawking)

<sup>3</sup> [http://en.wikipedia.org/wiki/Amyotrophic\\_lateral\\_sclerosis](http://en.wikipedia.org/wiki/Amyotrophic_lateral_sclerosis)

<sup>4</sup> BlockEscape

<sup>5</sup> <http://sourceforge.net/projects/opencvlibrary/>

<sup>6</sup> <http://note.sonots.com/SciSoftware/haartraining.html>

<sup>1</sup> <http://www.tobii.com/>

Lucas-Kanade method<sup>7</sup>. This location is used as pointing value and blinking is used as left click mouse event.

- 4) Ref. [12] developed a human-computer interface by integrating eye and head position monitoring devices. The system was controlled based on user sight and blinking. The user command could be translated by system via sight and blinking. Also, they modified calibration method for reducing visual angle between center of target and the intersection point (derived by sight). It was reported that this modification could allowed 108 or more command blocks to be displayed on 14 inch monitor. Also, it was reported that it has hit rate of 98% when viewed at the distance of 500mm. For triggering, the blinking was used to invoke commands.

The bio-potential-based method estimated user behavior (eye behavior) by measuring user's bio-potential. The bio-potential measurement instrument is required for measuring eye behaviors. The example of bio-potential-based has been applied into application of electric wheelchair controlled using Electrooculograph (EOG)<sup>8</sup> analyzed user eye movement via electrodes directly on the eye to obtain horizontal and vertical eye-muscle activity. Signal recognition analyzed Omni directional eye movement patterns [13].

In this paper, we propose eye-based HCI allowing phoning, reading E-Book, E-Leaning and E-Comic, and TV information extraction. E-Book, E-Learning, E-Comic contents can be accessible through Internet. The proposed Eye Based Tablet PC: EBTPC allows read the contents [14]. One segmentation TV signal can be acquired with tuner. Sometime users would like to get information for purchasing products introduced from the TV program. The conventional system need human resources to extract such information for purchasing the products then create sales product database for consumers. They used to sell access fees for getting information. The proposed TV information extraction allows users to extract information automatically [15]. The objective of this research is how we could use to replace the use of touch screen that always rely on hand. The use of touch screen to input a command has been widely used in many applications. The use of it is still limited only for normal person who could input a command by touching the screen directly using hand. Unfortunately, the handicap person will not be able to use it since he could not use his hand to input a command via touch screen like the normal person. In this research, besides allowing the use of it for handicap person, it should improve the response time of typing since the sight is faster than hand control. If we input a command using hand, it could be fast if hand have recognized the location of the key, unfortunately the actual speed rely on the distance between key. If the bigger size of keyboard is used, the hands-typing speed will decrease (It happens if the distance among the keys is farther than the finger covered area). For the condition with distance among keys is farther, the sight will be faster than the hands. In this research, we propose eye-based HCI by utilizing camera mounted on user glasses to estimate the user sight. By fixing the users head position, we estimate the sight of user to

display. We use the sight detection result to input command such as navigate E-Comic/E-Book/E-Learning reader, call a phone number for phoning, and TV information extraction.

## II. PROPOSED METHOD

In this paper, we propose a new system of eye-based HCI allowing phoning, browsing internet, reading E-book/E-Comic, and TV information extraction. Such system will help handicap person using mobile E-Learning system.

The mobile E-Learning system utilizing mobile phone or smart phone for accessing E-Learning content from server have to be prepared also for handicap student. The handicap student who has difficulty to use hands will face problems using E-Learning system. Beside the application for E-Learning system, we also prepare it for making call (allowing user type phone number and making phone call), internet browsing, E-Comic/E-Book/E-Learning content reader, and TV information extraction. In this system, we design user interface that will be explained as follows,

### A. Proposed User Interface

**Error! Reference source not found.** shows the interface's flow of our system. The startup menu will show buttons consisting of Phone Dial Pad, E-book/E-Comic Reader, Internet browsing, and TV Information extraction. These buttons provide different functionality that could be choosing by user. User chooses one of this menu buttons to enter the sub menu. The main menu is shown in **Error! Reference source not found.**

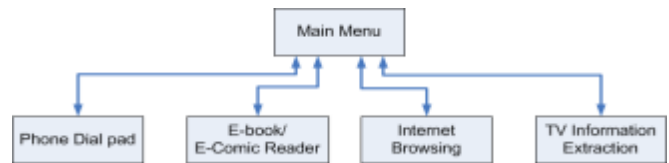


Figure 1 User Interface



Figure 2 Main menu

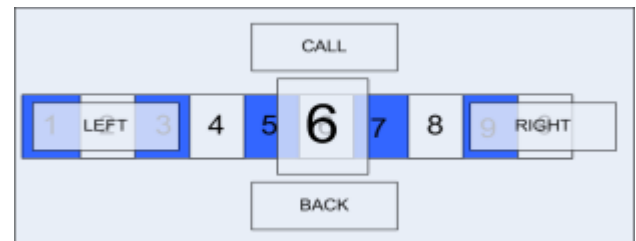


Figure 3 Phone dial pad sub menu

It consists of four buttons placed in top (Phone dial pad), down (TV), right side (Internet Browsing), and left sides (Read E-Comic/E-Book). To select the button, user have to look at it and hold within few seconds (using timer mode) or

<sup>7</sup> [http://en.wikipedia.org/wiki/Lucas%E2%80%93Kanade\\_method](http://en.wikipedia.org/wiki/Lucas%E2%80%93Kanade_method)

<sup>8</sup> <http://en.wikipedia.org/wiki/Electrooculography>

blinking (using blinking mode) for execution. In our system, we maintain the number of button is five (top, down, left, right, and center) for maintaining selection accuracy (As we know that the increasing the number of key/button will decrease the accuracy). Also, we design these buttons with same distance among them for making all buttons become a button with same characteristic with others.

After user select the button on main menu, the sub menu of selected button will be shown. If user selects the phone dial pad button, the interface such in **Error! Reference source not found.** will appears.

It contains four buttons with single line moving keyboard. The four buttons are used to move the moving keyboard to left or right, call the selected phone number, and return to main menu via "BACK" button. The single line moving keyboard consist of the number characters and symbol that as used in usual phone dial pad. We only use single line moving keyboard because the number of character for phone dial pad is few, so it does not need multi line moving keyboard. User could select the phone number by navigating the left and right button to move the single line moving keyboard. User have to locate the candidate of selected number to center by using these two navigator buttons ("LEFT" and "RIGHT"). To locate the candidate of selected number to center (for instance "4"), user could look at "LEFT" within 2 steps (if the initial condition is "6") until the "4" moves to center (the number located in center will be shown in bigger size to help user distinguish it easily).

The other sub menu is E-book/E-Comic reader as shown in **Error! Reference source not found.** This sub menu consist of four buttons: "SELECT" for selecting the title of E-book/E-Comic, "BACK" for return to main menu, "PREVIOUS" to go to previous page, and "NEXT" to go to next page of opened content. Before user could read the content, user have to select the title of the E-book/E-Comic by navigating "PREVIOUS" and "NEXT" button. After the desired title is shown, user opens it by selecting the "SELECT" button and the content of selected file will be opened and shown on display.

The sub menu for internet browsing is shown in Figure 5. This sub menu will allow user surfing around the world through web site. User could use our interface for browsing internet by utilizing his eye only. First, user input the URL address via moving keyboard navigated using four buttons: "UP" is for moving the layout go to upward, "DOWN" is for moving the layout go to downward, "LEFT" is for moving the layout go to leftward, and "RIGHT" is for moving the layout go to rightward. After the URL address is input by user, the web page will be shown on bottom part of our interface.

The last sub menu is TV Information Extraction is shown in **Error! Reference source not found.** It will allow user extract information from Digital TV (Usually used to extract information such as schedule, price of advertising item, sub title, etc). To extract the information, user could navigate our interface using four buttons: "LEFT" and "RIGHT" are for changing the type of information, "BACK" is for returning to main menu, and "EXTRACT" is for executing the TV Information follows the type of information.

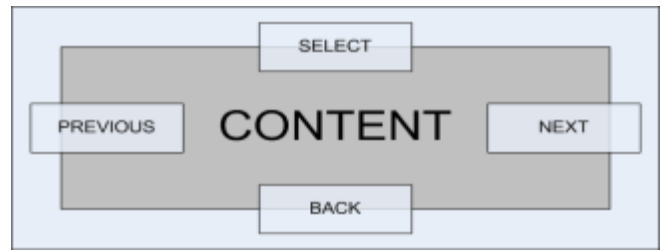


Figure 4 Sub menu of E-book/E-comic/E-Learning content reader

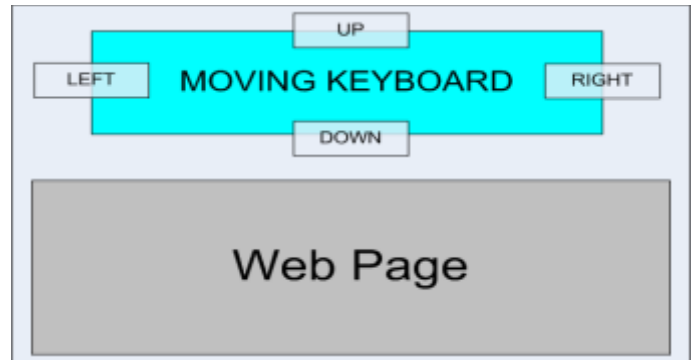


Figure 5 Sub menu of Internet browsing

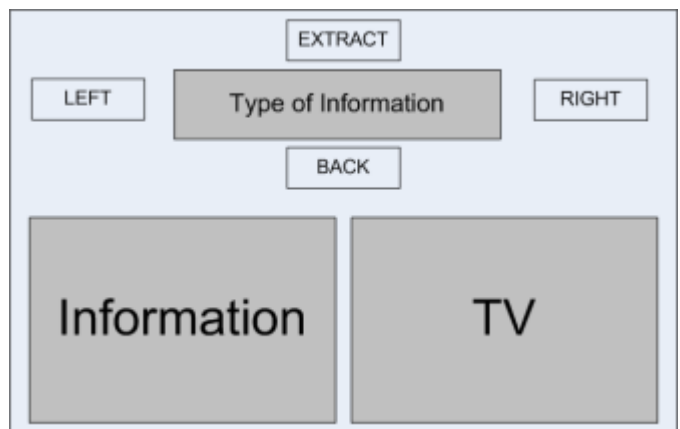


Figure 6 Sub menu of TV information extraction

### B. Implementation

We implement our system by utilizing Infrared: IR Camera, NetCowBoy DC NCR-131 mounted on user glasses to acquire user image. We modified the position of 7 IR LED for adjusting illumination and obtaining stable image even illumination of environment changes as shown in Figure 7.

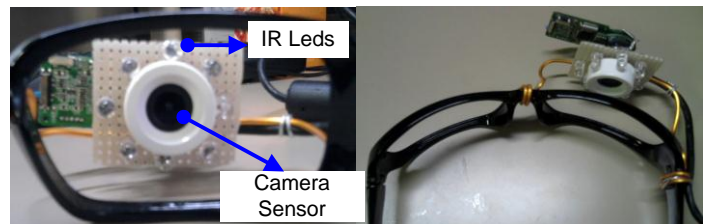


Figure 7 Modified camera sensor

Our software is developed under C++ language of Visual Studio 2005 and OpenCv, image processing library, which can be downloaded as free at their website. The advantages of this

camera mounted on glasses have been explored in ref [4]. It was success to minimize problems such as vibration, illumination changes, head movements, etc.

In this system, we search pupil location on an eye image by using our method that has been published in the reference [16]. We estimate the sight by converting the obtained position of pupil to sight angle. After the sight angle is estimated, we control the mouse cursor by using this sight.

The use of typical web camera has merit in low cost and easy to make, unfortunately it has demerit in noise, flicker, low resolution, etc. These demerits influence the stability of our system. Also, the various type of user's eyelash, deformable phenomenon of eye shape due to eye movement, existence of light source, etc often influence our sight result become unstable.

To solve this stability problem, there are many approach such as improving hardware stability, filtering, etc. In this system, we solve this problem by developing interface allowing user type characters; navigate an E-book/E-Comic reader, etc. We maintain the typing accuracy by developing the interfaces that have been explained previously.

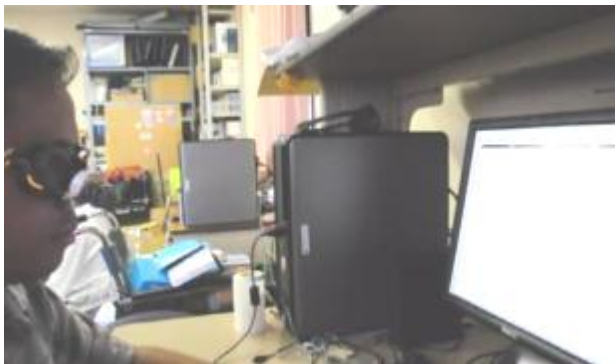


Figure 8 Use of the proposed system

To use our system, user just wears the glasses sensor in front of display as shown in Figure 8. Before user uses our system, the calibration step should be passed first (because we only use single camera). The calibration step will synchronize eye trajectory on image (which is acquired by camera) with the sight trajectory on the display. The eye trajectory in image has different pattern compared with eye trajectory in display. Due to the difference of camera placement, the sight estimated result may have nonlinear output with display as shown in Figure 9.

If the camera placement is not in center of pupil exactly, it means the plane of camera is not in a line between center of pupil and center of display, we have to transform the trajectory output of camera to trajectory of display. The calibration points that resulted from calibration step influenced by the different camera placement are shown in Figure 10.

From Figure 10 we can see that between target points on

display has different pattern to calibration points on camera image. We can see that the calibration points is little bit nonlinear and symmetry. To solve this non linearity problem, we use perspective transformations to transform the nonlinear eye estimated location to trajectory of display as shown in Figure 11.

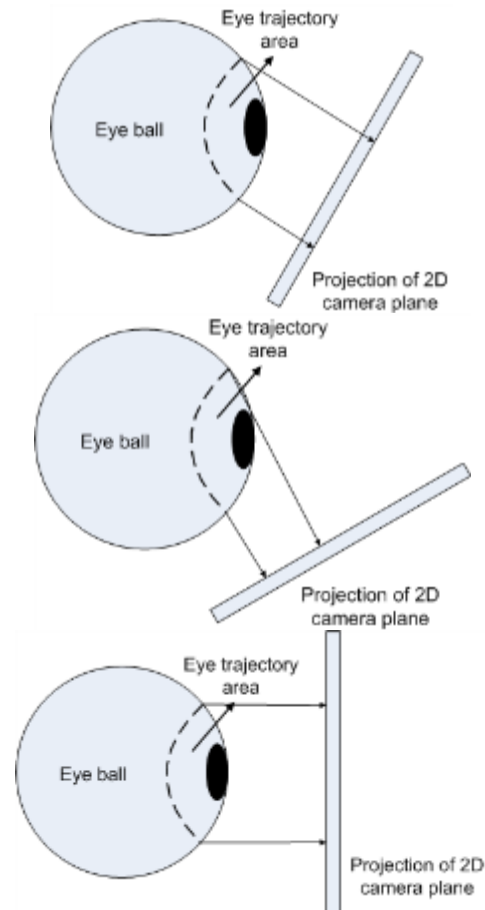


Figure 9 Phenomena of different camera placement to the relation between eye trajectories to display-trajectory area

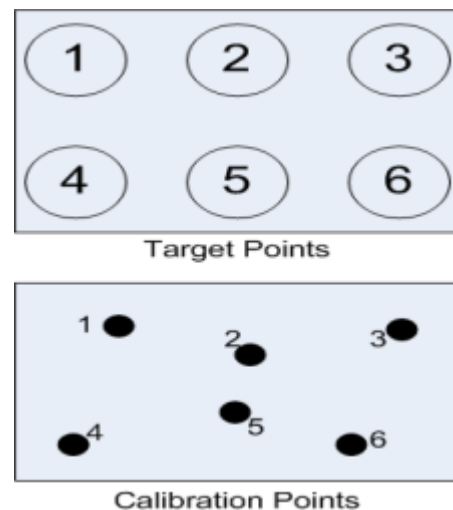


Figure 10 Effect of different camera placements



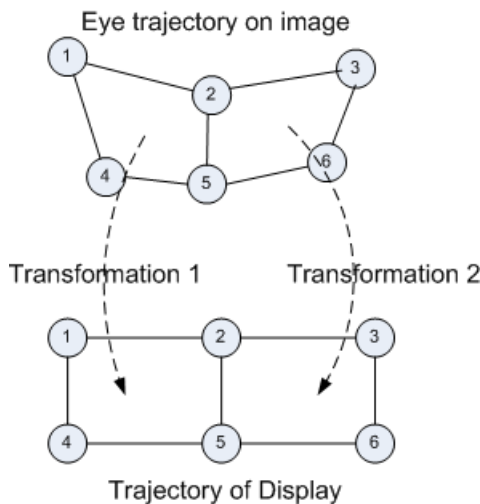


Figure 11 Transformation from eye trajectory to display trajectory

### III. EXPERIMENTAL RESULTS

To measure the effectiveness of our proposed method, we tested its performance by conducting accuracy experiment for eye detection and the sight.

#### A. Eye Detection Accuracy

The perfect performance of eye detection method is mandatory in every eye-based HCI system. It determines accuracy in early step. The next process only will gain the result. This experiment involved six users with different nationality (The difference of nationality identical with various eye shapes, various skin colors of eye, etc). The TABLE 1 shows the accuracy performance of our method compared with two other methods: Adaptive Threshold<sup>9</sup> and Template Matching<sup>10</sup>. The result shows that our method is superior with accuracy of 96.73% and could maintain accuracy against different user by variance is 16.27%.

Also, we test our method to light changing. We give adjustable lighting to system and measure the effect. The result shows that our method could maintain the accuracy from 0 Lx until about 1500 Lx. It means that our method does work without any light because we have IR LED to adjust the illumination. The proposed method failed if the light of environment is more than 2000 Lx (direct sun light).

TABLE 1. EYE DETECTION ACCURACY

User Types	Nationality	Adaptive Threshold (%)	Template Matching (%)	Proposed Method (%)
1	Indonesian	99.85	63.04	99.99
2	Indonesian	80.24	76.95	96.41
3	Sri Lankan	87.8	52.17	96.01
4	Indonesian	96.26	74.49	99.77
5	Japanese	83.49	89.1	89.25
6	Vietnamese	98.77	64.74	98.95
<b>Average</b>		<b>91.07</b>	<b>70.08</b>	<b>96.73</b>
<b>Variance</b>		<b>69.75</b>	<b>165.38</b>	<b>16.27</b>

<sup>9</sup> <http://homepages.inf.ed.ac.uk/rbf/HIPR2/adpthrsh.htm>

<sup>10</sup> [http://en.wikipedia.org/wiki/Template\\_matching](http://en.wikipedia.org/wiki/Template_matching)

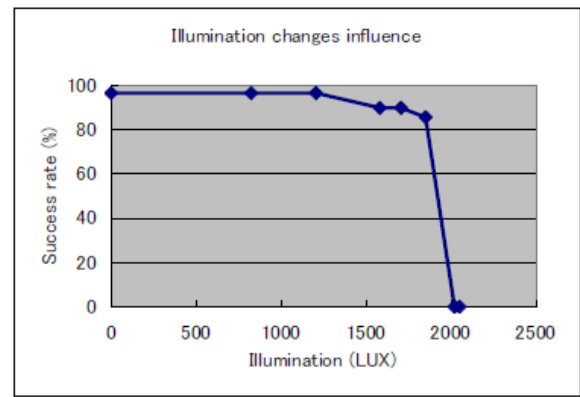


Figure 12 Influence of light changing to eye detection accuracy

#### B. Stability of Sight Estimated Point

The sight stability is measured to know the radius of sight error. This radius determines the maximum number of key that still could be used. The bigger radius of sight error causes the maximum number of key become decreases. Otherwise, the minimum radius of sight error could elevate the maximum number of key. In this experiment, user was looking at the target point and it moved serially on six locations. The Figure 13 is shown the sight stability. It shows that on key 5, the radius of sight error is high compared with other keys (it was caused by light source disturbed the eye detection method).

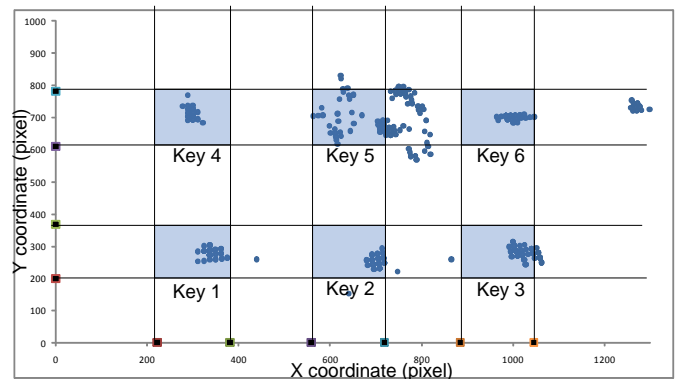


Figure 13 Sight Stability

#### C. Effect of High Number of Key to Accuracy

This experiment measured the effect of high number of key to accuracy. We conduct the experiment by modifying the number of key and measuring the accuracy. We start our experiment by using the number of key is 2; continue with 3, 5, 7, 9, and 15. The effect of number key to accuracy is shown in Figure 14. It shows that the raise of number of key will increase the error (decrease the accuracy). Also, we made a simulation that could figure the relation of accuracy to the distance among keys. This relation could be draw to a model as shown in Figure 15, with the assumptions are sight instability follow circle distribution (non parameter) with the radius is R, the distance among key is AB, and the error is represented by the slices among the circles.

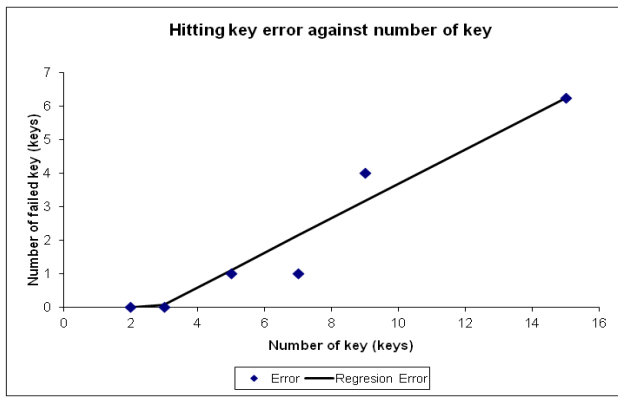


Figure 14 Effect of number of keys to accuracy

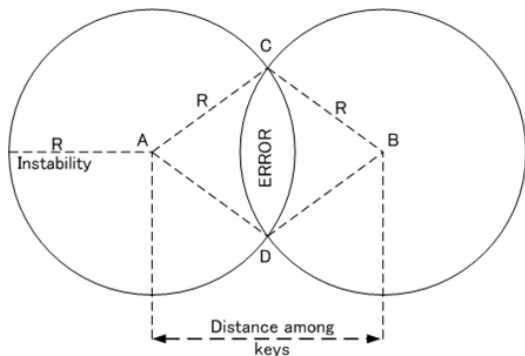


Figure 15 Model of relation accuracy to distance among keys

The result of simulation is shown in Figure 16. It shows that the approach of distance among keys causes the accuracy decreases. Otherwise, the widening of distance among keys causes the accuracy become maximal.

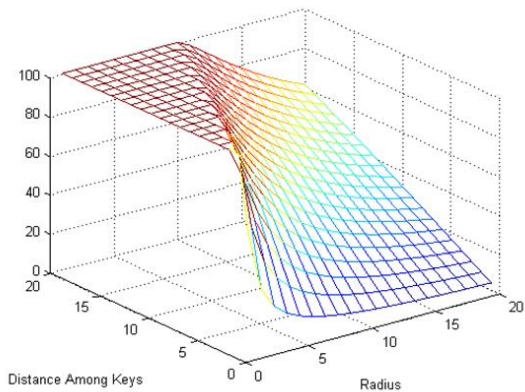


Figure 16 Simulation result of relation accuracy to distance among keys

#### D. Key-in Accuracy and Process Speed

In the last experiment, we conducted the experiment for measuring typing accuracy and speed. In this system, we use several model interfaces including standard navigator interface with five key, single line of moving keyboard, and multi-line of moving keyboard. In this experiment, we measured the typing accuracy when user was using multi line of moving keyboard and also recorded the typing speed.

In this experiment, we involved six users including two beginner users and four expert users. The beginner user means they ever used our system less than ten times. Otherwise, the expert user ever used our system more than ten times. How much time they ever used our system determines their expertise level. We compared our result to Fixed Keyboard Model as shown in TABLE 2.

TABLE 2 THE ACCURACY OF MULTI LINE OF MOVING KEYBOARD

User	Expertise	Moving Keyboard (%)	Fixed Keyboard (%)
1	Expert	100.00	92.86
2	Expert	100.00	76.19
3	Beginner	82.14	71.43
4	Expert	100.00	85.71
5	Beginner	71.43	78.57
6	Expert	100.00	66.67
<b>Average</b>		<b>92.26</b>	<b>78.57</b>

The experiment result shows that our system has better accuracy compared with fixed keyboard model. It was caused by our key was bigger than the fixed keyboard. Also, we only used five keys to navigate the moving keyboard while the used fixed keyboard in this experiment has thirty keys. According to our simulation result in Figure 16 that the higher number of key will has worse accuracy. Because our method used lower number of key, it causes our accuracy better than the fixed keyboard model.

Beside the measurement of typing accuracy, we measured the typing speed. By using same methodology of typing accuracy experiment, we recorded the typing speed. The experiment result is shown in TABLE 3. It shows that our method is faster than the fixed keyboard model because the use of smaller key (in fixed keyboard model) made user become difficult to input a character and it made the typing speed become slower. Otherwise, our method used bigger key and it made user still possible input a character easily. The result shows that our method is faster with the typing speed of 134.69 seconds while the fixed keyboard has slower typing speed of 210.28 seconds.

TABLE 3 TYPING SPEED

User	Expertise	Moving Keyboard (s)	Fixed Keyboard (s)
1	Expert	117.50	154.00
2	Expert	138.67	195.33
3	Beginner	180.50	275.00
4	Expert	101.00	197.33
5	Beginner	161.50	213.00
6	Expert	109.00	227.00
<b>Average</b>		<b>134.69</b>	<b>210.28</b>

#### E. All the Functionalities

All the functionalities, phoning, reading E-Book, E-Learning, E-Comic contents, Internet browsing, watching TV and the required information for purchasing products extraction from TV commercial are confirmed.

#### IV. CONCLUSION

It is concluded that the proposed eye-based HCI system works well for selecting and determining five keys for navigation of functionalities. Expertise is required for perfect key-in. In other words, 100% of key-in success rate can be achieved through exercises of using eye-based HCI system. Comparative study between the conventional fixed keyboard and the proposed moving keyboard shows that key-in speed of the proposed system is much faster than that of the conventional system by around 35%. All the functionalities, phoning, internet browsing, reading E-book/E-Comic/E-learning contents, and TV information extraction are confirmed. These functions are available in a real time basis.

#### ACKNOWLEDGMENT

Authors would like to thank to the graduate students who contributes to the performance evaluation experiments of the proposed system.

#### REFERENCES

- [1] FOUNDATION, Christopher and Dana Reeve. Disponivel em: <<http://www.christopherreeve.org>>. Acesso em: 7 October 2011.
- [2] EYEWRIter: low-cost, open-source eye-based drawing system. Disponivel em: <<http://www.crunchgear.com/2009/08/25/%20eyewriter-low-cost-%20open-source-eye-%20based-drawing-system/>>. Acesso em: 7 October 2011.
- [3] LEFF, R. B.; LEFF, A. N. 4.954.083, 1990.
- [4] Arai, K.; R. Mardiyanto, A Prototype of ElectricWheelchair Controlled by Eye-Only for Paralyzed User. *Journal of Robotic and Mechatronic*, 23, 1, 66-75, 2011.
- [5] Djoko P., R.Mardiyanto, and K.Arai, Electric wheel chair control with gaze detection and eye blinking. *Artificial Life and Robotics, AROB Journal*. 14, 694,397-400, 2009.
- [6] Arai K; K. Yajima, Robot Arm Utilized Having Meal Support System Based on Computer Input by Human Eyes Only. *International Journal of Human Computer Interaction (IJHCI)*, 2, 1, 120-128, 2011.
- [7] EYE tracking and eye Control. Disponivel em: <<http://www.tobii.com/>>. Acesso em: 8 out. 2011.
- [8] OFFICIAL website of Professor Stephen W. Disponivel em: <<http://www.hawking.org.uk>>.
- [9] John J. M. et al. EyeKeys: A Real-Time Vision Interface Based on Gaze Detection from a Low-Grade Video Camera. 2004 Conference on Computer Vision and Pattern Recognition Workshop. 159, 2004.
- [10] Changzheng L., K. Chung-Kyue, P.Jong-Seung, The Indirect Keyboard Control System by Using the Gaze Tracing Based on Haar Classifier in OpenCV. *Proceedings of the 2009 International Forum on Information Technology and Applications*. 362-366, 2009.
- [11] Zhu H, L. Qianwei, Vision-Based Interface: Using Face and Eye Blinking Tracking with Camera. *Proceedings of the 2008 Second International Symposium on Intelligent Information Technology Application*. 306-310, 2008.
- [12] Parks K.S., L. K. T. Eye-controlled human/computer interface using the line-of-sight and the intentional blink. *Computers and Industrial Engineering*. 463-473, 1993.
- [13] Barea, R. et al. System for Assisted Mobility using Eye Movements based on Electrooculography. *IEEE Transaction on Neural System and Rehabilitation Engineering*, 10, 209-218, 2002.
- [14] Arai K. and T. Herman, Automatic e-comic content adaptation, *International Journal of Ubiquitous Computing*, 1,1,1-11,2010.
- [15] Arai K., and T. Herman, Method for extracting product information from TV commercial, *International Journal of Advanced Computer Science and Applications, Special Issue on Artificial Intelligence*, 2, 8, 125-131, 2011.
- [16] Arai, K.; R.Mardiyanto, Improvement of gaze estimation robustness using pupil knowledge. *Proceedings of the International Conference on Computational Science and Its Applications (ICCSA2010)*. 336-350, 2010.

#### AUTHORS PROFILE

**Kohei Arai** received a PhD from Nihon University in 1982. He was subsequently appointed to the University of Tokyo, CCRS, and the Japan Aerospace Exploration Agency. He was appointed professor at Saga University in 1990. He is also an adjunct professor at the University of Arizona since 1998 and is Vice Chairman of ICSU/COSPAR Commission A since 2008