

# Development of a User Interaction System that Presents Relevant Information Based on Gaze Line

Takahide Otomo<sup>1</sup>, Shinya Mochiduki<sup>1</sup>, Eriko Ishii<sup>2</sup>  
Yuko Hoshino<sup>1</sup>, Mitsuho Yamada<sup>1</sup>

<sup>1</sup>Tokai University 2-3-23 Takanawa Minato-ku Tokyo Japan

<sup>2</sup> Kagoshima Prefectural College 1-52-1 Shimo-Ishiki Kagoshima-shi Kagoshima Japan

Keywords: Gaze, Eye movement, User interface, Web

## ABSTRACT

We developed a new user interaction system using non-contact eye tracking device. This is a system that extracts and uses words that are estimated to be of interest to the user from a web page, based on the gaze point and gazing time automatically. This article describes this system.

## 1 BACKGROUND

The development of information and communication technology has brought many benefits to our society. In particular, browsing web pages is an important way to utilize and obtain information and we browse web pages of various designs in daily life. There are over 1.7 billion web sites in the world [1] also, the age group of internet user has expanded with the spread of smartphones. Hence, it can be said that a web page which are easy to be accepted is an important theme for many users. In addition, the five-sense interface and sensing technology are also progressing thus, a user interaction system that can help user's operations has a great meaning. In recent years, cost reduction of gaze input devices is progressing. Therefore, even a general user can easily use high-performance devices. Conventionally, a gaze input device has been very expensive. For this reason, studies of using gaze input for general application or system have not progressed much. And the searching engines using voice recognition have various types, it can be said that voice recognition has spread widely general users. On the other hand, there is not many effective use method or system for gaze input. This time, we focused on gaze input devices and attempt to develop a system that acquires interest words from a browsing page and uses them for user interaction.

## 2 EYE MOVEMENT

Eye movements are driven by extraocular muscles and they can be sorted into binocular or monocular type. Moreover, monocular eye movements can be classified as fixation eye movement, pursuit eye movement and saccades [2-3]. The eye movement when we read a sentence is realized by repeating saccades and temporary stopping [4]. Therefore, gazing and saccade are important for reading a sentence but also, web browsing. Fig. 1 shows the gaze line and the mouse movement, while user is selecting a content in Tokai University official web page.

The mouse movement is shown with the blue line. Also, the red and green line represents movement of gaze. Especially, the green line indicates the saccade for changing gazing contents.



Fig. 1 Image of gaze line and the mouse movement, while user is selecting a content

The gaze jumped with saccade (Green line), the yellow arrows and number represents the order of movement. These data were acquired with Eye Tribe Tracker

## 3 GAZE AND WEB BROWSING

A gaze information is often used for evaluation of usability. There are various examples of usability evaluation of web pages based on gaze or computer operation. It has been reported that the gaze retention time becomes longer at parts (point) of interest in the previous researches that examined browsing and gaze time [5]. In addition, Toda et al. reported that gazing time while subject is searching information is shorter than gazing time when a viewer judges its information is match the purpose or not [6]. Nakamichi et al. investigated the relationship between Web pages with low usability and viewer's behavior. As a result, they reported that the amount of mouse scrolling can be an indicator for detecting a low usability page [7].

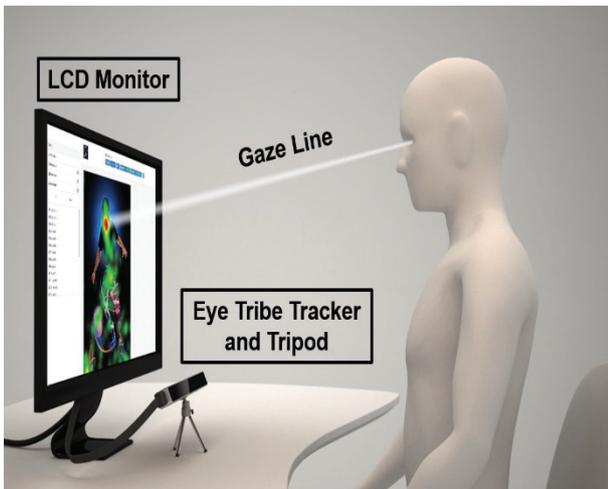
## 4 WEB PAGE AND HTML

HTML stands for Hypertext Markup Language, which is the standard language for writing a hypertext. Position of elements such as images and text in many web pages

are specified by this. A web page is composed of multiple html tags, and a page creator assign each content to a necessary element, making it a page that can be read by a web browser. For example, "div" tag is used as a block of document to be displayed on the screen. There is also a research example that uses HTML to acquire the content arrangement of a web page, and examines how the subject is browsing based on eye movements [8].

## 5 INTERACTION SYSTEM OVERVIEW

We used Eye Tribe Tracker (The Eye Tribe) that can make various programs using gaze input as a contactless gaze measurement device. The sampling rate was 60Hz and the device can fix on a tripod. Fig 2 shows image when using Eye Tribe Tracker. This time we targeted the most used web browser chrome in the world. This system acquires the user's gaze line position when reading text on a Web page, also detects a temporary stopping of gaze. If the stopping time exceeds a threshold, the word is determined as a user's object of interest. Thereafter, the acquired word is used for user interaction for displaying recommended information on the screen. These functions are performed automatically thus, user operation is not required. This user interaction is realized by two different programs, the one is gazing position detection program and the other is word extraction program. The former acquires user's gaze line position by screen pixel unit and generates a mouse click event on its coordinate value, when interest word was determined. The latter performs extraction processing. After a mouse click event occurs, a word on a Web page under the coordinate value is extracted from the sentence, and an interaction processing is performed on a Web browser.



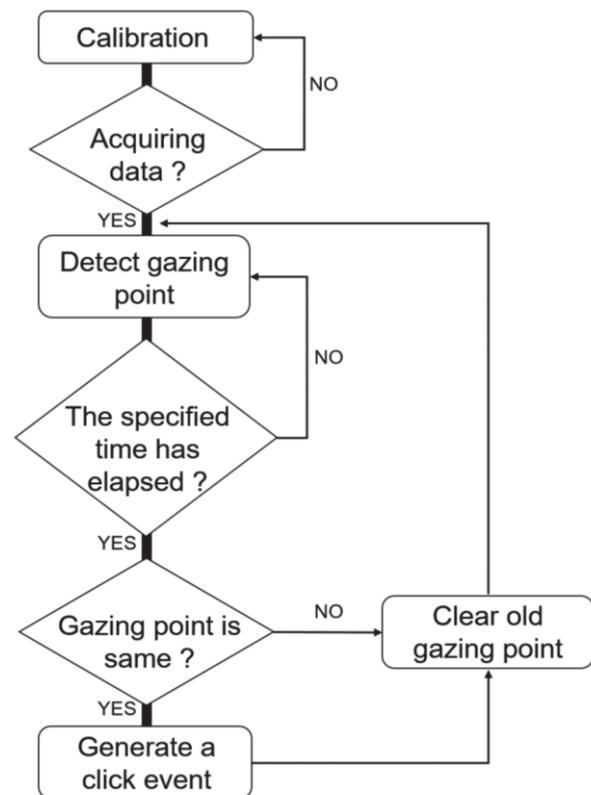
**Fig. 2 Image when using Eye Tribe Tracker**

User's gaze line to the display was acquired using Eye Tribe Tracker. (Image: <https://theeyetribe.com/dev.theeyetribe.com/dev.theeyetribe.com/general/index.html>)

### 5.1 Gazing position detection program

Gazing position detection program for acquiring user's

gaze data and recorded it, the program was developed with C#. The program calculates a temporary stopping point of gaze with screen coordinate values from Eye Tribe Tracker and Fig.3 shows Flowchart of the program. Time of temporary stopping is not always constant. Hence, it is considered that user's interest part or word can be detected by measuring of gazing time (time of temporary stopping). This time, we set threshold of velocity and time threshold. The threshold of velocity was set 2 deg [9] per 1 sample, if user's velocity of eye movement exceeds this value, the program judged that user changed gazing point, also subsequent data was set another gazing point data. Then, the time threshold is used to determine the gazing point with high interest. We set the threshold of gazing time was set 500 msec. Hence, if user kept gazing at same point more than 500 msec, the program generates a mouse double click event to a computer. The click event is performed based on the coordinate value of the gaze position acquired by Eye Tribe Tracker. Also, this program has a function that can visualize the user's gaze line position. Using this function, the user's gaze is drawn on the screen with a circular index and Fig.4 shows image of screen when using this function.



**Fig. 3 Flowchart of Gazing position detection program**

This program detects the gazing point and measures gazing time. Also, generates a click event. The process is repeated until program is finished. And this program can close at any time.



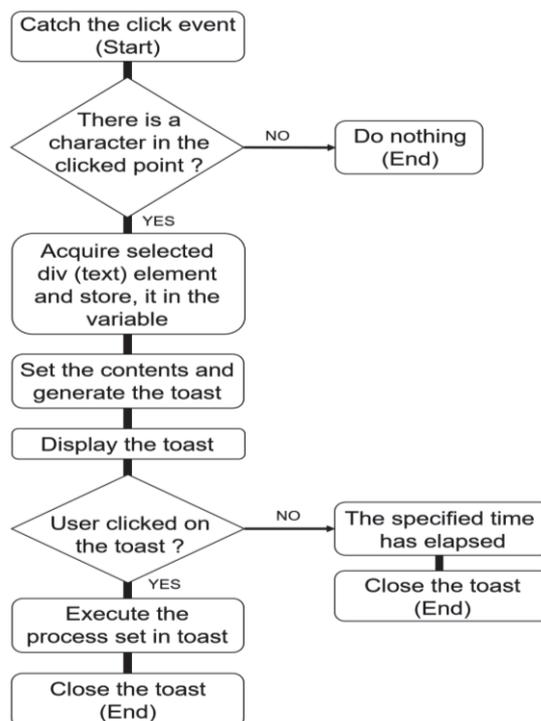
**Fig. 4 Image of screen when gaze line is superimposed and displayed on web page**  
The gaze position is indicated by this green circle.

### 5.2 Word extraction program

This program was developed with JavaScript and it was run on the web browser using a chrome extension which called Tampermonkey. We can run own JavaScript program by using this. The program performs function of word extraction after catching a double click event. Fig.5 shows Flowchart of the program. The reason of using double click is, it can easily select range in units of English words. The program performed selected word extraction on a point where occurred double click. Concretely, this script acquires character elements set by the div tag within a HTML under the coordinate value. After the extraction, the extracted word was recorded and use it for user interaction. This time, we created a function of displaying toast-type notification as an interaction for presenting information. (toast is a function of application that temporarily pops up a small message window for information notifications. It is also called a notification banner and this function is mainly used for app notifications, and when clicked, detailed information is displayed.) This notification function was created by adding two JavaScript libraries jQuery and Toastr to the source code in Tampermonkey. The content of the toast and its display time can be set freely in JavaScript. Fig. 6 shows the image of the toast used. This time, the selected word and the link to the page of the result of searching its word with google were displayed. The link was set to change with the selected word using "href" tag and variable.

## 6 SYSTEM VERIFICATION TEST

We verify the system can acquire a user's gazing word and information during web browsing. The distance between the subject and the monitor was 60 cm, but the head was not fixed with a chin rest. A verify was performed with one subject, the subject browsed the news page of Tokai University. The system configuration is shown in Table 1.



**Fig. 5 Flowchart of Word extraction program**

Start processing after the click event. After a prescribed period of time elapses the toast is closed.



**Fig. 6 Image of the toast**

This toast was set that it can open the specified page by clicking it and the green bar represents the display time. (upper image: remaining long time)

**Table. 1**

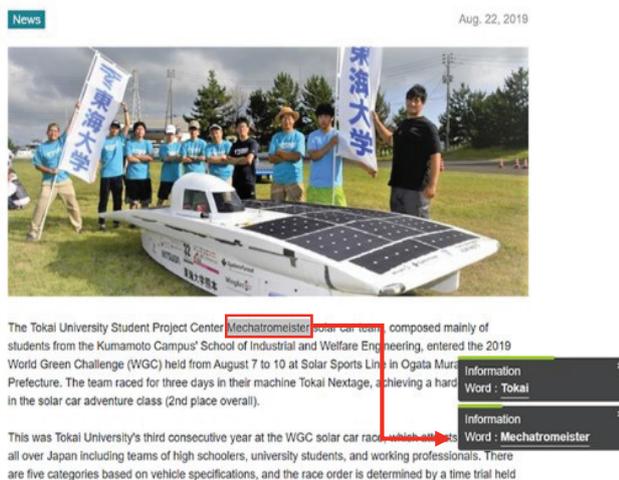
PC	Frontier FRXN770/C Intel® Core™ i7-7700HQ Windows 10
LCD Monitor	KA240H (1920×1080) (Acer Incorporated)
Web Browser	Google Chrome (Google LLC)

### 6.1 Results

The result image when the system is running shown in Fig.7. It was confirmed that English words whose gazing time was longer than the threshold could be extracted from the text in the page, also the presented link worked normally. In the figure, two words "Tokai" and "Mechatromeister" are extracted by the system and are presented its information on the screen. The number of notifications by toast can be set arbitrarily, and this time,

up to 3 can be displayed. About the gazing words, the subject was directed to gaze "Tokai" by experimenter to confirm system running. Thus, the extraction of "Tokai" is not caused by natural browsing. On the other hand, Mechatromeister is a proper noun hence, subject's attention was high. We need to investigate the setting of a suitable threshold for extracting the gazing word. In addition, it was confirmed that a character size of 17 points or more (viewing distance is 60 cm), which is larger than the accuracy of the device, is required for proper extraction.

### Tokai Team Wins Class at 2019 World Green Challenge



**Fig. 6 The processing procedure**  
Image when two words "Tokai" and "Mechatromeister" are displayed on the screen by the system.  
(Red squares and lines were added by the edit)

## 7 SUMMARY

We developed a user interaction system that can obtain user's interest information with a non-contact eye tracking device and confirmed that the system can extract the word of gaze position and display it for the user interaction. Currently, simple word detection by double-clicking is performed, but in the future, in order to support Japanese, we will add a morphological analysis function. In addition, we plan to create a page title analyze function (e.g. when the browsing page title is delicious food, add a positive element as "deliciousness" to word of interaction). Various low-cost gaze input devices for consumers such as Tobii eye tracker were sold already. Therefore, the importance of the gaze input interface is increasing. Hence, to realize a user friendly interaction system using eye tracker which can help web browsing has great meaning.

## ACKNOWLEDGMENTS

This work was supported by JSPS KAKENHI Grant Number 17K02129.

## REFERENCES

- [1] <https://news.netcraft.com/archives/2018/05/29/may-2018-web-server-survey.html>
- [2] K.Ukai: "Eye Movement: Characteristics and Method of Measurement", Japanese journal of optics publication of the Optical Society of Japan, 23(1), pp. 2-8, 1994 (in Japanese).
- [3] S.Martinez, S.L.Macknik and D.H.Hubel: "The Role of Fixational Eye Movements in Visual Perception", NATURE REVIEWS NEUROSCIENCE, 5, pp. 229-240, 2004.
- [4] T.Kondo, R.Reiko and K.Takehi: "Effects of Lexical Properties and Word Order on Japanese Sentence Processing", Cognitive Studies, 9(4), pp.543-563, 2002 (in Japanese).
- [5] A.Fukasawa, M.Omata and A.Imamiya: "Identification of User's Interested Areas based on Eye Fixation Duration and BVP Height during Web Browsing", IEICE technical report, 107(462), pp1-6, 2008.
- [6] K.Toda, N.Nakamichi, K.Shima, M.Ohira, M.Sakai and K.Matsumoto "An Information Exploration Model based on Eye Movements during Browsing Web Pages", IPSJ SIG technical reports, 52, pp.35-42, 2005 (in Japanese).
- [7] N.Nakamichi, M.Sakai, K.Shima, K.Matsumoto "Detecting Low Usability Web Pages using Data of User's Behavior", The transactions of Human Interface Society, 8(1), 2006 (in Japanese).
- [8] T.Matsunobe "Gaze Analysis Method for Web Site Usability Evaluation", The transactions of Human Interface Society, 16(4), 2014 (in Japanese).
- [9] M.Yamada, T.Fukuda: "Definition of gazing point for picture analysis and its applications", The IEICE Transactions, D. 69(9), pp.1335-1342, 1986 (in Japanese).
- [10] <https://theeyetribe.com/theeyetribe.com/about/index.html>