

Development of a Tourist Information Search Support System That Reflects Interests Based on User's Gaze

Taiga Mori¹, Yuko Hoshino¹, Mitsuho Yamada¹ and Eriko Ishii²

¹Tokai University, Tokyo, Japan

²Kagoshima Prefectural College, Kagoshima, Japan

1cjinm016@mail.u-tokai.ac.jp

hoshino@tokai.ac.jp

myamada@tokai.ac.jp

erikoishii@k-kentan.ac.jp

Abstract: In 2014, Tobii Tech and Eye Tribe released the first inexpensive non-contact eye tracking device, and eye tracking technology has been attracting more and more attention. It has long been said that gaze information can be used to infer a person's interests and attention, and the analysis of gaze position by eye tracking technology is also used for Search Engine Optimization (SEO). In addition, with the spread of information and communication technology, the Internet has become the most used media for "obtaining information on hobbies and entertainment," surpassing books, magazines, and television for all age groups except those in their 60s in Japan. The majority of information gathering has shifted to web browsing. As a result, the information gap between those who can use the Internet to collect information and those who cannot use it well, so-called digital divide, has become one of the social problems. Based on these background, we attempted to develop an interaction system that presents search keyword to get related information by using the user's eyes while browsing the web. Based on analysis of user's gaze and accumulation of web pages viewed by each user, this system can provide search keywords that reflect a user's interest, such as Tourism information or history information. We believe that this system can support users' web browsing by assisting them in gathering information on sightseeing and travel, presenting them with unexpected facilities that they did not know about, and making it easy for them to obtain a variety of related information.

Keywords: user's gaze, eye tracking, web browsing, user's interests, keyword recommendation, related information

1. Introduction

It has long been believed that gaze information can be used to infer a person's interests and attention, and it is now being used in a variety of fields. For example, in the field of automobile driving, Panasonic (2022) presented an in-vehicle HUD (Head-Up Display) with a gaze-tracking function at CES2022, a trade fair for the latest technologies. The HUD can automatically adjust its focus according to the driver's gaze and can detect whether the driver is drowsy or not, thus improving driving safety. In the field of sports, it is said that effective education can be achieved by comparing the gaze information of skilled athletes and amateurs, and understanding the gaze behavior peculiar to skilled athletes. (Manabu, N. et al, 2016 and Takuya, S. et al, 2021) In addition, not only players' eyesight but also referees' eyesight has been analyzed to improve referees' skills (Kenichiro, M. et al, 2021). In 2014, Tobii Tech and Eye Tribe released the first inexpensive non-contact eye tracking device. The International Market According to a market study by the International Market Analysis Research and Consulting Group (IMARC), the global eye tracking market is expected to grow at a compound annual growth rate (CAGR) of approximately 26% between 2021 and 2026. The demand for the use of eye information is expected to increase further in the future. Thus, the analysis of gaze information is used in all fields and is a market that is expected to grow further in the future. (IMARC Services Private Limited,2021)

In recent years, with the spread of information and communication technology, the proportion of information gathering using the Internet, such as Web browsing and SNS, has been increasing. According to a report by the Ministry of Internal Affairs and Communications (MIC), the Internet will be the most used media for "obtaining information on hobbies and entertainment" as of FY2020, surpassing books, magazines, and TV for all age groups, and much of the information gathering is shifting to Web browsing (Ministry of Internal Affairs and Communications, 2021). As a result, the information gap between those who can use the Internet to collect information and those who cannot use it well, so-called the digital divide, has become one of the social problems. In addition, in recent years, however, the number of people who collect information from recommendations has been increasing, and the demand for information collection support by recommendation systems has been increasing.

Based on this background, we tried to develop an interaction system to present related information by using the user's eyes while browsing the Web. We believe that this system can support users' web browsing by helping them to collect information about sightseeing and travel, presenting unexpected information that users did not know, and allowing them to easily obtain various related information. We also believe that the system will be able to provide search support to some elderly and physically disabled people who have difficulty in gathering information from the Internet, because they can obtain additional information just by browsing.

2. Background

2.1 Support for information gathering based on keyword recommendations

Tourist often use web to gather information before departure and in the field. In addition, there have been attempts to extract keywords from browsing web pages to support information gathering. In Watanabe et al (2011) research, they extracted unique expressions in the text as keyword candidates, and applied keyword scoring using WebIDF (Masayuki, O. et al, 2011), which is a score calculated by how frequently words occur on the Web, to them. We are developing a search support UI that notifies and presents the top scoring keywords to users. Yamada et al (2020) focused on the information that exists in the web pages that users do not browse. They focused on the problem that users cannot obtain new and necessary knowledge efficiently by browsing randomly from the search results and proposed a search support method that presents keywords as extended snippets that can predict the contents of unvisited pages. Ono et al (2015) focus on the possibility that words that the user needs exist in the web pages of links that the user has not browsed in the search results and propose a method to recommend words that are suitable for the search intention in the pages that the user has not browsed as keywords.

2.2 Web browsing and gaze information

Umemoto et al (2013) attempted to estimate search intentions during real-time web browsing based on the assumption that words with higher attentional value are more suitable for users' search intentions. The authors developed an estimation method using tf-idf, which evaluates the importance of words in a document based on the spatial vector representation of words and the degree of word occurrence, and systematized it in the form of. Shobayashi et al (2018) studied a method for extracting users' interest words from their eye movements while reading Web news articles, and showed that it was possible to extract interest words under certain conditions, which could not be extracted using linguistic information alone. Toda et al (2005) investigated the relationship between information search and gazing time on a Web page, and found that the gazing point stops when a Web page visitor decides what information to obtain. In addition, they reported that the gaze pausing time was longer when the user judged whether the information was what he wanted than when he searched for information.

As above, there have been many researches to support users' web browsing by extracting keywords from the web. There are also many studies on methods to estimate users' interests based on gaze information during web browsing. On the other hand, there are only a few studies of practical systems for general consumers that support web browsing by combining keyword extraction from the web and gaze, such as the study by Otomo et al (2021) In this study, we focused on the use of gaze information in a search support system, and tried to develop an interaction system that presents keywords to obtain related information based on the user's gaze while browsing the Web.

3. Gaze input device

In this chapter, we describe the Tobii Pro Nano (Figure 1) (Tobii, 2021), a non-contact gaze input device used in this study. The device uses the Pupil Centre Corneal Reflection (PCCR) as an eye movement detection method. In the PCCR method, a light reflection point is generated on the cornea, and its image is captured by a camera, which identifies the light reflection point on the cornea and the pupil from the captured image of the eye. The direction of the eye is calculated based on the reflection point and other geometric features. The sampling rate is 60 Hz, and the system is capable of extracting the gaze every 16.7 msec.



Figure 1: Gaze input device (Tobii pro nano)

4. Interaction system overview

Figure 2 shows a flowchart of our Interaction system. This system consists of three processes:

1. text acquisition process
2. interest extraction process
3. related information presentation process

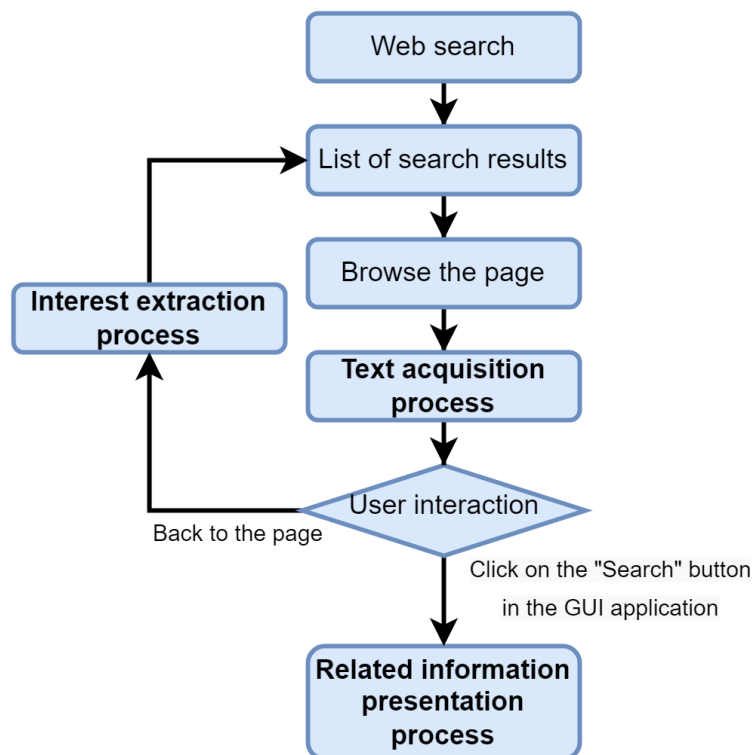


Figure 2: Flowchart of our interaction system

4.1 Text acquisition process

In this process, the system automatically retrieves text elements of the web page being browsed based on the user's gaze. Figure 3 shows the GUI application we have created. From this GUI application, the user browses a web page using Google Chrome, which is opened by Selenium. After that, the system extracts the user's gazing point from the Tobii Pro Nano and runs JavaScript based on the coordinates of the gazing point to retrieve the text in the html elements of the web page that the user is browsing, such as the h tag that contains the heading and the p tag that contains the text in the body of the page. Figure 4 shows the example of text extraction range. The range of the text extraction is chosen to be a sentence unit rather than a single character unit, because the later process also analyses the sentence unit. Furthermore, unread sentences are not extracted in order to extract sentences from eye coordinates. In addition, there is a possibility that sentences that are read in a stream are sentences that the user is not interested in. Therefore, we set a threshold value calculated from the average reading speed of Japanese sentences (Shinya, S., 2004 and Jumpei, K. et al, 2018) and the recommended font size of Web pages (IMJ Corporation, 2013), so that sentences judged to have been skimmed are not extracted.

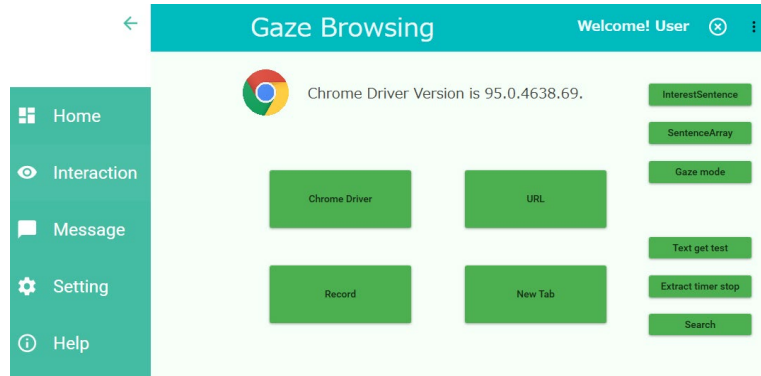


Figure 3: GUI application



Figure 4: Example of text extraction range

4.2 Interest extraction process

This process uses BERTSUM (Yang, L., 2019) to create and save the summary sentences of the browsed web pages as the user profile used to present the recommended keywords that reflect the user's interests in the related information presentation process. This process is performed on the sentences acquired by the text acquisition process when the user leaves the page being browsed, and each page is summarized into three sentences and saved. This process and the text acquisition process continue until the user clicks the "search" button on the GUI application, and the summary text of the browsed page is saved as an index to measure the user's interest.

4.3 Related information presentation process

This process is executed when the user presses the "search" button on the GUI application screen, and it recommends search keywords through three processes:

4. interest understanding process
5. important words extraction process
6. search keywords recommendation process

4.3.1 Interest understanding process

In this process, the system uses Sentence-BERT (Nils, R. et al, 2019) to compare the similarity between the text that was being browsed when the "search" button was clicked (hereinafter referred to as "browsing text") and the summary text that was saved in the preference extraction process (hereinafter referred to as "browsed

text"). Sentence-BERT generates sentence vectors by fine-tuning BERT (Jacob, D. et al, 2018) and calculates the similarity between documents by the cosine distance of the sentence vectors. The closer the cosine distance is to zero, the more similar the documents are to each other. This time, we calculate the distance between documents of browsing text and browsed text. Then, the sentences whose distance between documents is less than 0.33 are temporarily saved as the sentences suitable for the user's interests (hereinafter referred to as "interests sentences").

4.3.2 Important words extraction process

In this process, we use important word analysis by tf-idf, which is an evaluation index of important words, to obtain important words in "interests sentences". First, our system performs morphological analysis using MeCab on the "interests sentences" and obtain proper noun words. After that, we analyze the morphological analysis results using tf-idf, which is an evaluation index of important words. tf-idf estimates the importance of a word in a sentence by using the equations (1), (2), and (3) in Figure 5. This time, the system saved the top 10 important words calculated by tf-idf and used these words for the following process.

<p>(1) tf (Term Frequency)</p> $= \frac{\text{Frequency of occurrence of word X in sentence A}}{\text{Total of frequencies of all words in sentence A}}$ <p>(2) idf (Inverse Document Frequency)</p> $= \log\left(\frac{\text{Number of whole sentences}}{\text{Number of sentences containing word X}}\right)$ <p>(3) tf-idf = tf × idf</p>
--

Figure 5: Calculation formula of tf-idf

4.3.3 Search keywords recommendation process

We use spatial vector representations of words to evaluate the relationship between words. In this study, we use word2vec (Tomas, M. et al, 2013) to create a vector space model of words. Based on this model, we finally decide the search keywords to recommend to users. This time, we calculate the similarity of words using word2vec for 10 words stored in the important words extraction process, and identify 5 words whose vectors are close to each other. After that, the system sends a pop-up notice to the user using the five words as search keywords. When the user clicks on this pop-up notice, Google search is automatically performed with the recommended search keywords and the search result is displayed. In this way, it is possible to provide additional information related to the web page that the user browsed.

5. System verification and consideration

We have verified whether our system can provide additional information to users who are browsing the Web, reflecting their search intentions. As one of the results of our analysis, we present the results of browsing the homepage of Goryokaku, a special historical site in Hokkaido, Japan (hereafter referred to as the "verification page"). In the previous study (Taiga, M. et al, 2020) the keywords of different topics, such as tourism information and historical information, were presented together regardless of the search intention. Therefore, in this study, in order to verify whether the recommended keywords change appropriately according to the search intention, we assumed two information search behaviors: one is interested in tourism and the other is interested in history. Therefore, in the interest extraction process, three articles on tourism in Hokkaido, Japan, and three articles on history and war in Hokkaido, Japan, were stored in the system. Table 1 shows the list of search keywords suggested by the system.

Table1: List of search keywords suggested by the system

Keywords	Topic	
	tourism	historical
1	Goryokaku Park	Hokkaido Development Commissioner
2	Goryokaku Tower	Hakodate War
3	Hakodate City	Ayasaburo Takeda
4	Tatsuoka Castle	Tatsuoka Castle
5	Goryokaku	Western style

In the case of a search for tourist information, "Goryokaku Park" and "Goryokaku Tower" are suggested keywords. These are tourist spots located near Goryokaku. In addition, we were able to propose new additional information by suggesting as a recommended keyword "Tatsuoka Castle", a Goryokaku that exists in the central part of Japan. However, "Hakodate City" is not suitable as a recommended keyword because it is unlikely to be searched by the name of the municipality in the search task assuming the acquisition of tourism information. As for "Goryokaku", it cannot expect to obtain any new information on it because the verification page is about "Hakodate Goryokaku". In the case of historical information, it was confirmed that the keywords such as "Ayasaburo Takeda" and "Hakodate War" could be retrieved for the designer and related wars. However, "Western style" is not a suitable keyword because it is unlikely to be searched by itself. From the above results, we confirmed that it is possible to recommend search keywords suitable for users' search intentions by using gaze information and user profiles, although some unsuitable keywords were recommended.

Next, we describe the issues we found in this study. First of all, the current system is slow in processing time. It took about 14 seconds from the time when the user clicked the "Search" button to the time when the keyword was recommended. According to Nielsen Norman Group, the response time of an interaction system is reported to be less than one second without stopping the user's flow of thought. However, if we prioritize the processing time, the current recommendation accuracy will not be maintained, so we need to adjust the system continuously. The next problem is that our system can only recommend search keywords from the text being browsed. We need to improve our system in order to support users to get new information. As a suggestion for improvement, we propose to add collocations to the current keywords as new recommended keywords, so that users can obtain new information. Another problem is that the current system searches automatically with all the recommended keywords presented, and the user cannot select the keywords to search. The purpose of this specification is to reduce the time and effort of user's decision-making and operation, such as "deciding the keyword" and "pushing the search button". However, since not all the recommended keywords capture the user's browsing intention, unnecessary information may be presented. Therefore, we think that it is desirable to add a check box type GUI and to leave the decision-making part to the user.

6. Summary

In this study, we have developed an interaction system that extracts sentences based on the user's gaze while browsing the Web, and through various analyses, can recommend search keywords suitable for the search intention. By verification, we confirmed that the recommended keyword changes according to the search intention such as tourism topic and historic topic.

Eye tracking technology has been attracting attention in recent years because eye tracking information can be used to infer the interests of users, and it is being used in various fields. We believe that by using this technology to support web browsing for tourist destinations, we can solve the digital divide and easy to search for people with handicaps. In the future, we plan to improve the issues we found and to evaluate the system quantitatively. In addition, our system does not currently support non-Japanese web pages. Therefore, in the future, the system will be updated to support English web pages and will be improved to support the information gathering of tourist destinations around the world.

References

- IMARC Services Private Limited (2021) "Eye Tracking Market: Global Industry Trends, Share, Size, Growth, Opportunity and Forecast 2021-2026"
- IMJ Corporation (2013) "Site Usability Study on Tablet Devices -Ease of Tapping and Readability on Different Screen Sizes-" (in Japanese)
- Jacob, D. Ming-Wei, C. Kenton, L. (2018) "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding", arXiv:1810.04805v2
- Jakob Nielsen (2010), "Website Response Times"
- Jumpei, K. Toshio, K. (2018)"Eye movements that cause individual differences in reading speed of Japanese texts"(in Japanese), The Institute of Image Information and Television Engineers, Vol 72, No.10, pp. J154-J159.
- Jumpei, Y. Daisuke, K (2020) "A Search Assistance Method Using Extended Snippets for Predicting Unvisited Pages of Users" (in Japanese), Information Processing Society of Japan Interaction 2020,1B-43
- Kazutoshi. U. Takehiro, Y. Satoshi, N. *etal.* (2013) "Estimating Real-time Search Intent Based on Attention term Extraction from Eye Movements" (in Japanese), Journal of Information Processing Society of Japan, Vol6, No.3, pp.120-131
- Ken-ichiro, M. Yutaka, S. (2021)"Research of Kendo Referee's Viewpoints by Eye-mark Recorder: Through the Judging of "Shikaketeiku-waza""(in Japanese), Research journal of Budo, Vol.41, No.2, pp1-11
- Kentaro, O. Yuki, T. Makoto, O. *etal.* (2015) " Keyword recommendation from unread pages using words that co-occur with the characteristic words in the Web" (in Japanese), Information Processing Society of Japan Research Report, Vol2015-GN96 No.18
- Koji, T. Noboru, N. Kazuyuki, S. *etal.* (2005) "An Information Exploration Model based on Eye Movements during Browsing Web Pages"(in Japanese), Information Processing Society of Japan Research Report, Vol.52, No6, pp.35-42
- Manabu, N. Yoshimitsu, K. Yukihiko, A. *etal.* (2016) "Characteristics of gazing points in volleyball" (in Japanese), Volleyball research, Vol.18, No.1, pp12-18
- Masayuki, O. Nayuko, W. Sumi, O. (2011) " Information retrieval support for mobile devices using touch operation" (in Japanese), TOSHIBA review, Vol66, pp46-49
- Ministry of Internal Affairs and Communications (2021), Information and Communications White Paper (in Japanese), Trends in the Use of ICT Services
- Nayuko, W. Masayuki, O. Masaaki, K. *etal.* (2011) "Search assistance UI for mobile devices based on keyword extraction from browsing web pages"(in Japanese), Information Processing Society of Japan Interaction 2011
- Panasonic (2021)" Panasonic Automotive Unveils Driveable AR HUD 2.0 with Patented Eye-Tracking-System at CES2022", <<https://na.panasonic.com/us/news/panasonic-automotive-unveils-driveable-ar-hud-20-patented-eye-tracking-system-ces-2022>>, viewed 3 January 2022
- Shinya, S. (2004) "Speed Reading and Eye Movements"(in Japanese), The Japanese Psychology Society, Vol.23, No.1, pp64-69.
- Taiga, M. Takahide, O. Yusuke, N. *etal.* (2020) "Development of a Web Browsing Support System Using Gaze Information", The 27th International Display Workshop, INPp1-2, pp.953-956.
- Takahide, O. Shinya, M. Eriko, I. *etal.* (2021) "Development of system that can obtain texts on web page based on gazing point"(in Japanese), Journal of the Japan Personal Computer Application Technology Society, Vol.14, No.1, pp.36-41
- Takuya, S. Kanji, K. Mitsuho, Y. (2021)"Skill analysis of athletes using eye movement measurement device"(in Japanese), The Transactions of Human Interface Society, Vol.23, No.1, pp29-42
- Tobii, (2021) "Pro Nano product Description", <<https://www.tobii.com/siteassets/tobii-pro/product-descriptions/tobii-pro-nano-product-description.pdf?v=2.1>>, viewed 31 October 2021
- Tomas, M. Ilya, S. Kai, C. *etal.* (2013) "Distributed Representations of Words and Phrases and their Compositionality", NIPS2013: Proceedings of the 26th International Conference on Neural Information Processing Systems, Vol.2, pp.3111-3119
- Yang, L. (2019) "Fine-tune BERT for Extractive Summarization", arXiv:1903.10318
- Yuta, S. Shinya, F. (2018) "Interest extraction using gaze behavior during web news article reading" (in Japanese), Forum on Information Technology 2018, J-035, pp.289-290