Detecting Low Usability Web Pages using Quantitative Data of Users' Behavior

Noboru Nakamichi
Nara Institute of Science and Technology
Nara, Japan
noboru-n@is.naist.jp

Makoto Sakai
SRA Key Technology Laboratory, Inc.
Tokyo, Japan
sakai@sra.co.jp

Kazuyuki Shima
Hiroshima City University
Hiroshima, Japan
shima@computer.org

Ken-ichi Matsumoto
Nara Institute of Science and Technology
Nara, Japan
matumoto@is.naist.jp

ABSTRACT
The purpose of this research is to detect low usability web pages from the behavior of users, such as browsing time, mouse movement and eye movement. We experimented to investigate the relation between the quantitative data viewing behavior of users and web usability evaluation by subjects. We analyzed the data to detect low usability web pages using discriminant analysis. Low usability web pages, 94.4% (17pages / 18pages = detectable pages / low usability pages) were detectable from the moving speed of gazing points and the amount of wheel rolling of a mouse. Moreover, this detection reduced the number of web pages which should be evaluated by half (46% = 89 pages / 192 pages = detected pages / all pages).

Categories and Subject Descriptors
D.2.8 [**Metrics**]: Performance measures; H.1.2 [**User/Machine Systems**]: Human factors; H.5.2 [**User Interfaces (D.2.2, H.1.2, I.3.6)**]: Evaluation/methodology; H.5.4 [**Hypertext/Hypermedia**]: User issues;

General Terms
Measurement, Performance, Experimentation, Human Factors

Keywords
gazing point, eye information, web usability, evaluation, performance

1. INTRODUCTION
The usability of a Web site is so important that it can influence the amount of sales, because users are unwilling to read Web pages with low usability, such as having pages that are hard to operate or understand, or pages that react differently from expectations[4]. To create easy-to-use Web pages, an evaluation of usability is required[3]. Web usability evaluation is performed to mainly discover problems on a Website. And designers consider a re-design about discovered problems.

Among various methods for usability evaluation, usability testing [3] is widely used. Usability testing is a method of discovering problems based on operation of the application by subjects. It tends to discover serious problems leading to trouble, and other problems which an evaluator cannot discover[3]. However, analyzing recorded data such as a user's utterance data and VTR takes time. Therefore, methods of supporting evaluation using quantitative data aiming at efficient and objective evaluation are proposed. These include operation time, mouse movement, and eye movement as quantitative data about users' behavior used for supporting evaluation. Operation time includes browsing time, goal time of a task, and the time interval of operation. Mouse movement includes moving distance, click positions, and the amount of wheel rolling. Eye movement includes the movement locus, moving distance, and moving speed of gazing points.

The method of supporting evaluation by using the quantitative data of users' behavior requires precise measurement of timing and position which is impossible for general users. Conventionally, to support evaluation using quantitative data, a specialist analyzes users' behavior from quantitative data based on experience and knowledge. For example, when using the locus of gazing points, a specialist finds areas where gazing crosses and returns[5]. Then a specialist analyzes the user's behavior in those areas, and evaluates usability problems on a screen. This quantitative data of behavior of a user supports analysis by the specialist. However, there is little support for efficiency of evaluation work in this approach. Companies which change a website frequently need to improve the efficiency of evaluation. Therefore, they need to first detect low usability pages on a website and then quickly narrow down the objects of usability evaluation.

Quantitative data about users' behavior may support detection of low usability Web pages without special knowledge. We

ICSE'06, May 20–28, 2006, Shanghai, China.

hypothesized that characteristics of various quantitative data used for evaluation are related to Web usability. If low usability Web pages are detectable from quantitative data, the efficiency of usability evaluation will increase without additional new work. However, it was not clear which data would be effective in detection of low usability Web pages among various quantitative data.

We experimented to investigate the relation between quantitative data viewing users' behavior and web usability evaluation by subjects for detecting low usability Web pages. By quantitative analysis, we showed clearly that the moving speed of gazing points is effective in detecting low usability Web pages. Moreover low usability web pages, 94.4% (17pages / 18pages = detectable pages / low usability pages) were detectable with the moving speed of gazing points and the amount of wheel rolling of a mouse. In addition, this detection reduced the number of web pages which should be evaluated by half (46% = 89 pages / 192 pages = detected pages / all pages).

2. RELATED WORKS

In this chapter, we describe methods of supporting evaluation objectively using quantitative data such as browsing time, mouse movement and eye movement.

Paganelli [9]analyzed the execution situation of a task from a user's operation event recorded by Java script. He supported the analysis based on quantitative data, such as page reference time and task execution time. He analyzed the usability of a Web page based on task execution time totaled for every Web page. Okada [8]developed the GUITESTER which extracts a common operation pattern from two or more users' operation history. If the tool is used, the operation pattern for mistaken operation can be extracted. And when the moving distance of a mouse cursor is long and an operation time interval is long, they suggest a possibility that a screen layout is bad.

Mon-Chu Chen [1]showed clearly that there is a strong correlation between the position of a mouse cursor and a gazing point when users browse. By using a mouse cursor, they have reported being able to predict a part with a user's interest or being able to reason about an intention of a user. However, they do not confirm whether the part of low usability is detectable by measuring a motion of a mouse cursor.

Focusing on the human interface, Mori [5]analyzed the movement of gazing points and revised the prototype screen to improve the effectiveness of prototyping with screen design in the development of an information system. They showed that revising redundant movements of the gazing point improves the operation speed and the users' satisfaction of usability.

WebTracer [10] can collect the operation log of users on the Web pages. Collectable data include the information on users' sight line (the coordinates of the gazing point on the computer screen), operation log of a mouse, and the displayed screen images, together with their time information. The data collected by WebTracer characterize Web pages and have the possibility of being used for supporting the usability evaluation. However, the relation between such data and the problems in the Web usability was merely an example of the characteristics of the Web pages. Quantitative evaluation of the relation to the usability of Web pages was not done.

In these conventional researches, the specialist had discriminated usability only using certain quantitative data. However, the effectiveness of each quantitative data was not compared. In addition, these studies did not confirm whether quantitative data can be used as a measurement standard of the usability of a Web page.

3. EXPERIMENT OF USABILITY TESTING

In this chapter, to clarify the quantitative data which can discriminate a low usability Web page for a user, we describe the experiment which investigated the quantitative data and the evaluation result by the user for every Web page.

In the experiment, we recorded the quantitative data of the subject who imposed the task (information search in a certain site) using WebTracer. After the subject finished the task, we used the operation history and did the questionnaire and an interview. In the questionnaire, we asked a question about the usability of each Web page. At the interview, we asked a question about the situation at the time of browsing.

3.1 Quantitative Data of Users' Behavior

Browsing, mouse movement, and eye movement are the quantitative data about users' behavior mainly used for web usability evaluation. This experiment recorded the following six quantitative data for every Web page:

- Browsing time (sec): Time since the user begins to see a certain page, until it changes to another page
- Moving distance of mouse (pixel): Moving distance of the mouse cursor on a screen
- Moving speed of mouse (pixel/sec) = Moving distance of mouse / Browsing time
- Wheel rolling (Delta): the amount of wheel rolling of a mouse
- Moving distance of gazing points (pixel): Moving distance of gazing points on a screen
- Moving speed of gazing points (pixel/sec) = Moving distance of gazing points / Browsing time

Gazing point is the point at the intersection of the users' look with the target screen[1].

When objective information cannot be found, this causes low usability. In such a case, we consider that browsing time becomes long. In addition, since a subject needs to check at various parts, we consider that moving distance of gazing point becomes long and speed of gazing point becomes high. Since it is reported that a mouse has the position of a gazing point and strong correlation[1], we consider that moving distance of mouse also becomes long and moving speed of mouse also becomes high. When a Web page is not settled in the window of a browser and a subject cannot find the information in a window, so that a subject may use a wheel, we consider that the amount of wheel rolling of a mouse becomes high.

3.2 Experimental Environment

The experiment environment used by this research is as follows.

- Display: 21 inches (Viewable screen size: H30 x W40cm)
- Resolution: 1,024×768 pixels
- Distance from subject's face to display: approx. 50cm
- Device for measurement of sight line: NAC, EMR-NC (View angle: 0.28, resolution on the screen: approx. 2.4mm)
- Recording and playing of sight-line data: WebTracer (Sampling rate: 10 times per second)

WebTracer[10] is an environment for recording and analyzing the users' operations in Web pages. Collectable data include the information on users' sight line (the coordinates of the gazing point on the computer screen), operation log of a mouse, and the displayed screen images, together with their time information. The recorded data and the pictures of Web pages are integrated and displayed.

3.3 Subjects and Tasks

Subjects are 10 frequent users of the Internet. They have never visited the sites used in the experiment.

Firstly, we requested the subject to read news out in a certain portal site so that the subject adapts oneself to experiment environment. Next, we requested the subject to perform five tasks of looking for the starting salary of a master from the site of five companies, as a main experiment. Moreover, the order of the task was performed at random for every subject.

3.4 Experimental Procedure

We experimented with usability evaluation in the following procedures to five tasks. We recorded quantitative data for every Web page, and investigated a user's evaluation result.

Procedure 1: The Web page for an experiment linked to the top page of each company is displayed by a subject. And the experiment is started from the time of a subject clicking the link.

Procedure 2: While subjects are doing the tasks, several types of data regarding browsing operations such as the gazing points of subjects are recorded using WebTracer. We do not take any interruptive measures, such as asking questions to the subjects. A task is ended, when a subject reports himself as having found the starting salary.

Procedure 3: The Web pages that subjects visited are displayed. Those pages are listed on the operation log recorded in Procedure1. We requested the subject to choose the ease of use for every visited Web page from the following five levels.

1. hard to use
2. relatively hard to use
3. relatively easy to use
4. easy to use
5. don't know

We questioned the subject, to investigate in what kind of quantitative data the usability which a subject feels appears.

We count going to visit a Web page with a subject once as 1 page. And when two or more subjects visit the same Web page, it counts as another page. Moreover, when a subject visits the Web page same two or more times, it counts as another page.

Procedure 4: We reproduce the operation history recorded by WebTracer, and a subject checks all the visited Web pages. At that time, when the subject searched for the information on target, we interviewed the subjects about the situation of their search.

3.5 Experimental Result

We recorded the quantitative data for 275 pages which the subject visited. We were not able to record correctly about 75 pages of them. The cause is a frequent blink and head movement. Moreover, there were eight pages which the subject answered "don't know" about the usability of the Web page. We measured the quantitative data in 192 pages except these pages as a result of the experiment. The average values of each quantitative data are shown below.

- Browsing time: 12.9 (sec)
- Moving distance of mouse: 1179.4 (pixel)
- Moving speed of mouse: 110.2 (pixel/sec)
- Wheel rolling: 277.5 (Delta)
- Moving distance of gazing points: 4848.7 (pixel)
- Moving speed of gazing points: 387.7 (pixel/sec)

4. ANALYSIS and DISCUSSION

Following the answers in the experiment, we classified the cases into two types: the cases with low usability and the other cases. If the subject in a case considered a Web page "hard to use" then we regard this as a case with low usability. If the subject answered "relatively hard to use" or "relatively easy to use" or "easy to use" then we regard this as an other case.

4.1 t-Test of Low Usability Pages and Others

We postulated a difference between the quantitative data for "low usability" and "others" pages which would allow detection of "low usability" pages. For each type of data, we performed a statistical test with the hypothesis that there is a difference between the two classes of cases.

Table 1 shows the average values and standard deviations for each type of quantitative data, together with the results of the above statistical tests. The results of the statistical tests in Table 1 show that the mean of each type of operation data for the cases with low usability is statistically different from that for the cases with others. However, the moving distance of the gazing point and the moving speed of the gazing point have both significantly different variances and means.

Table 1. t-test of low usability pages and other pages every kind of quantitative data

Quantitative data of users' behavior for each pages	Evaluation result by subjects				t-test (significance probability P)
	Low Usability Pages (18pages)		Other Pages (174pages)		
	average	standard deviation	average	standard deviation	
Browsing time (sec)	17.7	12.8	12.5	11.5	0.06882
Moving distance of mouse (pixel)	1267.9	717.4	1170.3	1186.0	0.61434
Moving speed of mouse (pixel/sec)	95.6	70.3	111.7	79.3	0.40922
Wheel rolling (Delta)	606.7	995.9	246.2	592.4	0.14885
Moving distance of gazing points (pixel)	8743.3	5808.3	4445.8	3815.9	0.00628
Moving speed of gazing points (pixel/sec)	515.6	102.5	374.4	126.9	0.00001

Table 2. discriminant function for every kind of quantitative data

Quantitative data of users' behavior for each pages	discriminant coefficient	constant term	discriminant boundary
Browsing time (sec)	17.7	12.5	11.5
Moving distance of mouse (pixel)	1267.9	1170.3	1186.0
Moving speed of mouse (pixel/sec)	95.6	111.7	79.3
Wheel rolling (Delta)	606.7	246.2	592.4
Moving distance of gazing points (pixel)	8743.3	4445.8	3815.9
Moving speed of gazing points (pixel/sec)	515.6	374.4	126.9

Table 3. The result of discriminant analysis for every kind of quantitative data

Quantitative data of users' behavior for each pages	Power of test (1-b)		Type II error b		Type I error a		Power of test (1-a)	
	pages	%	pages	%	pages	%	pages	%
Browsing time (sec)	9	50.0	9	50.0	37	21.3	137	78.7
Moving distance of mouse (pixel)	8	44.4	10	55.6	57	32.8	117	67.2
Moving speed of mouse (pixel/sec)	11	61.1	7	38.9	99	56.9	75	43.1
Wheel rolling (Delta)	7	38.9	11	61.1	29	16.7	145	83.3
Moving distance of gazing points (pixel)	8	44.4	10	55.6	31	17.8	143	82.2
Moving speed of gazing points (pixel/sec)	14	77.8	4	22.2	45	25.9	129	74.1

Power of test (1-b): The evaluation results by subject are low usability pages. And the discrimination results are low usability pages.

Type II error b: The evaluation results by subject are low usability pages. But the discrimination results are others.

Type I error a: The evaluation results by subject are others. But the discrimination results are low usability pages.

Power of test (1-a): The evaluation results by subject are others. And the discrimination results are others.

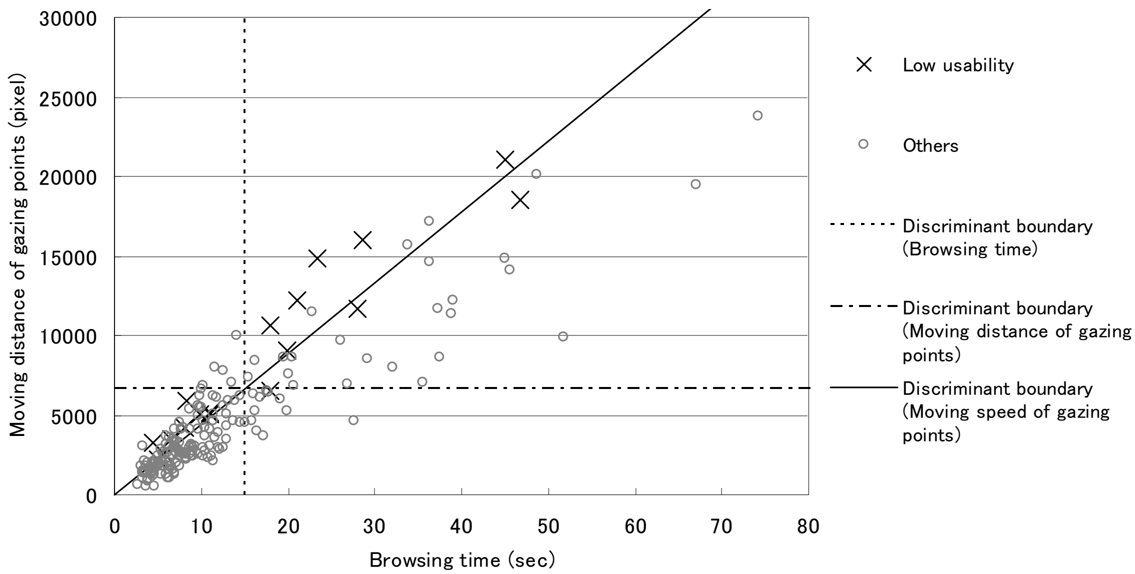

Figure 1. Scatter plot of low usability pages and other pages

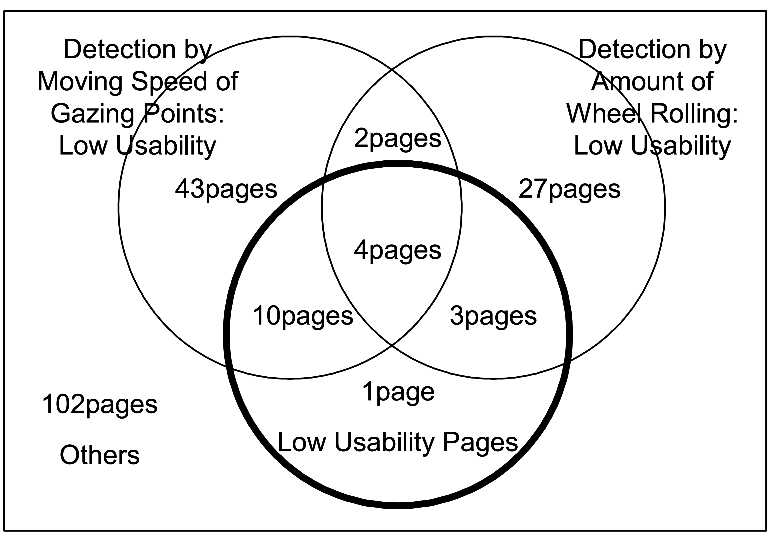

Figure 2. Venn diagram of the result of discriminant analysis for moving distance of gazing points and wheel running

This result showed clearly that the quantitative data measures that are effective in Web usability evaluation are the moving distance of the gazing point and the moving speed of the gazing point. In addition, on the low usability pages, we found that the moving distance of the gazing point was long, and the moving speed of the gazing point was high.

4.2 Discriminant Analysis of Low Usability Pages

To clarify which quantitative data can detect pages that are "hard to use," we analyzed the recorded data using discriminant analysis. The discriminant function for the quantitative data is shown in Table 2. Moreover, Table 3 shows the discriminant results of low usability pages and others. From Table 3, the quantitative data with the highest power of test (1-b) is the moving speed of the gazing point. 14 of 18 low usability pages were detectable (= 77.8%).

This result showed clearly that the quantitative data which can best detect low usability pages is the moving speed of the gazing point. We can detect low usability pages with 77.8% probability using the moving speed of the gazing point that are more than 445.0 (pixel/second). You can confirm this result by examining Fig. 1. Fig. 1 is a scatter plot with browsing time as the X-axis, the moving distance of the gazing point as the Y-axis, and low usability pages and others plotted on it.

4.3 Discrimination by Moving Speed of Gazing Points and the Amount of Wheel Rolling of a Mouse

To make the power of test (1-b) higher, we analyzed pages which were detected by the moving speed of the gazing point. We analyzed the behavior of subjects in Type II error b (4 pages) which were not detected by the moving speed of the gazing point. Specifically, there was a large amount of wheel rolling of the mouse on three of the four pages.

Based on this observation, we tried modifying the discriminant to include in low usability pages the case where the moving speed of the gazing point was higher than 445.0 (pixel/sec) and where the amount of wheel rolling was greater than 425.1 (Delta). The power of test (1-b) improved to 94.4% (17 of 18 pages) as a result of this discriminant analysis of low usability pages. The relation between the discriminant result by the moving speed of gazing point and the discriminant result by the amount of wheel rolling is shown in Fig. 2.

When low usability pages are discriminated the moving speed of the gazing point and the amount of wheel rolling, you can narrow down pages requiring Web usability evaluation by about half (89 of 192 pages = 46%) without evaluation by a specialist.

4.4 Improvement in Power of Test

We have shown that the power of test (1-b) improves when low usability pages are discriminated using the moving speed of the gazing point and the amount of wheel rolling. We analyzed the subject's behavior in the experiment, finding that while the subject was scrolling the Web page using the mouse wheel, the subject's gazing point moved very little. When subjects do not scroll a Web page to search for the objective information, the subject's gazing point goes to various parts of the Web page. However, while scrolling a Web page using a wheel, even if the gazing point focuses on one part of the screen, the contents of the Web page change with scrolling. The same action as moving the gazing point up and down while searching for the objective information is possible. Thus, if a subject uses a mouse wheel, even if they seldom move the gazing point, they can still search for the objective information on a Web page. On pages with large amounts of wheel rolling, the moving distance of the gazing point may be short even on low usability pages. Therefore, the moving speed of the gazing point becomes low, so that low usability pages were undetectable using only the moving speed of the gazing point. We found that the power of test (1-b) became higher by applying the discriminant result of wheel rolling along with the discriminant result of the moving speed of the gazing point.

4.5 Cause for False Discrimination

We analyzed the subjects' comments in an interview to clarify the subjects' situation on pages which could be discriminated as low usability. The following two kinds of comments were obtained in regard to the discriminated 17 pages.

"I couldn't easily find the link which leads to the objective information."

"I got lost because the menu layout is bad."

These two kind of comments both identify the same kind of situation, where subjects get lost because a link is not found.

Moreover, we analyzed the comment also on one page which was not able to be discriminated. The subject commented, "Although I found the link which leads to the objective information, since it was a mistake, I evaluated that it was low usability." On this Web page, when the link of "recruiting information" is clicked, there is a starting salary which is the goal of a task. However, the subject believes and clicks the link of a "welfare program." This page has

Table 4. The relation of the result of discriminant analysis for moving speed of gazing point and wheel rolling for each subject

Evaluation result by subjects	discriminant analysis for moving speed of gazing point	discriminant analysis for wheel running	pages	pages for each subject									
				A	B	C	D	E	F	G	H	I	J
low usability	low usability	low usability	4	0	0	0	0	0	1	0	1	2	0
low usability	low usability	others	10	0	0	0	0	7	0	0	0	0	3
low usability	others	low usability	3	1	0	1	0	0	1	0	0	0	0
low usability	others	others	1	0	0	0	0	1	0	0	0	0	0
others	low usability	low usability	2	0	0	1	0	0	1	0	0	0	0
others	low usability	others	43	3	7	5	0	8	2	2	2	2	12
others	others	low usability	27	4	0	2	13	0	4	0	1	3	0
others	others	others	102	10	15	8	27	9	7	8	12	6	0
summary			192	18	22	17	40	25	16	10	16	13	15

the Web usability problem that the anchor text does not describe the content of the linked page for the subject. However, the subject did not stray but clicked the link smoothly. We considered that this page was not able to be discriminated for that reason.

From these results, the pages which can be discriminated are pages where the subjects get lost. However, even if a page contains a Web usability problem, pages where subjects do not have trouble selecting a link cannot be discriminated.

4.6 Individual Difference of Quantitative Data

Table 4 shows the eight logical combinations of the user evaluation and discriminant analysis of the moving speed of the gazing point and the wheel rolling for each subject. The first three rows indicate those pages which are correctly identified as low usability. The fourth row contains the one page which the user considered low usability, but discriminant analysis of moving speed and wheel running did not correctly identify. This was the page with poor labeling of the links, and layout is poor. The fifth through seventh row contain pages which discriminant analysis incorrectly indicates may be low usability. While this is a large number (72 pages), the total which must be inspected is still less than half (90 pages) of the total pages.

In addition, the moving speed of the gazing point indicates individual differences in the difficulty of the task of searching for objective information. We consider this not as a capability difference but as a difference in similar experiences. It is experience based on whether the subjects have visited a site which was similar in the past.

Next, we consider individual differences in the amount of wheel rolling. In Table 4, there are two subjects (B and J) who do not use a wheel. So the amount of wheel rotations is a fixed quantity data with large individual differences. However, evaluation of a subject who does not use a wheel can be discriminated only by the moving speed of the gazing point. Moreover, we consider that it is necessary to apply the discrimination result by the wheel in the case of the subject who does use a wheel.

Based on these considerations, 45 pages which are incorrectly identified as possible low usability might be correctly classified using individual measurements. While this would significantly improve the accuracy of identification, the current data collection and analysis did not attempt to do this. Future research may want to investigate this further.

4.7 Discriminant Boundary

Various steps are required to obtain a general discrimination boundary. It is necessary to choose many subjects from a set of all users at random. It is necessary to choose many websites from a set of all websites at random. However, a huge budget is required to choose such subjects. Therefore, we experimented with limited group of subjects. Moreover, since the experimental result may be opened, when a website is used for an experiment, it is necessary to obtain a Web administrator's permission. Therefore, we were not able to experiment with many websites. However, if the number of subjects and the number of websites can be increased from now on, we expect that a general discrimination boundary may be determined. If such a discrimination boundary can be analyzed, it will become unnecessary to perform the discriminant analysis at every usability testing.

5. CONCLUSION

We experimented to investigate the relation between quantitative data about viewing users' behavior and web usability evaluation by subjects. As a result of this quantitative analysis, we showed clearly that the moving speed of the gazing points is effective in detecting low usability Web pages. In addition, 94.4% of low usability web pages (17pages / 18pages = detectable pages / low usability pages) were detectable using the moving speed of the gazing points and the amount of wheel rolling of a mouse. This detection reduced the number of web pages which should be evaluated by half (46% = 89 pages / 192 pages = detected pages / all pages).

The quantitative data which we evaluated is already used in existing usability evaluations. This analysis result indicates a possibility for doubling the efficiency of evaluation without increasing the cost of data collection. However, it is necessary to increase the sites and subjects for evaluation. Moreover, we must also clarify the behavior of a subject on a page with features that cannot be discriminated, to remove it, and to investigate higher discrimination.

Since the cost which Web usability evaluation takes is increasing in software development, increasing efficiency is required. Before a specialist analyzes the quantitative data of users' behavior, the specialist needs to evaluate all Web pages which the user visited to discriminate low usability pages. We expect that this research can lead to reduced evaluation costs by setting only the discriminated low usability pages as the evaluation target. Moreover, we consider that it also may become possible to point out detailed problems in a Web page quantitatively by investigating changes in the quantitative data containing the gazing point when facing a Web usability problem within a Web page.

6. ACKNOWLEDGMENTS

This work is supported by the Comprehensive Development of e-Society Foundation Software program of the Ministry of Education, Culture, Sports, Science and Technology (MEXT) and Grant-in-Aid for Japan Society for the Promotion of Science (JSPS) Fellows (Research No: 16005035).

7. REFERENCES

[1] Mon Chu Chen, John R. Anderson, Myeong Ho Sohn. What can a mouse cursor tell us more?: correlation of eye/mouse movements on web browsing. *CHI '01 extended abstracts on Human factors in computing systems table of contents*, 2001, 281-282.

[2] Andrew T. Duchowski. Eye tracking methodology: theory and practice. Springer, 2003.

[3] J. S. Dumas, J. C. Redish. A Practical Guide to Usability Testing. Ablex Publishing, 1993.

[4] Kelly Goto, Emily Cotler. Web ReDesign. Pearson Education, 2002.

[5] Masatoshi Mori, Tetsuo Ui. Research on effectiveness of gazing point movement analysis in screen design, Japan

Society for the Study of Office Automation (JSSOA), Vol.16, No.3, 1995, 49-56. (In Japanese).

[6] Jakob Nielsen. Designing Web usability. New Riders, 2000.

[7] Hidehiko Okada. Usability and its evaluation technique. *The Institute of Systems, Control and Information Engineers*, Vol.45 No.5, 2001, pp.269-276. (In Japanese)

[8] Hidehiko Okada, Toshiyuki Asahi. GUITESTER: a log-based usability testing tool for graphical user interfaces. *IEICETrans. on Information and systems*, Vol.E82-D No.6, 1999, 1030-1041.

[9] Laila Paganelli, Fabio Paterno: " Intelligent analysis of user interactions with web applications," Proceedings of the 7th international conference on Intelligent user interfaces table of contents, pp.111-118 (2002).

[10] Makoto Sakai, Noboru Nakamichi, Jian Hu, Kazuyuki Shima, Masahide Nakamura. Webtracer: A New Integrated Environment for Web Usability Testing. In 10th Int'l Conference on Human - Computer Interaction (HCI International 2003), Crete, Greece, Adjunct Proceeding, June 2003, 289-290.

[11] Melody Y. Ivory, Marti A. Hearst. The state of the art in automating usability evaluation of user interfaces. *ACM Computing Surveys*, Vol.33, Issue 4, 2001, 470-516.