Development and Evaluation of A Usability Evaluation Tool: WebTracer

N. Nakamichi[1], M. Sakai[2], J. Hu[1], K. Shima[1], M. Nakamura[1]

[1]*Nara Institute of Science and Technology*
[2]*SRA Key Technology Laboratory, Inc*
noboru-n@is.aist-nara.ac.jp, sakai@sra.co.jp, {jian-hu, shima, masa-n}@is.aist-nara.ac.jp

1. INTRODUCTION

Usability evaluation of Web application is very important because the usability level will influences the sales of a company that adopts web application in business [2]. Therefore, the usability evaluation technique that can evaluate easily, and its support tool are required. The various usability evaluation techniques were studied until now [3]. Moreover, the automatic usability evaluation technique and evaluation support tool were developed too.

For usability evaluation, we consider that display of a screen, eye mark and operation history of a user are also effective like that in software analysis [4]. Eye mark is a user's gazing point here. Moreover, We consider that the tool which integrated these needs to be used [6]. Therefore, we developed the tool WebTracer by which the operation record of Web page was recorded and analyzed. There are some other tools that can use logs on web server or JavaScript to record user operations [1][5]. Scan converters and video recorders (e.g. analog, MPEG 2/4) can be used to record computer screen on videotapes. Some tools (e.g. Lotus ScreenCam) can record computer screen on hard disk. Such conventional (non-integrated) tools cannot synchronously playback users operations and computer screen.

It is considered that the WebTracer can find the problem that cannot found by the typical usability evaluation techniques such as Thinking Aloud etc.

2. FUNCTIONS OF WEBTRACER

WebTracer has the following three functions.

(1) The record function

It records the screen data, the eye mark, and mouse and keyboard data of user operation. Also the display event of a Web page was recorded. Since WebTracer captures one still image of a web page in true resolution in front of every page navigations, the size of the recorded data is about 1/10 compared with MPEG 2.

(2) The reproduction function

The recorded data and the picture of a Web page are integrated and displayed. As shown in Fig. 1, when reproducing user's operation using a web browser, an
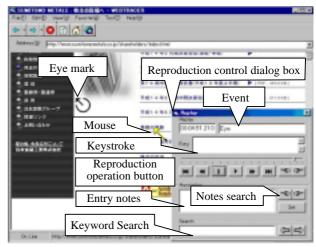

Fig. 1. The reproduction function of WebTracer

icon is displayed according to the coordinates of a mouse and an eye mark that appear on the screen of a Web page.

(3) The analysis function

This function consists of a time-series-analysis function and a collected data analysis function. The time-series-analysis function can display and search notes recorded at the time of reproduction. Therefore, detailed analysis of user operation is supportable. The collected data analysis function can display the total data and mouse/eye mark speed list/graph of operation records of the every web page.

3. EVALUTION EXPERIMENT

We have conducted an experiment, in order to evaluate the effectiveness of WebTracer in Web usability evaluation. In the experiment, we imposed the examinees a task to find objective information within a website of a company. Then, with the support of WebTracer, a hearing to the examinees was performed and examinees' comments were recorded.

3.1 Experiment Procedure

The experiment was conducted by the following processes.

P0: An examinee performs the task. The operations are

being logged in the background by WebTracer.

P1: WebTracer shows the analysis results. Based on the results, we perform the hearing.

P2: We focus on the pages problematic with respect to the usability (e.g. Movement of eye mark is quick, There is many movement of eye mark or mouse) and perform the hearing again with playing back the operations.

P3: A hearing for whole the task is performed.

P4: With fast-forwarding the operation record, we did the final hearing.

The task for the examinees is to gather the following five information from the website of a company.

T1: The way to a certain place in the company

T2: The number of employees

T3: Welfare program of the company

T4: A specific news

T5: A technical method of construction

3.2 Experiment results

The average number of comments given by three examinees is summarized in Table 1. The comments include the usability problems, thinking during operation, etc. We suppose that the unit of a comment is every sentence of utterance of an examinee.

The average execution time per task in the experiment was 2 minutes and 48 seconds. Also, the average time taken for the hearing was 19 minutes. So, the total time spent for whole processes was 21 minutes and 48 seconds. For the execution time, we obtained 16 comments on average. The number of comments becomes material for an evaluator judges it as the problem about usability.

4. DISCUSSION

We consider that the experiment shown in Section 3.1 would be very difficult without WebTracer, even though applying other conventional usability evaluation methods. We summarize, in Table 2, the applicability of the other methods to the experiment conducted.

In Table 2, "Good" indicates the method is suitable for the process. "Possible" shows that the method can be used for the evaluation. "Weak" represents that it may not be used for some case. "N/A" means that the method is not originally intended to be used for the evaluation.

For example, let us take "U: Observation in a usability laboratory." First, U cannot be used for process P1, since no quantitative result is available. It can be applied to P2 and P4, if the evaluator records the examinees' operations with video camera, etc. P3 is also possible by performing questionnaire-based evaluation in the end of the task.

By this experiment only, we cannot do quantitative and fair comparison with other conventional methods. However, if we assume that only "U: Observation in a usability laboratory" is available, then process P1 cannot

Table 1

The average number of comments by three examinees

	T1	T2	T3	T4	T5	Total	(%)
P1	0.7	2.3	2.3	1.7	2.3	9.3	11.6
P2	7.7	13.7	12.0	8.0	18.7	60.0	74.4
P3	2.3	0.0	0.0	1.3	0.3	4.0	5.0
P4	0.7	0.7	1.7	1.0	3.3	7.3	9.1

Table 2

Comparison with other evaluation methods

	P	T	U	I	WT
P1	Possible	N/A	N/A	Weak	Good
P2	N/A	N/A	Possible	N/A	Good
P3	Possible	Possible	Possible	Possible	Possible
P4	N/A	N/A	Possible	N/A	Good

P: Performance measurement
T: Thinking aloud
U: Observation in a usability laboratory
I: Usability Inspection methods
WT: WebTracer

be performed. Consequently, 9.3 comments in Table 1 would not be available. As a result, "U: Observation in a usability laboratory" would miss 11.6% of whole comments shown in Table 1.

5. CONCLUSION

We have developed a new record/analysis tool, WebTracer, for the effective Web usability evaluation. We have also conducted experimental evaluation. The evaluation using WebTracer can cover the feature of other evaluation techniques. Our future work is to conduct more evaluation and comparison with other evaluation techniques quantitavely.

REFERENCES

[1] Michael Etgan, and Judy Cantor: "What does getting WET (Web Event-logging Tool) mean for web usability?," HFWEB99, 1999.

[2] Kelly Goto, and Emily Cotler: "Web ReDesign," Peason Education, 2002.

[3] Jakob Nielsen: "Usability Engineering," Academic Press, 1993.

[4] Hidehiko Okada, and Toshiyuki Asahi: "GUI TESTER: log-based usability testing tool for graphical user interfaces," IEICE Trans. on Information And Systems, Vol.E82-D, No.6, pp.1030-1041, 1999.

[5] Laila Paganelli, "Intelligent analysis of user interactions with web applications," IUI'02, 2002.

[6] Koji Torii, Ken-ichi Matsumoto, Kumiyo Nakakoji, Yoshihiro Takada, Shingo Takada, and Kazuyuki Shima: "Ginger2: An environment for computer-aided empirical software engineering," IEEE Trans. on Soft. Eng., Vol.25, No.4, pp.472-492, 1999.