Development of Integrated Analysis System and Tool of Perception, Recognition, and Behavior for Web Usability Test: With Emphasis on Eye-Tracking, Mouse-Tracking, and Retrospective Think Aloud

Byungjoo Kim, Ying Dong, Sungjin Kim, and Kun-Pyo Lee

Abstract. Recent researches reveal effort to observe user's experience from user's point of view in order to estimate usability of a web site. Eye-tracking and mouse-tracking to record and analysis what user sees and how user acts can be proper examples. However, although eye-tracking and mouse-tracking are used practically, not only difficult to find the case that uses both, but also the case is rare that considers what user is thinking. Hence, this paper introduced EMT System that tracks eye and mouse, and records user's thinking. And for applying EMT system, this paper developed EMT Tool, which helps a researcher to do usability test by recording the user's experience, and reproducing it visually. EMT Tool is consist of EMT Tracker which is responsible for observing and collecting user's experience and EMT Analyzer synthesizing and analyzing data from EMT Tracker.

Keywords: Web Usability Test, Integrated Analysis System, Eye-Tracking, Mouse-Tracking, Retrospective Think Aloud.

1 Introduction

Since World Wide Web (WWW) was firstly introduced by Tim Berners-Lee in 1989, it has been widely used by public nowadays, and influence the society and culture in many aspects. Following web popularization, web business is growing rapidly. As the competitions among global companies were deepen, the concern of web usability was more and more emphasized. Jacob Neilson said: "web users undergo the usability before using a web or deciding a latent purchasing",[1] which is to emphasis web usability through realistic problems in order to get web business success.

People began to consider the importance of web usability. Many researches which relate to usability evaluation methods are actively carried out in order to improve the web usability. Various usability evaluation methods such as user questionnaire, Heuristics evaluation, Usability test, and Card sorting are widely used for web usability evaluation. But, such usability evaluation methods were originally developed for products and software, some problems including subjective evaluation from evaluators' point of view, difficulties of user observation, and limitations of grasping users' needs come out if apply those methods to web usability evaluation. Moreover, as the asynchronous communication is expending with the web2.0's appearance, the

---

[1] Jacob Nielsen, Designing Web Usability, New Riders Publishing, 2000.

observing method of users' interaction through Web Server's Data Logging is loosing the function of evaluating web usability.

This study tries to overcome the limitations of existing usability evaluation, and aims to develop analysis methods based on users' experiences instead of the web usability evaluator's perspectives.

2   Web Usability Methods and User Experience Review

2.1   Classification of Web Users' Experience

Existing web usability evaluation methods have not satisfied the sub-conscious, tacit, emotional, and qualitative characteristics of the web observation and analysis required for the web usability evaluation. And those have been used on the evaluators' point of view. So that above problems comes out. It is very important to observe users' experience in users' perspectives in order to evaluate web usability.

As mentioned above, it is necessary to understand users as the core of the usability to observe web users' experience at firsthand. In the web environment, users could have various and complex experiences in a short time. So it is proper to apply observation method according to different users' experience sections.

By comparing and synthesizing Norman's seven Stage of Action, Trumbo's Spatial Environment in Multimedia Design, and Kantowitz's Human Information Processing Model, three sections such as perception, recognition, and behavior of users' experience can be drew, as the table 1 shows.

Table 1. Comparison of users' experience categories

	Perception	Recognition	Behavior
Norman's Seven Stage of Action	Perceiving the State of the World	Interpreting the Perception, Evaluation of Interpretations, Goals, Intention to Act, Sequence of Actions	Execution of the Action Sequence
Trumbo's Spatial Environment in Multimedia Design	Perceptual Space	Conceptual Space	Behavioral Space
Kntowitz's Human Information Processing Model	Perceptual Stage	Central Processing	Action Stage

Users' perception is the procedure of knowing and long-time memorizing the outer world stimulus through the sensible organs. The recognition is the procedure of setting the action plan for realizing aims by knowing and evaluating the inputted information. Behavior is the procedure of expressing the real executions by following the action plan. Users' experience sections such as perception, recognition, and behavior are based on users' recognition models. So it can be used not only on the interaction between physical objects like products and users, and also the interaction between web and users.

2.2   Observation Methods for Each Specific Experience

By following web users' experience section of perception, recognition, and behavior as mentioned above, using different user observation method according to each specific experience is effective on web usability evaluation.

Eye tracking is suitable for web users' perception experience observation. While web users staring at computer interface, web contents such as text, graphic image, and various multimedia can be perceived. In the graphic user interface of web environment, user's visual sense takes great proportion in perceiving. The basic principle of eye tracking is to analyzing eye movement by measuring eyeballs' movement such as fixation and searching.

In order to observe web users' recognition experience, 'Think Aloud' is an existing usability evaluation method which lets users to speak out what they think during the web usability evaluation process. Because users can explain their perception process if they speak out,[2] evaluator can observe users' recognition process through their words. Usually, 'Think Aloud' is used to let users to speak out their thought when they are performing tasks. But, at the same time, 'Think Aloud' results users' unnatural behavior and less concentration during the performance. It is not suitable for eye tracking method which is used for users' perception observation. Another 'Think Aloud' method, which is called 'Retrospective Think Aloud', asks users to speak according their memories after they finish performing those tasks. The possibiliy of reliance problem exists because users provide untruthful description in order to make them sounds more reasonable. But, Zhiwei Guan[3] proofed that the retrospective think aloud method has satisfied reliability and validity. The retrospective think aloud was used as web user's recognition experience observation method in this study.

Mouse tracking is suitable for web users' behavior experience observation. In most interaction of GUI web interface, cursor's movement can be captured.[4] Compare to Video Ethnography which is suitable for observing users' movement in the physical world through the device such as video camcorder, mouse tracking is useful for observing web users' movement which appears on the monitor's screen background.

3   EMT System: Integrated Evaluation System of Perception/ Recognition/Behavior

3.1   Introduction of EMT System

Web usability evaluation aims to understand users' latent think or interaction pattern by observing web users' usage. In order to observe web users' experience from all sides, this paper proposed observation method for users' specific experience and "EMT (Eye, Mouse, Think) System; Perception/Recognition/Behavior integrated

---

[2] Jeeny Preece, Human Computer Interaction, Addison-Wesley, 1994.

[3] Zhiwei Guan, Shirley Lee, Elisabeth Cuddihy, Judith Ramey, The validity of the stimulated retrospective think-aloud method as measured by eye tracking, Proceeding of CHI2006, 2006.

[4] ChangMin Park, A Study on the Visulization & Analysis of User Interaction on the WWW, Master Thesis at KAIST, 2001.

evaluation system" which can analyze web usability problems. EMT System which includes eye tracking, mouse tracking, and retrospective think aloud which are parallel used for observing users' specific experience is a systematic method, which synthetically analyzes observed users' perception process from specific experience.

3.2 Constitution and Progress of EMT System

To observe meaningful web usability problems from users, EMT System is based on the Usability Test which can get much detail data from users.
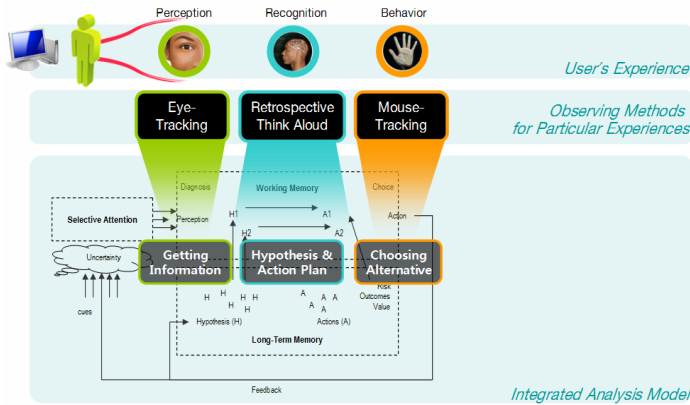

Fig. 1. Constitution of EMT System

In the usability test as showed in figure 1, web users' experiences were observed when they were performing the given tasks. The perception of each user was observed by eye tracking, the recognition was by retrospective think aloud method, and the action was by mouse tracking. Each specific experience data of observed users were analyzed according to the integrated analysis framework.

The usability evaluation experiment was carried out according to the usability test matrix in which the evaluation system was referred is showed in the figure2 bellow.
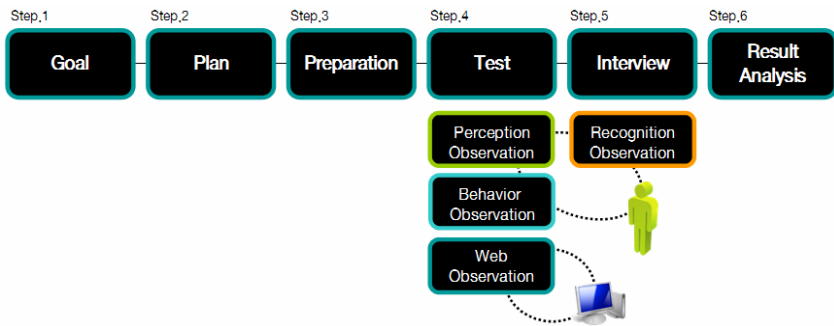

Fig. 2. Progress of EMT System

3.3   Analysis Frameworks

The EMT System analyzes the collected web users' specific experience observing data synthetically through human's information processing model of decision making. During performing the task in web users' evaluation experiment, user makes decision among many alternatives. This process related analysis can be used for analyzing web usability problems. Decision making basically contains 3 phases: getting information, making hypotheses, choosing plan and execution. According to the decision making processing model, users make many alternatives or hypotheses by using a few clues from external stimuli and then choose one of the alternatives for the action plan based on users' memory and finally carry out.[5] Those 3 phases of the information processing model of decision making, as showed above, can be applied to and be analyzed with Perception/Recognition/ Behavior phase of human experience.

Table 2. User's Particular Experiences and Steps of Decision Making

User's Experience	Step of Decision Making	Description
Perception	getting information	Getting information and clues from external stimuli
Recognition	making hypotheses	Making hypothesis and action plan based on long-term memory
Behavior	choosing plan and execution	Executing the chosen action plan

In the EMT System, the process of users' perception can be understood as the process of getting information in the decision making model. Through eye-tracking data, users' picked information and clues from web page can be analyzed. Next, the process of recognition can be understood as the process of making hypotheses. With the description from retrospective think aloud, it is possible to analyze how users set up the hypotheses and action plans for the task, under exposure to visual stimulation or information from the web page. Finally, the process of behavior can be understood as the process of choosing alternatives. By following the hypotheses and action plans in users' recognition process, how to execute specifically can be analyzed through mouse-tracking data.

The information processing model of decision making is different depending on users' different degree of expertness. Because the novice and expert have differences in terms of information processing, task execution path, background knowledge and experience, they would act differently in each phrase of getting information, making hypothesis, and choosing alternatives. Based on Gordon's skill-rule-knowledge model, the EMT System divides users' degree of expertness into analytical behavior, intuitive behavior, and automatic behavior.

As showed in table 3, in the process of performing the given tasks, the proportion of each specific experience is different according to different degree of expertness. Because the novice needs to process much information in the process of Perception/Recognition/Behavior, the observed data from each experience are also huge. Especially, the data from the complex thinking process are huge. On the contrary, the intermediate level user understanding the outer stimulus and using rules

---

[5] C.D. Wickens, S.E. Gordon, and Y. Liu, Intro. Human Factors Engineering, Longman, 1988.

Table 3. Degrees of User's Expertness and User Observation Data

Degree of Expertness	Behavior	User Observation Data		
		Perception	Recognition	Behavior
Novice	Analytical Behavior			
Intermediate	Intuitive Behavior	○	○	○
Expert	Automatic Behavior	○		○

and habit in the recognition process, the observed data are smaller. Especially, since the experts skip the analysis of stimulus from outside, and since they go directly into the task execution, the data collected from the recognition process stage would be insufficient.

In the integrated Perception/Recognition/Behavior evaluation system, the usage problems that vary on the degree of expertness can be classified as follows:

1. Novice users: usage problems can be found on analytic recognition stage, such as limited short-term memory, biases on developing hypothesis or action, cognitive clinger.
2. Intermediate level users: usage problems that are related to the mis-application of rules, based on misled decision making.
3. Expert users: Usage problems can be observed in the cases that users pay attention to unimportant places. On the contrary, sometimes users pay too much attention to the task itself. The over-concentration on the task disturbs the processes of the task.

4   EMT Tool: Integrated Evaluation Tool of Perception/ Recognition/Behavior

4.1   Development of EMT Tool

The EMT System collects the user's experience data and web environment data from the usability test experiment. The collected data is not analyzed individually, but instead it is connected to other kinds of data and the user's task progress situation as well as the recognition process of the user is analyzed. In order for this type of integrated analysis, the EMT System requires the EMT Tool; Integrated evaluation tool of Perception/Recognition/Behavior that can effectively manage and analyze various kinds of data.

The requirement for EMT Tool can be summarized into the following categories of data logging, data synchronization, visualization of generalized data, reduction of video analysis time, and memory recollection material.

4.2   Scenario of EMT Tool's Application

The EMT Tool is utilized from 'Step 3. Preparation' to 'Step 6 Results analysis' of the progress scenario of EMT System (refer to Fig. 2). In order to see the role and progress process of EMT Tool in a scenario, refer to Fig. 3. After the calibration, EMT Tool collects the user's eye and mouse tracking data and web environment data in the Test stage and stores the data. And in the Interview stage all the data are
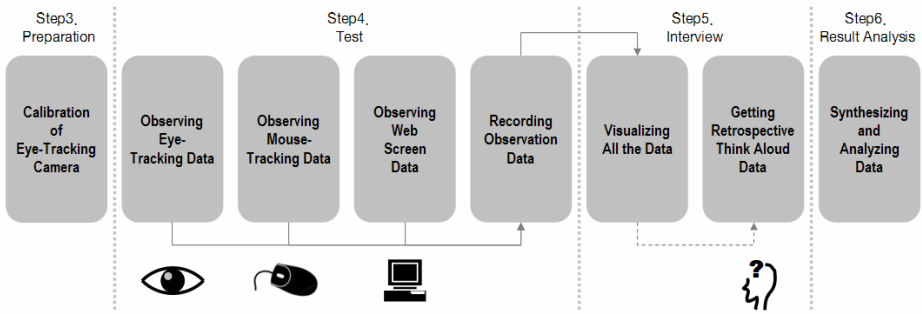
Fig. 3. Scenario of the EMT Tool

collectively reproduced visually while the user records the retrospective think aloud accounts using it as a reference.

4.3 Realization of EMT Tool

The EMT Tool is a tool that realizes the execution of the EMT System. The EMT Tool acts as the "Observance of the user's experience and web environment" and "the analysis of collectively observed data". To carry out these two tasks, EMT Tool consists of two different tools which are EMT Tracker, which is the observe/collect tool, and EMT Analyzer, which is the synthesize/analyze tool. The EMT Tracker is executed in the 1) zero point adjusting of the eye tracking device ~3) data storage stages of the tool scenario [Fig. 3]. The EMT Tracker's process is as follows: First the zero point is adjusted between the testee and the sight tracking equipment, and then it observes the user and the web to collect data, and finally stores this data in the data storage stage.

The EMT Analyzer is executed in the 4) Visualization of data ~ 6) Integrated Analysis stage [Fig. 3]. The EMT Analyzer loads the data that was stored by the EMT Tracker, carries out the analysis, the sight, mouse, and web screen data is collectively visualized, and the testee's Think Aloud description is inputted to provide an analysis frame.

The screen of the EMT Tracker is shown in [Fig.4]. The calibration and the starting point and the end button of eye-tracking recording are provided as a toolbar. The screen of the EMT Analyzer is shown in [Fig. 5]. In the background, the web screen video data is played while the user's eye movement path (red) and mouse movement path (blue) is shown above the screen video. In addition, 'EMT information palette' for data adjustment and confirmation, 'Time slider palette' for data play search, 'EMT analyzer palette' for consciousness/thinking/movement integrated analysis are provided. The user can move the palette and also can set it to disappear from sight and can also move the palette while the player is being played. Also, when a singularity is found, contents can be inserted at any time and place, and the place of insertion is indicated on the player bar so that an effective analysis is possible.
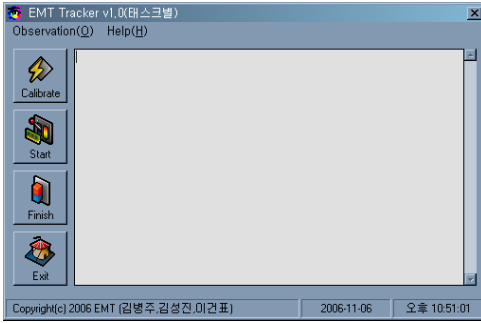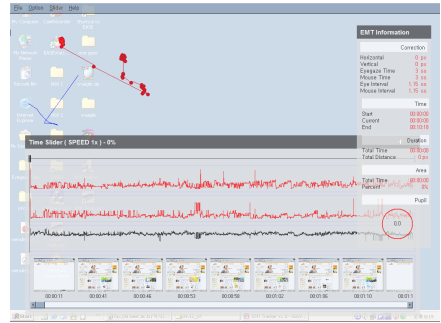
Fig. 4. EMT Tracker

Fig. 5. EMT Analyzer

As it was described above, the EMT Tracker and EMT Analyzer of the developed EMT Tool send and receive data and progress linked to each other. [Fig.6]
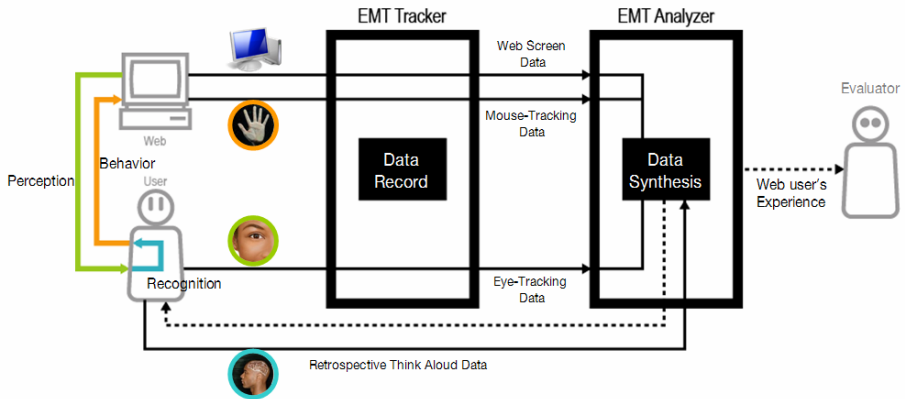

Fig. 6. Overview of EMT Tool

5   Conclusion and Future Study

For a new web usability evaluation method, this research proposes EMT System that can directly observe the web user's experience, and can analyze the user's potential thinking procedure, and at the same time EMT Tool was developed in order to process the evaluation system. If we elaborate in more detail, first, the user's experience is divided into perception, recognition, and behavior and each item's experience observation methods are as follows: eye-tracking for perception, retrospective think aloud for recognition, mouse-tracking for behavior. Second, EMT System was proposed from the generalized user's observation method for each detailed experience, and that process of usability evaluation experiment and data observation and the process of collection, and data analysis frame work were explained. Third, EMT Tool was developed to execute EMT System.

The EMT Tool, in comparison to other eye-tracking and mouse-tracking methods utilized in previous web usability evaluation methods, has the advantage where both eye-tracking data and mouse-tracking data can be shown on the web usage screen and that it is able to complementarily analyze two different data simultaneously. Also, it can spontaneously input retrospective think aloud accounts from revived memory of what the user saw and reacted during the process of web experience through eye-tracking and mouse-tracking data. EMT Tool is not a one-way interface where the evaluator questions the user. It is a bi-directional analysis interface where the user spontaneously records what he or she felt and thought during the web experience process and the evaluator analyzes it.

For future research, first, we require an improvement of interface for increase in usability of EMT Analyzer as well as an addition of result output function for increase in practical usage. Although it is possible to use various functions through a menu, since accessibility can be an issue when using a menu, a tool bar is needed for user's convenience. Also, for an immediate comprehension for the user, the text-centered interface needs to be modified to the icon-centered type. In addition, documentation of analysis results or a function to output contents in commercial file types for editing is required. Second, in order to verify the EMT System and the practical usage of EMT Tool and its wide usefulness, research on various user and environment in different cases is required. Even though EMT System and EMT Tool were developed for web usability evaluation purposes, it can observe web in computer environments as well as in various software. Therefore, research utilizing EMT System and EMT Tool in various situations must be carried out.

References

1. Park, C.M.: A Study on the Visulization & Analysis of User Interaction on the WWW, Master Thesis at KAIST (2001)
2. Wickens, C.D., Gordon, S.E., Liu, Y.: Intro. Human Factors Engineering. Longman, Longman (1988)
3. Donald, A.N.: The Design of Everyday Things, Doubled Currency (1988)
4. Gordon, S.E.: An Information-Processing Model of Naturalistic Decision Making. Presentation at Annual Meeting of the Idaho Psychological Association, Sun Valley, Idaho (1997)
5. Nielsen, J.: Designing Web Usability, New Riders Publishing (2000)
6. Trumbo, J.: The Spatial Environment in Multimedia Design: Physical, Conceptual, Perceptual, and Behavioral Aspects of Design Space, Design Issues, vol. 13(3)
7. Preece, J.: Human Computer Interaction. Addison-Wesley, London (1994)
8. Kantowitzm, B.H.: The role of human information processing models in system development. In: Proceedings of the human Factors Society 33rd annual meeting, Santa Monica, CA, Human Factors Society (1989)
9. Guan, Z., Lee, S., Cuddihy, E., Ramey, J.: The validity of the stimulated retrospective think-aloud method as measured by eye tracking. In: Proceeding of CHI2006 (2006)