On Using Eye Tracking in Empirical Assessment of Software Visualizations

Huzefa Kagdi, Shehnaaz Yusuf, and Jonathan I. Maletic
Department of Computer Science
Kent State University
Kent Ohio 44242
{hkagdi, sdawoodi, jmaletic}@cs.kent.edu

ABSTRACT

The paper advocates the use of eye movement measurements in conducting empirical studies of software engineering tools, especially visualization techniques. Traditionally, measures such as accuracy and performance time have been used to assess and compare different tools for a given set of tasks. These measures are typically collected after the conclusion of an assigned task. Eye tracking adds a new additional dimension to the assessment arsenal by allowing access to the gaze activity of human subjects. The gaze activities can be captured quite precisely while a task is being performed. Thus, providing a unique opportunity to include measures of how exactly humans use a tool and ratiocinate their conclusions. A brief discussion on using the eye movements for assessing UML class diagram layouts is also presented.

Categories and Subject Descriptors

D.2.7. **Software Engineering**: Distribution, Maintenance, and Enhancement – documentation, enhancement, extensibility.

General Terms

Experimentation, Measurement, Documentation

Keywords

Eye tracking, Assessment metrics, Empirical studies

1. INTRODUCTION

Empirical studies designed as surveys, case studies, and experiments are a rigorous means of evaluating and comparing software engineering tools including visualization metaphors and techniques. Studies designed to validate software visualization techniques typically involve human subjects answering (performing) a set of questions (tasks). Arguably, irrespective of the very careful soundness consideration in the study design (e.g., selecting appropriate tasks and their distribution among participants), the validity of the achieved results directly corresponds to the type and quality of the used measures.

Traditionally, objective measures such as the accuracy/level of the response and time needed are collected from these studies [3, 4]. Additionally, subjective data such as a subject's experience, comments, preferences, and any other feedback are also collected. These measures are used directly or indirectly to draw conclusions and/or meet other objectives of the performed study. A wide majority of these traditional measures are collected retrospectively. For example, human subjects are asked to report their final answers on the completion of a given task and their response time is recorded. We term such measures as *black box* measures as they only record the final outcome after a specific task conclusion. That is, no other data is collected, at least not implicitly, while a human subject is performing a given task. There is almost no measurement taken that helps understand how and why a subject chose a particular (correct or incorrect) answer or solution. Additionally, black box measures raises a potential threat to the validity of the study, namely the match/disparity between the subjects' responses on completion of a task and the "reality" they observed while performing that task. For example, a subject may forget to report (or misreport) an observation after a lengthy task. Alternatively, subjects could be asked to note their observations while working towards their answers, albeit at the potential risk of obtrusiveness and distraction.

Our position is to use eye-tracking equipment to implicitly collect a subject's activity data in a non-obtrusive way while they are performing a given task. The equipment collects three forms of pertinent data including the eye gazes on the visual display and an audio/video recording of the subject's session. Eye gazes are substantiated with the measurement of various eye movements. This eye movement data could provide a much valuable insight as to how and why subjects arrive at a certain solution. Therefore, we term the eye gaze measures collected from eye tracking as *white box* measures. These measures add a new additional dimension in assessing visualization tools claim of supporting software comprehension tasks. Here, we briefly discuss the concepts of eye tracking in the context of our recently reported study on assessing UML class diagram layouts [5]. Also, we discuss how eye movement measures could be used to assess exploration, examination, and navigation support.

2. EYE TRACKING

The underlying basis of an eye tracking equipment is to capture various types of eye movements that occur while humans physically gaze at an object of interest. Fixation and saccade are the two most common types of eye movements.

<u>*Definition*</u>*: Fixation* is the stabilization of eyes on an object of interest for a certain period of time.

<u>*Definition*</u>*: Saccades* are quick movements that move the eyes from one location to the next (i.e., refixates).

WEASELTech'07, November 6, 2007, Atlanta Georgia, USA

Definition: *Scanpath* is a directed path formed by saccades between fixations.

The general consensus in the eye tracking research community is that the processing of visualized information occurs during fixations, whereas, no such processing occurs during saccades [2]. Humans use saccades to locate interesting parts in a visual scene to form a mental model.

Figure 1 shows the recording of eye positions superimposed on a UML class diagram. The numbered circles represent fixations and lines between them represent saccades. The size of a fixation (i.e., area of a circle) is proportional to its time duration. The numbering of circles represents the ordering of fixations. For example, in Figure 1, the fixation labeled with the number 35 on the class *NTuple* happened before the fixation labeled 36 on the class *NTupleController*. That is, the class *NTuple* was looked at before the class *NTupleController*. The scanpath in this case is directed to the left and downwards. A big circle on the class *PyNTuple* shows that a large amount of time was spent on this class. The eye-tracker captures fixation and saccades in the form of XY coordinates of the visual screen from which we can determine what was being looked at in a visual presentation.

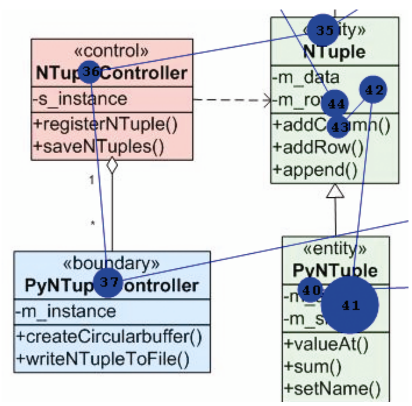

Figure 1. Gaze Information on a UML Class Diagram. Fixations are represented with circles and saccades with lines.

We used a *Tobii 1750* eye-tracker (www.tobii.se) to capture eye movements and collect eye gaze data. In this equipment, the two cameras used to track the eyes are built into a 17 inch flat-panel screen. Therefore, no restraints such as wearing a headband or goggles are placed on the human subject. This was not the case in older eye tracking equipment. This provides a normal computer-operating environment during the study. Moreover, the *Tobii 1750* eye-tracker is very accurate with an error rate of less than 0.5 degrees and a sampling rate of 50MHZ. Software that records the XY screen coordinates of eye gazes and supports analysis of eye movements is also provided along with the eye-tracker system. An audio/video recording is also made of each study session.

Now, we describe the use of white box measures such as fixations and saccades in the evaluation of UML class diagram layouts. From layout perspective, the support for exploration, explanation, and navigation are of general interest.

An exploration activity deals with how subjects perform searches on the UML class diagram to locate objects required for a given task. The number and size of fixations could help identify areas of the layout that smoothly assisted or created bottlenecks in the exploration activity. Also, the scanpaths provide the order and directionality information in which the search space was traversed. For example, do the recorded scanpaths justify a particular layout's strategy of placing certain classes at a particular position? Were only the relevant classes immediately visited and only once?

An examination activity deals with how subjects visualize, in detail, whole or parts of a specific class and relationship. In our experience, fixations can be recorded at the granularity of a specific line (i.e., class, attribute, method names). Thus, fixations could be used to assess questions/tasks that are related to a specific class. Also, the durations of fixations give information about which parts of a specific class receive the most attention.

A navigation activity deals with how subjects move from one object of interest to the next after their discovery. Once again fixations and saccades could be used to justify a layout's strategy in supporting navigation.

In our previous study [5], we used the number of fixations as an indicator of the required human effort. Fewer number of fixations on a layout means that the subject needs less effort to answer the associated question. If the total number of fixations is high then the classes and relationships are possibly laid out in a way that leads to an inefficient visual exploration, explanation, and navigation. Such poor arrangement spans the attention of a subject across a number of objects instead of systematically narrowing down to only the relevant area of interest.

3. SUMMARY

The basic premise of eye tracking methods lays in the strong eye-mind hypothesis [1], which states that human gazes directly correspond to their thinking and cognitive process. We believe that white box measures from eye tracking are a promising step towards developing objective metrics for software comprehension and cognitive load. The measures should be used synergistically with the traditional black box measures in the empirical assessments of software Visualization (generally engineering) tools. Considering a number of recent advancements in eye-tracking technology, empirical researchers now have a very effective and unprecedented tool.

4. REFERENCES

[1] Hyona, J., Radach, R., and Deubel, H. *The Mind's Eye:Cognitive and Applied Aspects of Eye Movement Research*, Amsterdam, 2003.

[2] Jacob, R. J. K. What you look at is what you get: eye movement-based interaction techniques in Proceedings of SIGCHI conference on Human factors in computing systems: Empowering people (Seattle, WA, USA, 1990), 11-18.

[3] Purchase, H., C., Colpoys, L., McGill, M., Carrington, D., and Britton, C. UML Class Diagram Syntax: An Empirical Study of Comprehension in Proceedings of Asia-Pacific Symposium on Information Visualisation (Sydney, Australia, 2001), 113-120.

[4] Ricca, F., Penta, M., Torchiano, M., Tonella, P., and Ceccato, M. The Role of Experience and Ability in Comprehension Tasks Supported by UML Stereotypes in Proceedings of International Conference on Software Engineering (ICSE'07) (Minneapolis, MN, US, 2007), 375-384.

[5] Yusuf, S., Kagdi, H., and Maletic, J. I. Assessing the Comprehension of UML Class Diagrams via Eye Tracking in Proceedings of International Conference on Program Comprehension (ICPC'07) (Banff, Canada, 2007), 113-122.