Low cost vs. high-end eye tracking for usability testing

Conference Paper · May 2011

DOI: 10.1145/1979742.1979744 · Source: DBLP

Low Cost vs. High-End Eye Tracking for Usability Testing

Sune Alstrup Johansen
IT University of Copenhagen
Rued Langgaards Vej 7
2300 Copenhagen S, Denmark
sune@itu.dk

Javier San Agustin
IT University of Copenhagen
Rued Langgaards Vej 7
2300 Copenhagen S, Denmark
javier@itu.dk

Henrik Skovsgaard
IT University of Copenhagen
Rued Langgaards Vej 7
2300 Copenhagen S, Denmark
hhje@itu.dk

John Paulin Hansen
IT University of Copenhagen
Rued Langgaards Vej 7
2300 Copenhagen S, Denmark
paulin@itu.dk

Martin Tall
Duke University
2424 Erwin Rd. (Hock Plaza)
Suite 302
Durham, NC 27705
info@martintall.com

CHI 2011, May 7–12, 2011, Vancouver, BC, Canada.
ACM 978-1-4503-0268-5/11/05.

Abstract
Accuracy of an open source remote eye tracking system and a state-of-the-art commercial eye tracker was measured 4 times during a usability test. Results from 9 participants showed both devices to be fairly stable over time, but the commercial tracker was more accurate with a mean error of 31 pixels against 59 pixels using the low cost system. This suggests that low cost eye tracking can become a viable alternative, when usability studies need not to distinguish between, for instance, particular words or menu items that participants are looking at, but only between larger areas-of-interest they pay attention to.

Keywords
Eye Tracking, Low Cost Eye Tracker, ITU Gaze Tracker, Usability Evaluation, Usability Testing, Performance Evaluation, Accuracy, Visual Attention.

ACM Classification Keywords
H.5.2 User Interfaces: Evaluation/methodology, H.1.2 User/Machine Systems: Human information processing.

Introduction

Eye tracking technology can measure visual attention during realistic use scenarios. Where do users look the most? What catches their attention first? What is not seen at all? This information is highly valuable for many usability practitioners.

Eye tracking has been criticized for being costly and tedious [8]. Newer generations of eye trackers have made it a lot easier to conduct experiments, to analyze data, and to present them in interpretable formats. However, the cost of the state-of-the-art equipment today is between US$20,000 and US$30,000 and while the quality of the systems has improved dramatically over the years, the price has remained more or less constant. Consequently, researchers have been looking for cheaper ways of uncovering user activities. Among the alternatives are think-aloud testing, mouse tracking [1], predictive models, image analysis [2], and recall [5]. However, none of these alternatives are fully accurate for studying visual attention and some may even disrupt the task flow, so there is a need for a low cost alternative to the high-end eye trackers.

The ITU *Gaze Tracker* is an open source eye tracking software [4] that can be used with low cost web cameras or off-the-shelf camcorders. The software tracks the pupil and corneal reflections of the users utilizing infrared illumination. The system makes use of an interpolation-based technique that maps the eye features from the camera images to the point of regard. The first version of the system was introduced and evaluated in 2009 [7]. Results indicated that a low cost eye tracker using a web camera could potentially have the same performance as commercial eye trackers. At that point in time the ITU *Gaze Tracker* was head-mounted and required a web camera to be placed very close to the user's eye, which was very intrusive, and the user had to sit completely still for the eye tracker to work appropriately. Therefore, this version was not suitable as a usability tool.

An updated version of the ITU *Gaze Tracker* software became available in October 2010, which provides the feature of doing remote eye tracking using a web camera modified with a simple change of the camera lens (from the standard wide angle lens into a narrow 16 mm lens). Including two infrared lights, the total cost of the ITU *Gaze Tracker* hardware setup is US$100, and the required software can be downloaded for free [4].

Accuracy is an important factor when using eye trackers for usability testing. If usability practitioners need to tell, for instance, on which particular word in a text the user fixates, the accuracy should be high (i.e. with a mean error of less than 40 pixels which is the size of a 5 character word with a font size of 12). However, many usability tests are reported using broader areas-of-interest, or so called heatmaps, like Poynter Institute's Eye Track studies showing the overall viewing patterns for home pages [6]. For this purpose, the accuracy needs not to be that high in order to tell what navigation element, text block or picture got the most attention.

Even though accuracy can be lower, it is important for all usability testing that accuracy remains stable throughout the test session. If not, the session may have to be interrupted to perform recalibrations, or data may have to be excluded, because of fall-out in data stream or high inaccuracy.

This study aims to investigate how well this new version of the ITU *Gaze Tracker* performs in terms of accuracy and stability during a usability test in comparison with a high-end remote eye tracker, *T60* from Tobii, one of the most common state-of-the-art eye trackers among usability practitioners [9].

Method

Participants

A total of 9 participants (7 male and 2 female), ranging from 16 to 36 years old, took part in this study. They received a gift card of US$20 for participating. None of them had previous experience with eye tracking.

Apparatus

The two different eye tracking devices tested were:

a. ITU *Gaze Tracker*, 30 Hz, monocular eye tracker with free open source software running with a modified Sandberg Nightcam 2 web camera that had a 16 mm lens mounted. Two Sony HVL-IRM infrared lamps were used to illuminate the eyes. The total cost of the hardware is US$100.
b. Tobii *T60*, 60 Hz, binocular eye tracker, price US$30,000.

The study was conducted on a PC with Intel®Core™2 2.6 GHz processor with 3 Gb RAM, and Windows XP SP3 installed. The screen used was the 17” TFT monitor embedded in the *T60* eye tracker. The screen had a resolution of 1280×1024 pixels and was placed 60 cm from the head of the user. Figure 1 shows the experimental setup.

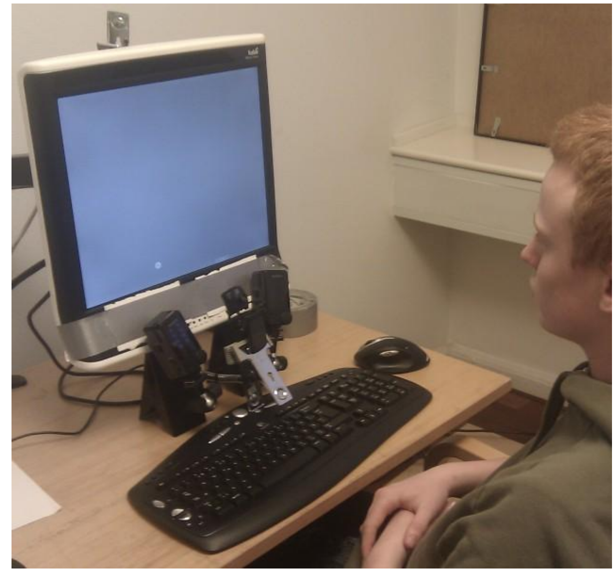

Figure 1: Experimental setup with T60 and ITU Gaze Tracker.

To measure the accuracy, each eye tracker was configured to control the mouse cursor with the eyes. The cursor position was then logged by an accuracy measurement software, *Accuracy Tool* [3]. Tobii *Eye Control Suite* v1.2.2 was used to control the mouse for the *T60*, and for the ITU *Gaze Tracker* the integrated eye control feature was enabled. The mouse cursor was made invisible to avoid distracting the participants. Tobii *Studio* eye tracking software was used to control stimuli presentation and executing *Accuracy Tool* for both systems.

Accuracy Tool required the user to look at a series of 16 targets displayed one by one on a 4x4 grid that covered the whole screen. Formally, this can be expressed as a series of N sample gaze positions, P, that are collected over M target positions, T. Fifty gaze positions were collected at 30 Hz for each of these target positions. Premature samples were avoided with a reaction delay of 600 ms. Samples further than $M \pm 3 \times SD$ were considered as outliers.

The distance between each sample gaze position P_i, and the corresponding target position T_i yields the error in pixels. This was used to calculate the effective accuracy in pixels A_{px} following Equation 1.

$$A_{px} = \sum_{i=1}^{M} \left(\frac{\sum_{j=1}^{N} \frac{\|T_i - P_{i,j}\|}{N}}{M} \right) \qquad (1)$$

Usually accuracy of eye trackers is given in degrees of visual angle. The physical size of a pixel S and the distance from user to screen D need to be known. The effective accuracy in degrees was calculated by means of Equation 2.

$$A^{\circ} = \frac{180}{\pi} \cdot \frac{1}{2} \cdot \tan^{-1} \left(\frac{\frac{A_{px}}{2} \cdot S}{D} \right) \qquad (2)$$

Procedure
Each participant was tested with both devices, where 4 participants were tested first with the ITU *Gaze Tracker*, and 5 participants first with the *T60*. The total test session lasted approximately 40 minutes. Before the test participants were encouraged to take a comfortable position in front of the computer.

For each device, participants were initially calibrated according to the standard procedures that each system has. This requires the user to look at 9 points on the screen, and it takes less than one minute to perform. Then, the first of four accuracy measurements was collected with *Accuracy Tool*. Immediately after this measure the participants were automatically sent to a news website, where they were asked to browse freely for two minutes. This procedure was repeated to collect a total of 4 accuracy samples, with 3 times two minutes of browsing (on different news websites) between each sample. The procedure was run automatically using Tobii Studio, which controlled the stimuli presentation, randomizing also the presentation order of the news websites among participants, and executing the Accuracy Tool.

Design
The experiment employed a within-subjects, 2x4 repeated measures design with the following factors and levels:

- Eye tracking device: *ITU Gaze Tracker, T60*
- Accuracy sample: 1, 2, 3, 4

The dependent variable was Error, which was measured in pixels. Pixels represent the most relevant measure for the practical implications when using eye trackers for usability studies; however, for comparison purposes with other eye tracking devices, we also report the mean Error in degrees. We investigated the effect of eye tracking device over time on the error during usability task solving. In total, there were 9 participants x 2 eye tracking devices x 4 accuracy samples x 16 target positions x 50 gaze positions = 57,600 gaze positions.

Results and Discussion

Analysis was performed using a 2x4 ANOVA with device and accuracy sample as the independent variables. 401 samples out of the 57,600 were considered outliers and removed from the analysis.

The analysis show an effect of eye tracking device that is statistically significant for Error between the ITU *Gaze Tracker* (M = 59.0, SD = 11.2) and the *T60* (M = 30.9, SD = 6.59); $F_{1,8} = 66.74$, $p < .0001$. For the ITU *Gaze Tracker* the mean Error in degrees was 1.48 (SD = 0.58), and for the *T60* it was 0.77 (SD = 0.35).

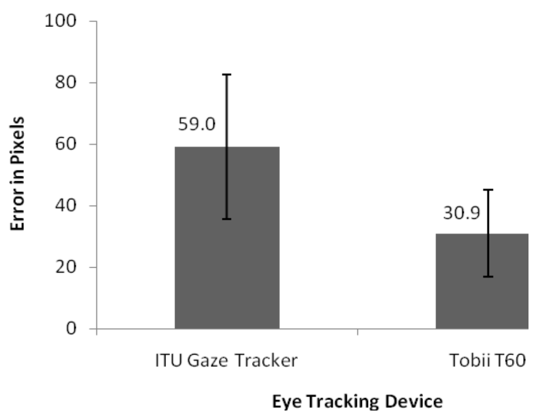

Figure 2: Mean error in pixels by eye tracking device

Figure 2 shows that the state-of-the-art eye tracker *T60* is approximately twice as accurate as the low-cost open source ITU *Gaze Tracker*.

The overall effect of accuracy sample was not significant for Error, $F_{3,24} = 2.47$, $p > .05$. Also, when analyzing each device separately, the effect of accuracy sample was not statistically significant, neither for the *T60* ($F_{3,24} = 1.11$, $p > .05$) nor the ITU *Gaze Tracker* ($F_{3,24} = 2.94$, $p > .05$).

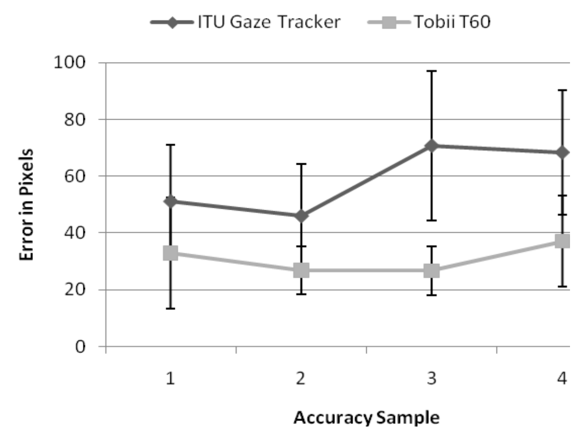

Figure 3: Error in pixels by accuracy sample for each device

These results indicate that both eye trackers had a relatively stable accuracy over the time, where a test session lasted approximately 20 minutes for each device. Figure 3 suggests that the Error increased between the first and the last accuracy sample for both devices, but this difference was not significant for either device.

The results of this study show a difference in accuracy between the eye tracking devices. However, the ITU *Gaze Tracker* is able to track eye movements with a mean error of 59 pixels, which in many cases could be sufficient for usability purposes. To partly overcome the lower accuracy of ITU *Gaze Tracker* one might consider collecting data for more participants than needed for a particular study, and then exclude all datasets from participants with a mean error above a certain threshold in a post-session accuracy test. Figure 4 illustrates the mean error when using the different eye tracking systems.

Figure 4: The potential mean error of the Tobii *T60* (red inner circle) and the ITU *Gaze Tracker* (yellow outer circle) illustrated on a screenshot from NY Times website.

The study shows that both systems are fairly stable over time, which is important when applying the eye trackers for usability testing. This indicates that the ITU *Gaze* Tracker could become a viable alternative to the commercial systems. However, in the current version it requires technical knowledge and skill to set up.

Our plans for future work includes making a large-scale longitudinal study, with a high number of participants, longer test session, and higher number of accuracy samples. Specifically we would like to explore how continuous recalibrations and repositioning of the participants, as well as how using different types of chairs, will affect the accuracy of the systems. In this study we would also like test various hardware setups for the ITU *Gaze Tracker* (e.g. better cameras), and different algorithms for calculating the point-of-regard.

Acknowledgements
We would like to thank EYEFACT for supporting the experiment, and the open source community for their help with improving the ITU *Gaze Tracker*.

References

[1] Bednarik, R. and Tukiainen, M. Visual attention tracking during program debugging. In *Proc. NordiCHI 2004*. ACM (2004), 331-334.

[2] Feng-GUI Science, How does it work? http://www.feng-gui.com/science.htm.

[3] ITU Gaze Group, Downloads. http://www.gazegroup.org/downloads/.

[4] ITU Gaze Group, Gaze Tracking Library. http://sourceforge.net/projects/gazetrackinglib/.

[5] Johansen, S.A. and Hansen, J.P. Do we need eye trackers to tell where people look? In Proc. *CHI 2006*, ACM, Press (2006), 923-928.

[6] Poynter Institute, EYETRACK III, Viewing Patterns for Homepages. http://www.poynterextra.org/eyetrack2004/viewing.htm.

[7] San Agustin, J., Skovsgaard, H., Hansen, J.P., and Hansen, D.W. Low-cost gaze interaction: ready to deliver the promises. In *Proc. CHI 2009*, ACM Press (2009), 4453–4458.

[8] Schnipke, S.K. and Todd, M.W. Trials and Tribulations of Using an Eye-tracking System, In *Proc. CHI 2000*, ACM Press (2000), 273-274.

[9] Tobii Technology. http://www.tobii.com.