The Pupillometric Precision of a Remote Video Eye Tracker

Jeff Klingner*
Stanford University

Abstract

To determine the accuracy and precision of pupil measurements made with the Tobii 1750 remote video eye tracker, we performed a formal metrological study with respect to a calibrated reference instrument, a medical pupillometer. We found that the eye tracker measures mean binocular pupil diameter with precision 0.10 mm and mean binocular pupil dilations with precision 0.15 mm.

CR Categories: B.4.2 [Hardware]: Input/Output and Data Communications—Input/Output Devices ; I.4.9 [Computing Methodologies]: Image Processing and Computer Vision—Applications; J.4 [Computer Applications]: Social and Behavioral Sciences—Psychophysiology

Keywords: eye tracking, pupil, pupillometry, metrology

1 Introduction

In order to measure gaze direction, most eye trackers gather high-resolution images of the pupil. These images enable an important secondary application of eye trackers: measurement of pupil diameter. Short-term changes in pupil diameter are linked to a variety of internal cognitive processes [Andreassi 2006], so the high-frequency, high-precision measurement of pupil diameter enabled by eye trackers has applications in fields such as learning [Paas and Van Merriënboer 1994], psychopathology [Steinhauer and Hakerem 1992], and human-computer interaction [Pomplun and Sunkara 2003].

Most eye trackers used in cognitive pupillometry use head-mounted cameras or chin rests, because a fixed camera-pupil distance enables good pupillometric precision. In contrast, remote eye trackers, which usually devote fewer pixels to each pupil and must correct for variations in the camera-pupil distance, exhibit worse precision. However, there are some applications which require remote, free-head eye tracking or pupillometry, such as studies with infants [Chatham et al. 2009] or investigations of small changes in anxiety, distraction, or mental effort [Porter et al. 2007]. Quantifying the accompanying loss of precision is important, both to guide equipment choices and to determine the number of participants and trials required to measure a given magnitude pupillary response using a remote eye tracker.

*e-mail: klingner@stanford.edu

ETRA 2010, Austin, TX, March 22 – 24, 2010.
© 2010 ACM 978-1-60558-994-7/10/0003 $10.00

2 Study Description

2.1 Evaluated Instrument

We evaluated the pupillometric performance of the Tobii 1750 remove video eye tracker [Tobii Technologies, Inc. 2007], shown in Figure 1(a). The Tobii 1750 measures the size of a pupil by fitting an ellipse to the image of that pupil under infrared light, then converting the width of the major axis of that ellipse from pixels to millimeters based on the measured distance from the camera to the pupil. According to Tobii, errors in this measurement of camera–pupil distance cause measurements of pupil diameter to have errors of up to 5% for fixed-size pupils [Tobii Technologies, personal communication].

This 5% figure is a good start, but for guiding experimental design, we need to extend it by a) distinguishing bias and precision components of the error, and b) determining the average-case, rather than worst-case performance, because it is usually the averages of many repeated pupil measurements which are used to quantify task-evoked pupillary responses [Beatty and Lucero-Wagoner 2000].

2.2 Reference Instrument

The reference instrument is a Neuroptics VIP-200 ophthalmology pupillometer, shown in Figure 1(b). The Neuroptics VIP-200 records two seconds of video of the pupil, then reports the mean and standard deviation of the pupil's diameter over those two seconds. This instrument has a precision of about 0.05 mm for these two-second averages, and was calibrated to zero bias when it was manufactured [Neuroptics, Inc. 2008].

(a) Tobii 1750 (b) Neuroptics VIP-200

Figure 1: *The eye tracker and reference pupillometer*

2.3 Procedure

Three volunteers participated in the metrology study, which took place in an eye clinic exam room. We took 336 double

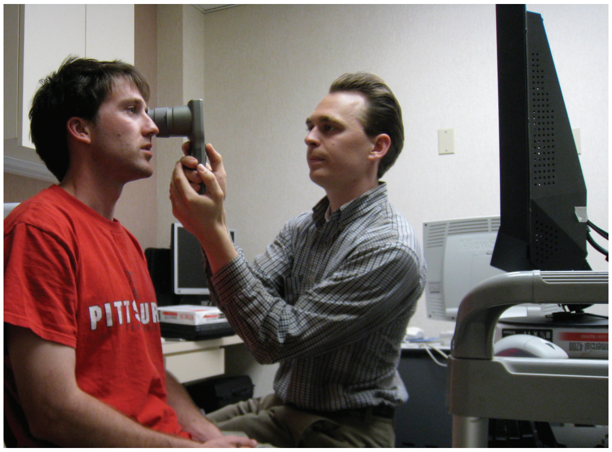

Figure 2: *Metrology study arrangement. An investigator is measuring the participant's left pupil using the reference pupillometer while the eye tracker simultaneously measures his right pupil.*

measurements in which we measured participants' pupils using the eye tracker and the pupillometer simultaneously. Because the pupillometer covers the eye it measures, we could not conduct simultaneous measurements of the same eye using both instruments, so for each double measurement, the pupillometer measured one of the participant's pupils while the eye tracker measured the other (Figure 2). The metrological validity of this study is therefore based on the strong correlation between the diameters of the left and right pupils [Loewenfeld 1999]. Measurements taken with the eye tracker were averages over the 100 camera frames gathered in the same two-second measurement window used by the reference pupillometer.

The measurements were conducted under various lighting conditions so that our measurements would span a variety of pupil states: half under normal room lighting and half under dim lighting, where a third of the time we switched the lights on or off during the few seconds between successive double measurements. In all trials, subjects looked at a small fixation target at the center of the eye tracker's screen, which was otherwise filled with 64 cd/m^2 medium gray. We excluded 120 measurements in which we did not get a clean reading with the pupillometer and 10 measurements in which we did not get a clean reading with the eye tracker, leaving 206 successful double measurements, analyzed below.

3 Metrology

We present two different metrological analyses of these double measurements: the first, based on pupil diameters, is simpler and can use all of the data but is limited by strong assumptions. The second analysis, based on dilations, uses weaker assumptions but is restricted to a subset of the available data.

3.1 Pupil Diameter Metrology

For both instruments, we model the measurement error as being additive and normally distributed:

$$\Pi = \pi + \epsilon \qquad \epsilon \sim N(\mu, \sigma)$$

where π is the diameter of the pupil, Π is the measurement of that diameter, and ϵ is the measurement error. π and ϵ are random variables that take on new values for each measurement. Each instrument's *bias* is the the fixed component of the measurement error μ, its *accuracy* is the magnitude of the bias $|\mu|$, and its *precision* is the standard deviation of the measurement error σ. For the reference pupillometer (pm), $\mu[\epsilon_{pm}] = 0$ mm and $\sigma[\epsilon_{pm}] = 0.05$ mm, according to information provided by its manufacturer. For the eye tracker (et), the parameters of the measurement error distribution $\mu[\epsilon_{et}]$ (bias) and $\sigma[\epsilon_{et}]$ (precision) are what we are trying to determine.

We can estimate these parameters by analyzing the differences between simultaneous measurements made with the eye tracker and the pupillometer:

$$\begin{aligned} \Pi_{et} - \Pi_{pm} &= (\pi_{et} + \epsilon_{et}) - (\pi_{pm} + \epsilon_{pm}) \\ &= \pi_{et} - \pi_{pm} + \epsilon_{et} - \epsilon_{pm} \end{aligned}$$

This is an equation of random variables. Considering the variance of each side:

$$\begin{aligned} \sigma^2[\Pi_{et} - \Pi_{pm}] &= \sigma^2[\pi_{et} - \pi_{pm} + \epsilon_{et} - \epsilon_{pm}] \\ \sigma^2[\Pi_{et} - \Pi_{pm}] &= \sigma^2[\pi_{et} - \pi_{pm}] + \sigma^2[\epsilon_{et}] + \sigma^2[\epsilon_{pm}] \\ \sigma^2[\Pi_{et} - \Pi_{pm}] &= \sigma^2[\pi_{et} - \pi_{pm}] + \sigma^2[\epsilon_{et}] + \sigma^2[\epsilon_{pm}] \quad (1) \\ \sigma^2[\Pi_{et} - \Pi_{pm}] &= \sigma^2[\epsilon_{et}] + \sigma^2[\epsilon_{pm}] \\ \sigma[\epsilon_{et}] &= \sqrt{\sigma^2[\Pi_{et} - \Pi_{pm}] - \sigma^2[\epsilon_{pm}]} \quad (2) \end{aligned}$$

The relationship in Equation 2 gives us a way to estimate the precision of the eye tracker based on the known precision of the reference pupillometer $\sigma[\epsilon_{pm}]$ and the variance in the differences between the simultaneous measurements $\sigma^2[\Pi_{et} - \Pi_{pm}]$. Similarly, we can compute the bias of the eye tracker based on the mean of those differences:

$$\begin{aligned} \mu[\Pi_{et} - \Pi_{pm}] &= \mu[\pi_{et} - \pi_{pm} + \epsilon_{et} - \epsilon_{pm}] \\ \mu[\Pi_{et} - \Pi_{pm}] &= \mu[\pi_{et} - \pi_{pm}] + \mu[\epsilon_{et}] - \mu[\epsilon_{pm}] \\ \mu[\Pi_{et} - \Pi_{pm}] &= \mu[\pi_{et} - \pi_{pm}] + \mu[\epsilon_{et}] + \mu[\epsilon_{pm}] \quad (3) \\ \mu[\epsilon_{et}] &= \mu[\Pi_{et} - \Pi_{pm}] \quad (4) \end{aligned}$$

Substituting the mean and variance of the actually observed differences $\Pi_{et} - \Pi_{pm}$ in Equations 4 and 2, the eye tracker's pupillometric bias is 0.11 mm, and its precision is 0.38 mm.

These figures are misleading, however, because the bias and precision of the eye tracker varied substantially between the three participants and between the two eyes of each participant. Figure 3 shows the results of all 206 successful simultaneous measurements and illustrates this inter-subject variation. For each eye individually, the measurement error has a much narrower spread, but the average is wrong by as much as 0.67 mm. The accuracy and precision varies from eye to eye like this because the eye tracker's pupil measurements depend on its estimate of the camera-pupil distance, which is affected by errors in the eye tracker's calibration to each eye's corneal shape.

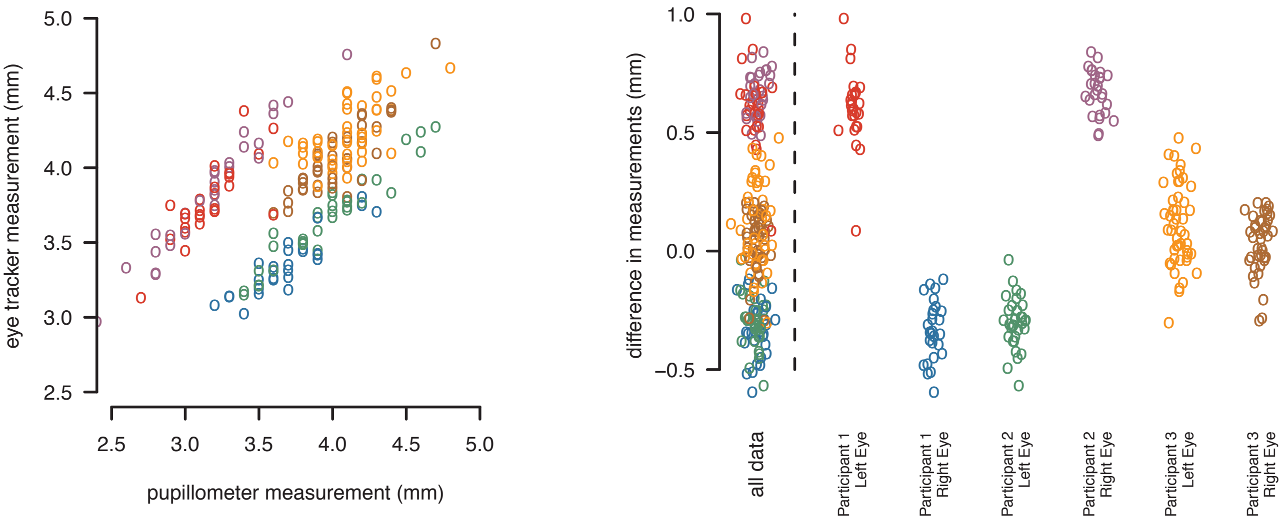

Figure 3: *The left graph shows the raw data of the metrology study, with each point representing a double measurement* (Π_{pm}, Π_{et})*. Data from each participant and each subject are plotted in a different color. The right chart shows the differences between the eye tracker and pupillometer measurements,* $\Pi_{et} - \Pi_{pm}$*, broken down by study participant and eye, showing how the eye tracker's pupillometric bias varies for each eye.*

When Equations 4 and 2 are applied to the data from each eye separately, we find an average bias of 0.34 mm (worse than the overall 0.11 mm) and an average precision of 0.12 mm (better than the overall 0.38 mm). Because it is differences in measurements for the same eye (dilations) that form the basis of most experimental use of pupillometry [Beatty and Lucero-Wagoner 2000], and because pupillometric experiments are usually conducted with several participants, these per-eye results for the eye tracker's bias and precision are the most relevant and are the ones summarized in Table 1.

In Equation 3 of the derivation for accuracy, the term $\mu[\pi_{et} - \pi_{pm}]$ was assumed to be zero. We ensured this zero mean left-right difference in pupil size by counterbalancing which of the two eyes was measured with which instrument within the trials for each participant.

Similarly, the cancelation in Equation 1 of the derivation for precision assumes that term $\sigma^2[\pi_{et} - \pi_{pm}]$ is zero. This assumption, that the difference in size between participants' left and right pupils is constant throughout the study, is much stronger. Judging from pupil data we've recorded in a variety of studies, it is true over short periods of time (a few minutes) but can sometimes drift over the 15–20 minutes it takes to make the measurements of each participant. Violations of this assumption would lead to an underestimate of the average error the eye tracker. A more conservative analysis, based on differences in short-term *dilations* measured by each instrument, provides an alternative estimate of the eye tracker's precision.

3.2 Pupil Dilation Metrology

We can determine the pupillometric precision of the eye tracker using differences in measurements of dilations rather than absolute pupil diameters. Using $\delta = \pi_2 - \pi_1$ to denote the dilation of the pupil from time 1 to time 2 and $\Delta = \Pi_2 - \Pi_1$ to denote the measurement of that dilation,

$$\begin{aligned}
\Delta_{et} - \Delta_{pm} &= (\Pi_{et_2} - \Pi_{et_1}) - (\Pi_{pm_2} - \Pi_{pm_1}) \\
&= [(\pi_{et_2} + \epsilon_{et_2}) - (\pi_{et_1} + \epsilon_{et_1})] \\
&\quad - [(\pi_{pm_2} + \epsilon_{pm_2}) - (\pi_{pm_1} + \epsilon_{pm_1})] \\
&= (\pi_{et_2} - \pi_{et_1}) - (\pi_{pm_2} - \pi_{pm_1}) \\
&\quad + \epsilon_{et_2} - \epsilon_{et_1} + \epsilon_{pm_1} - \epsilon_{pm_2} \\
&= (\delta_{et} - \delta_{pm}) + \epsilon_{et_2} - \epsilon_{et_1} + \epsilon_{pm_1} - \epsilon_{pm_2} \qquad (5)
\end{aligned}$$

As before, now considering the variance of the random variables on each side of Equation 5:

$$\begin{aligned}
\sigma^2[\Delta_{et} - \Delta_{pm}] &= \sigma^2[(\delta_{et} - \delta_{pm}) + \epsilon_{et_2} - \epsilon_{et_1} + \epsilon_{pm_1} - \epsilon_{pm_2}] \\
&= \sigma^2[\delta_{et} - \delta_{pm}] + \sigma^2[\epsilon_{et_2}] \qquad (6) \\
&\quad + \sigma^2[\epsilon_{et_1}] + \sigma^2[\epsilon_{pm_1}] + \sigma^2[\epsilon_{pm_2}] \\
&= \sigma^2[\epsilon_{et_2}] + \sigma^2[\epsilon_{et_1}] + \sigma^2[\epsilon_{pm_1}] + \sigma^2[\epsilon_{pm_2}] \\
&= 2\sigma^2[\epsilon_{et}] + 2\sigma^2[\epsilon_{pm}] \qquad (7) \\
\sigma^2[\epsilon_{et}] &= \tfrac{1}{2}\sigma^2[\Delta_{et} - \Delta_{pm}] - \sigma^2[\epsilon_{pm}] \\
\sigma[\epsilon_{et}] &= \sqrt{\tfrac{1}{2}\sigma^2[\Delta_{et} - \Delta_{pm}] - \sigma^2[\epsilon_{pm}]} \qquad (8)
\end{aligned}$$

The cancellation in Equation 6 is based on the assumption that the difference between the left eye's dilation and the right eye's dilation is constant over a short period of time. We observed this fact in an earlier study we conducted on lateralized pupillary responses, in which we tried several stimulus based ways of inducing different dilations in subjects' two pupils but never succeeded in causing any significant left-right differences. We abandoned the effort after learning that the neuroanatomy of pupil size regulation renders such differences extremely unlikely [Loewenfeld 1999]. Step 7 relies on the assumption that the bias of the measurement error is stable over time for both instruments ($\sigma[\epsilon_{pm_1}] = \sigma[\epsilon_{pm_2}]$ and $\sigma[\epsilon_{et_1}] = \sigma[\epsilon_{et_2}]$).

assumption	data	eye tracker diameter accuracy per-eye (mm)	eye tracker precision (mm)			
			pupil diameter		dilation magnitude	
			monocular	binocular mean	monocular	binocular mean
The difference in size between the left and right pupils is constant over the study.	206 double measurements	0.34	0.12	0.08	0.17	0.12
The difference between the left eye's dilation and the right eye's dilation is constant over 30 sec.	84 pairs of double measurements	NA	0.15	0.10	0.21	0.15

Table 1: *Summary of the Tobii 1750's pupillometric performance. Figures for monocular diameter accuracy and precision are the results of the metrological analysis above. Other figures in the table were then derived from these primary results. Dilation measurement precision is larger (worse) by a factor of $\sqrt{2}$, because it is based on the difference of two diameter measurements. When both eyes are measured and averaged, the precision in the estimate of their mean dilation (or diameter) improves by a factor of $\sqrt{2}$ over the monocular case.*

Among the 206 successful double measurements, there are 84 pairs of double measurements that took place within 30 seconds of each other. That is, there were 84 dilations with duration less than 30 seconds with starting diameters and ending diameters that were both measured with the two instruments simultaneously. Substituting the observed $\Delta_{et} - \Delta_{pm}$ in Equation 8 gives pupillometric precision of the eye tracker as 0.15 mm, slightly worse than the diameter-based precision of 0.12 mm.

4 Summary

We made many measurements of pupil size using a remote video eye tracker and medical pupillometer simultaneously. We analyzed these simultaneous measurements in two ways to assess the pupillometric performance of the Tobii 1750 remote video eye tracker.

The first analysis, diameter-based metrology (Section 3.1), provided an estimate of the eye tracker's pupillometric accuracy and—via a relatively strong assumption—a lower bound on the eye tracker's pupillometric precision. The second analysis, dilation-based metrology (Section 3.2), provided an alternative estimate of precision relying on fewer assumptions but also with less applicable data. The results of both analyses are summarized in Table 1, together with the resultant derived precision for binocular and dilation measurements.

Acknowledgements

This work was funded by the Stanford Regional Visual Analytics Center, through the U.S. Department of Energy's Pacific Northwest National Laboratory. Our eye tracker was funded by the Stanford MediaX project and the Stanford School of Engineering.

References

ANDREASSI, J. L. 2006. Pupillary response and behavior. In *Psychophysiology: Human Behavior and Physological Response*, 5th ed. Routledge, ch. 12, 289–307.

BEATTY, J., AND LUCERO-WAGONER, B. 2000. The pupillary system. In *Handbook of Psychophysiology*, J. T. Cacioppo, L. G. Tassinary, and G. Berntson, Eds., 2nd ed. Cambridge University Press, 142–162.

CHATHAM, C. H., FRANK, M. J., AND MUNAKATA, Y. 2009. Pupillometric and behavioral markers of a developmental shift in the temporal dynamics of cognitive control. *Proceedings of the National Academy of Sciences 106*, 14, 5529–5533.

LOEWENFELD, I. 1999. *The Pupil: Anatomy, Physiology, and Clinical Applications*, 2nd ed., vol. 1. Butterworth-Heinemann, Oxford, UK.

NEUROPTICS, INC., 2008. Instruction manual, VIP-200 pupillometer, revision A.

PAAS, F., AND VAN MERRIËNBOER, J. 1994. Instructional control of cognitive load in the training of complex cognitive tasks. *Educational Psychology Review 6*, 4 (Dec.), 351–371.

POMPLUN, M., AND SUNKARA, S. 2003. Pupil dilation as an indicator of cognitive workload in human-computer interaction. In *Proceedings of the International Conference on HCI*.

PORTER, G., TROSCIANKO, T., AND GILCHRIST, I. D. 2007. Effort during visual search and counting: Insights from pupillometry. *The Quarterly Journal of Experimental Psychology 60*, 2, 211.

STEINHAUER, S. R., AND HAKEREM, G. 1992. The pupillary response in cognitive psychophysiology and schizophrenia. *Annals of the New York Academy of Sciences 658*, Psychophysiology and Experimental Psychopathology: A Tribute to Samuel Sutton, 182–204.

TOBII TECHNOLOGIES, INC., 2007. Tobii 1750.