Validation of a prototype hybrid eye-tracker against the DPI and the Tobii Spectrum

Kenneth Holmqvist
Nicolaus Copernicus University
Torun, Poland;
University of the Free State
Bloemfontein, South Africa;
Universität Regensburg
Regensburg, Germany
kenneth.holmqvist@ur.de

Saga Lee Örbom
Universität Regensburg
Regensburg, Germany
saga.oerbom@stud.uni-regensburg.de

Michael Miller
EyeWay Vision Ltd.
Tel Aviv, Israel
michael.m@eyeway-vision.com

Albert Kashchenevsky
EyeWay Vision Ltd.
Tel Aviv, Israel
Albert.K@eyeway-vision.com

Mark Shovman
EyeWay Vision Ltd.
Tel Aviv, Israel
mark.shovman@eyeway-vision.com

Mark W. Greenlee
Universität Regensburg
Regensburg, Germany
mark.greenlee@ur.de

ABSTRACT

We benchmark a new hybrid eye-tracker system against the DPI (Dual Purkinje Imaging) tracker and the Tobii Spectrum in a series of three experiments. In a first within-subjects battery of tests, we show that the precision of the new eye-tracker is much better than that of both the DPI and the Spectrum, but that accuracy is not better. We also show that the new eye-tracker is insensitive to effects of pupil contraction on gaze direction (in contrast to both the DPI and the Spectrum), that it detects microsaccades on par with the DPI and better than the Spectrum, and that it can possibly record tremor. In the second experiment, sensors of the novel eye-tracker were integrated into the optical path of the DPI bench. Simultaneous recordings show that saccade dynamics, post-saccadic oscillations and measurements of translational movements are comparable to those of the DPI. In the third experiment, we show that the DPI and the new eye-tracker are capable of detecting 2 arcmin artificial-eye rotations while the Spectrum cannot. Results suggest that the new eye-tracker, in contrast to video-based P-CR systems [Holmqvist and Blignaut 2020], is suitable for studies that record small eye-movements under varying ambient light levels.

CCS CONCEPTS

• **Hardware** → *Hardware reliability screening*; *Emerging tools and methodologies*; • **Computer systems organization** → **Sensors and actuators**; • **Human-centered computing** → **Ubiquitous and mobile computing design and evaluation methods**.

KEYWORDS

hybrid eye-tracker, DPI, video-based P-CR, data quality, microsaccades

ACM Reference Format:
Kenneth Holmqvist, Saga Lee Örbom, Michael Miller, Albert Kashchenevsky, Mark Shovman, and Mark W. Greenlee. 2020. Validation of a prototype hybrid eye-tracker against the DPI and the Tobii Spectrum. In *Symposium on Eye Tracking Research and Applications (ETRA '20 Full Papers), June 2–5, 2020, Stuttgart, Germany*. ACM, New York, NY, USA, 9 pages. https://doi.org/10.1145/3379155.3391330

1 INTRODUCTION

Video-based eye-trackers using the pupil and corneal reflection (the P-CR method) today entirely dominate both the research and the applications markets. When P-CR eye-trackers first appeared in the late 1960s, the primary design goal was to *relax requirements on head movement restriction* [Holmqvist and Andersson 2017, p. 77-80]. Exhibiting excellent data quality and having few artefacts was not a priority at the time. Over the years, however, as camera technology and image processing algorithms improved, it became a common belief that video-based eye-tracking is the method also for the most demanding future eye-tracking research and applications.

The view that P-CR eye-trackers can be trusted to deliver reliable data in demanding research has been challenged by several recent studies of data quality in eye-trackers. Some of these findings have been frequently replicated, for instance the pupil artefacts on accuracy [Drewes 2014; Drewes et al. 2012], the post-saccadic oscillations[1] [Hooge et al. 2016; McCamy et al. 2015], and the effects of movement in the headboxes [Niehorster et al. 2017]. New artefacts are being reported, for instance that P-CR eye-tracking results in amplitude mis-measurements of saccades with amplitudes up to 2° [Holmqvist and Blignaut 2020], comparing 11 eye-trackers. Holmqvist [2015] provide an in-depth comparison of 12 eye-trackers along several criteria. These studies show that the choice to use cameras, as well as the P-CR method itself, comes with several drawbacks.

---

[1] Part biology, part artefact.

ETRA '20 Full Papers, June 2–5, 2020, Stuttgart, Germany

ACM ISBN 978-1-4503-7133-9/20/06…$15.00
https://doi.org/10.1145/3379155.3391330

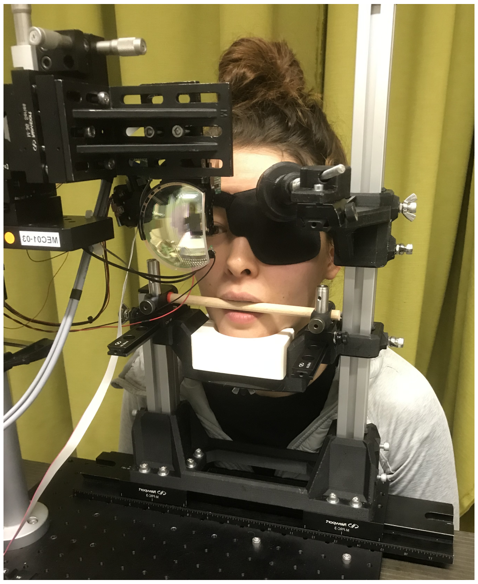

Figure 1: The EWET1 prototype with stabilized participant.

In the 1990s, camera-based P-CR eye-trackers replaced two technologies that, in terms of data quality, are superior: scleral search coils and the Dual-Purkinje-Imaging (DPI) system. The number of these superior eye-trackers is dwindling owing to the higher purchasing and maintenance costs, the learning threshold for usage and the obsolete electro-mechanical technologies. Video-based eye-tracking has been packaged to appear more user friendly and to require less maintenance. However, despite the fast development of new video-based eye-trackers, and marketing of the established ones, their known issues make it an open question whether P-CR eye-trackers will (or should) dominate the future of eye-tracking.

It is clear however, that in the middle-long term, the DPI will cease to be an option for researchers who are interested in correctly measuring small eye-movements and accurately determining gaze. New eye-tracking methods are also needed for the development of virtual and augmented reality devices. To meet with this demand, in this paper, we present a prototype hybrid (analogue/digital) eye-tracker and benchmark it against the analogue DPI [Crane and Steele 1985] and the digital/camera-based Tobii Spectrum (www.tobii.com) in five different paradigms that test its performance.

2 THE EWET1 PROTOTYPE EYE-TRACKER

The tested eye-tracker, referred to here as the EyeWay Eye-Tracker 1 (EWET1), is a hybrid system combining signals from an optoelectronic, MEMS-based, tracker of corneal reflection (the **CR signal** at 4000 Hz), and a dual camera-based tracker of 3D translational movements of the eye and head relative to the device (the **TR signal** at 120 Hz).

From these two sensor signals, the direction of the optical axis (the **GZ signal** at 4000 Hz) is calculated. The exact process by which CR and TR are combined to GZ has not been revealed by EyeWay Vision Ltd. at the time of this study, but involves upsampling the TR signal followed by geometrical fusion of the two signals. The calculation results in an adjusted line of the optical axis (uncalibrated GZ). A standard 9-point polynomial calibration protocol is used in this study, which also results in a GZ roughly calibrated for the visual axis.

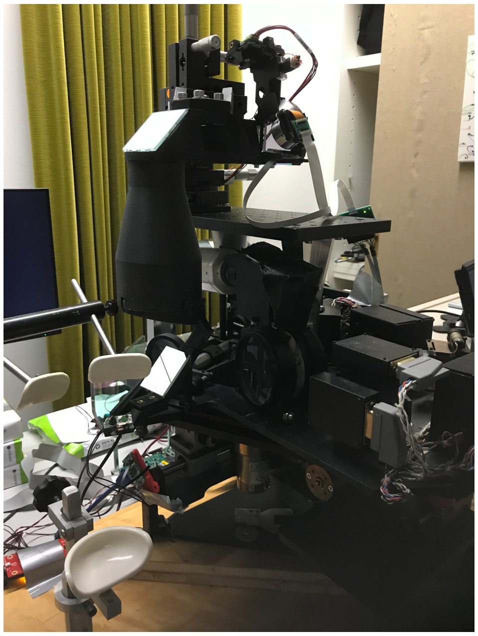

Figure 2: The EWET1 built on top of the DPI, positioned such that simultaneous co-recording is possible.

The very high sampling frequency and low power consumption of an optoelectronic CR tracker is combined with an accurate TR signal at a much lower sampling rate to compensate for the smaller and slower translational movements. This principle is similar to combining the signals of pupil and CR in P-CR trackers, or 1st and 4th Purkinje images in the DPI, but differs in two fundamental ways. Firstly, tracking eye features which are unrelated to the pupil should reduce artefacts on the estimates of gaze due to variations in ambient lighting [Drewes 2014]. Secondly, employing an optoelectronic CR measurement signal avoids the subpixel estimation problem that is now believed to cause mis-measurements of small eye-movements [Holmqvist and Blignaut 2020]. Furthermore, a different choice of sensor features can help resolve post-saccadic oscillations [Nyström, Hooge, and Holmqvist 2013], which are known to be troublesome to P-CR and DPI trackers alike.

Each of the three signals can be used separately, for relevant analyses. CR can be used for event detection of eye movements; GZ for applications where gaze location is important; and TR for detection of head-movement or slippage of a device on the head.

Figure 1 shows this prototype system in its state of September 2019, with an optional bite bar to make sure we can control for the effect of small head movements.

The system has been successfully tested for eye safety by an independent accredited testing house (The Standards Institute of

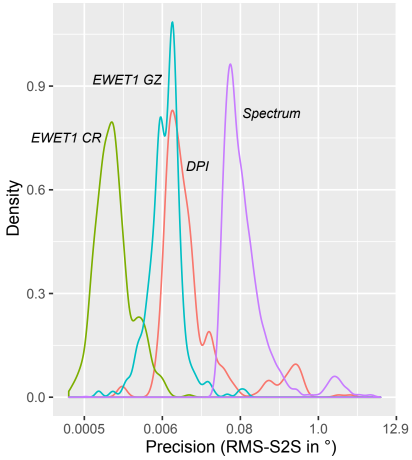

Figure 3: Density plot of the RMS-S2S precision in degrees for each of the three eye-trackers. For the EWET1, the CR and the combined GZ precision distributions are plotted separately. Note the log-scale on the horizontal axis.

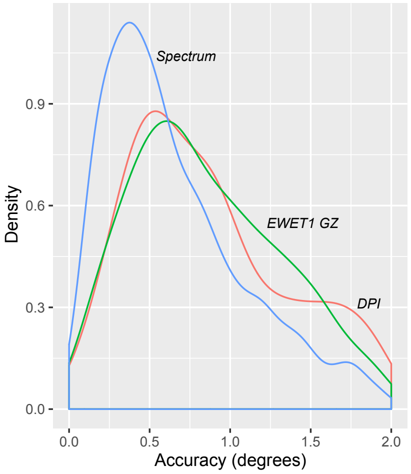

Figure 4: Density plot of the accuracy in degrees for each of the three eye-trackers. For the EWET1, the GZ signal was used.

Israel). We have applied for ethical clearance for this research at the ethics board of the University of Regensburg (number 19-1618-101).

The EWET1 eye-tracker tracks the optical axis and needs to be calibrated for the visual axis (similar to a DPI), and for the relative position of the device to the stimulus display. In this study, we used the same 9-point polynomial calibration algorithm for the EWET1 and for the DPI, and for fair measure, we also used this same calibration algorithm to recalibrate the Tobii Spectrum data.

3 PARTICIPANTS: EXPERIMENTS 1 AND 2

A total of 28 participants (15 M, 13 F, mean age 33.79 (SD 11.69)) were recruited for this study. The data from 4 further participants had to be excluded from analysis, as they met the exclusion criteria (taking medication that might affect their eye-movements) or due to technical issues on one of the trackers. Written informed consent was obtained from each participant after they had been briefed on the nature of the study. They were compensated for participation.

4 RECORDING ENVIRONMENT AND EYE-TRACKERS

We set up the EWET1 120/4000 Hz, the Tobii Spectrum 600 Hz and the DPI Gen 5.5 at 4000 Hz in a dark, window-less and soundproofed room with no noticeable vibrations. The EWET1 existed in two versions: one standalone, which we placed on a 300 kg antivibration workbench (Figure 1), and one version of the EWET1 with the same sensors but a slightly different optic path that was added to the DPI to allow for co-recording the same (right) eye on both systems (Figure 2).

The DPI manufacturer and maintenance specialist Warren Ward (www.wardtek.com) was present during all set-up and recordings to ensure that the DPI was operating at its very best.

The entirely camera-based Tobii Spectrum was chosen because it is a high-end video-based P-CR eye-tracker, the flagship eye-tracker of the Tobii company (www.tobii.com). Many details of its operation are not public. The operation of the analogue DPI is fully published by Crane and Steele [1985]. The DPI was chosen because of its superior data quality and its excellent reputation in the scientific community.

5 EXPERIMENT 1: COMPARISON OF EWET1, TOBII SPECTRUM AND THE DPI

The purpose of Experiment 1 was to compare the standalone EWET1 of Figure 1 against two high-end eye-trackers, the DPI Gen 5.5 and the Tobii Spectrum, using a within-subjects comparison modelled after and extending Holmqvist [2015]. We counterbalanced the order in which participants were recorded on the eye-trackers.

5.1 Stimuli and instructions

All stimuli were dot-in-a-cross patterns [Thaler, Schütz, Goodale, and Gegenfurtner 2013] of a 0.3° diameter, displayed on a black background. The monitor on which the stimulus was displayed was placed at a 70 cm viewing distance. On the Tobii Spectrum we used the built-in 60 Hz monitor, while the other eye-trackers were recorded with stimuli presented on HP25x gaming displays at 144 Hz. The same custom-built experimental software was used with all three systems, and all stimuli were scaled to the size and gamma profile of the specific monitor.

Each recording session comprised of 5 trials, presented in a fixed order. Between trials there was a 4-second rest period, during which the participants were instructed to blink and rest their eyes, but to keep their head still. The trials were:

(1) Calibration: one point at a time was shown for 2 seconds, in a 9-point calibration pattern, covering ±4° range vertically and horizontally.
(2) Long fixation: a single fixation point was shown for 10 seconds. Dependent variables are number and amplitude of microsaccades. We also looked for tremor. This is a very

common task in the study of such intra-fixational eye movements.

(3) Grid: one point at a time was shown for 2 seconds, in a 5 x 5 grid, equidistant (3°), randomly ordered, covering 12° horizontally and 12° vertically. Dependent variables are accuracy, precision and data loss. This was the main trial in Holmqvist [2015].

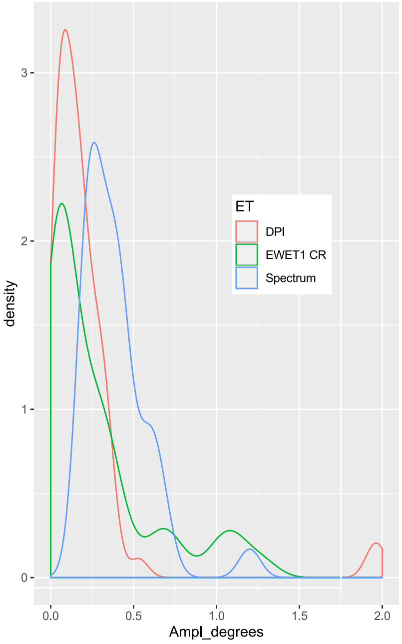

Figure 5: Density plot of microsaccade amplitudes. Total number of microsaccades used in this plot are for the DPI: 64 microsaccades, Tobii: 30, EWET1 CR: 74.

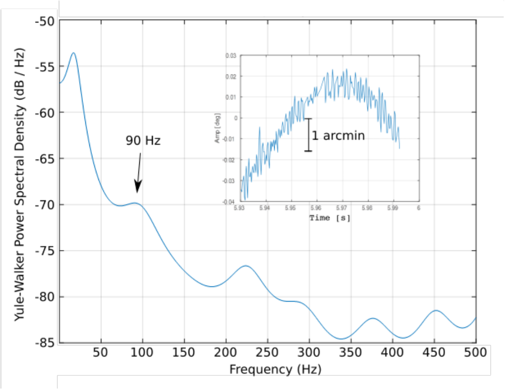

Figure 6: Possible tremor in EWET1 CR data from the long fixation condition detected by a power spectral density analysis. The inlaid plot of the horizontal CR signal shows that the potential tremor has an amplitude of 1 arcmin.

(4) Variations in luminance: the stimulus consisted of one fixation cross (described above) placed on a background that slowly changed from black to white and back with a frequency of 0.125 Hz following a sinusoidal temporal profile. Maximum luminance was 62.0 cd/m^2 measured with a spot photometer (MINOLTA CS 100), and minimum at 0.07 cd/m^2. Presentation time was 80 s (10 cycles of 8 s). The participants were asked to continuously fixate the marker without blinking during stimulus presentation. We investigated whether horizontal and vertical gaze oscillates with the variation in luminance, following several studies by Drewes et al. [2012] and others.

(5) Rhombus: a 5° x 5° rhombus (3.55° each side). The participants were instructed to execute saccades from corner to corner at high speed in a clockwise direction for 20 seconds. We plotted the resulting scanpaths to illustrate differences in saccade dynamics. This task is taken from Viviani et al. [1977].

5.2 Results

Accuracy, precision and data loss. For the eye-movement data recorded when the participants look at the points from the 5 x 5 grid, we first determined a sample selection period of 1.6 seconds, 400ms after the stimulus onset, and then calculated accuracy for those data, precision as root mean square of the sample-to-sample distances (RMS-S2S), and the standard deviation as the square root of the pooled variance in horizontal and vertical (STD). The data are shown in Table 1 as means and medians. For an in-depth discussion of these data quality metrics, see Holmqvist and Blignaut [2020], Holmqvist and Andersson [2017, Chpt 6], Holmqvist [2015], and Niehorster et al. [view].

Table 1: Accuracy and precision measurements in degrees.

	RMS-S2S	STD	ACC
Spectrum Mean	0.25	0.365	0.778
Spectrum Median	0.073	0.234	0.56
DPI Mean	0.059	0.846	2.86
DPI Median	0.012	0.329	1.65
EWET1 CR Mean	0.0017	0.325	NA
EWET1 CR Median	0.0012	0.211	NA
EWET1 GZ Mean	0.0091	0.263	1.72
EWET1 GZ Median	0.0045	0.234	0.91

Figure 3 shows a density histogram of the RMS-S2S values for each eye-tracker. It is clear that the EWET1 CR has a precision around 10 times better than the DPI with these human data, which the means and medians in Table 1 also confirm. Even the combined EWET1 GZ has better precision than the DPI. In comparison, the average RMS-S2S of commercial video-based eye-trackers reported by Holmqvist and Andersson [2017] varied from 0.025 to 0.77 degrees.

The STD precision values in these data come across as high, possibly an effect of a crude sample selection method. Holmqvist and Andersson [2017] report averages of 0.06° for the DPI and 0.08 - 0.63° for commercial video-based eye-trackers.

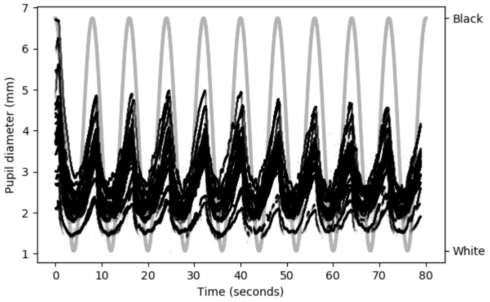

Figure 7: Luminance variation in gray and resulting pupil dilation data in black. Recorded during 80 s of maintained fixation on the Tobii Spectrum.

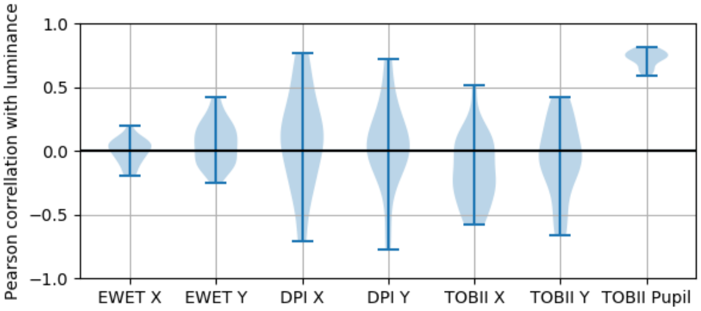

Figure 8: Violin plots of correlations between gaze data and luminance, and in the TOBII Pupil plot between luminance and pupil size (the data in Figure 7). Recorded during 80 s of maintained fixation.

The gaze signals from both the DPI and the EWET CR are strongly coloured, measured as the quotient of RMS-S2S and STD being close to 0 [Holmqvist and Andersson 2017, p. 182], while the colour of the Tobii Spectrum data is white ($\frac{\text{RMS-S2S}}{\text{STD}}$ being much higher). Niehorster et al. [view] show that tested P-CR video eye-trackers all produce white noise at the image processing stage, and that filters smoothen the data and give it colour. The DPI and EWET CR data are coloured too, but for a different reason: their use of analog opto-mechanics as sensors produces a smooth signal from the start.

The accuracy of the EWET1 is significantly lower than those of the DPI and Tobii Spectrum (Figure 4 and Table 1). Other video-based eye-trackers also have better accuracy, ranging from 0.6 to 1.5 degrees [Holmqvist and Andersson 2017].

Data loss, measured as the percentage of missing samples during the grid task, was 0.22% for the DPI, 18.8% for the EWET1, while the Tobii Spectrum lost 3.72% of the samples. In comparison, commercial video based eye-trackers tested by Holmqvist [2015] were reported to have a data loss between 2 and 18%.

Microsaccades and tremor. We used the first 3.5 seconds of the long fixation condition to manually determine the number and amplitudes of microsaccades. Manual detection was preferred as we were comparing three different systems with differing data quality, and algorithms for microsaccade detection have been suspected to build on assumptions regarding the data that yield artefacts [Fang, Gill, Poletti, and Rucci 2018]. For the EWET1, we looked for microsaccades in the CR signal, as it exhibits superior RMS-S2S precision compared to the GZ signal (Figure 3), which in laymens' terms implies that fewer of the small eye movements drown in the noise [Holmqvist and Blignaut 2020].

Results show that, on average, more microsaccades were detected by the EWET1 CR (3.08 ± 1.74) and the DPI (2.88 ± 2.45) than were found in the data from the Spectrum (1.76 ± 3.30). Because we suspected that small microsaccades are not correctly reported by the Spectrum, we made a density plot of microsaccade amplitudes found on each of the three eye-trackers. Figure 5 indeed shows that the EWET1 CR and the DPI reveal more small microsaccade amplitudes compared to the Spectrum. This difference could possibly be explained by the finding by Holmqvist and Blignaut [2020] that small eye-movements in video-based data are often mismeasured, such that eye rotations that are shrunk cannot be detected in the data files, while other movements are artificially extended. Figure 5 is also consistent with observations by Kowler [2011] and others that microsaccades have been reported to be larger with video-based P-CR eye-tracking than with earlier, superior technologies.

The power spectral density (PSD) calculations resulted in clearly visible 90 Hz peaks (suggesting oculomotor tremor) for six DPI recordings and for 1 Tobii Spectrum recording. For the EWET1 CR, we found weak peaks in the PSD plots indicative of tremor for the vast majority of the participants. The horizontal data from one participant, and a PSD plot of them, are exemplified in Figure 6. A similar PSD calculation for the recordings with artificial eyes (Figure 13) resulted in no 90 Hz peak for either of the four eye-trackers.

Effect of variations in luminance. The data collected from Tobii Spectrum and the EWET1 TR in the condition where we sinusoidally varied background luminance were first checked for whether the luminance manipulation had the expected effect on participants' pupil dilation. Pupil dilation plots of the Tobii data are plotted in Figure 7. We then correlated recorded gaze against luminance, which resulted in the violin plots of Figure 8.

For the EWET1 GZ, we found 16.8 % data loss on average, but also very frequent SWJs. There are no clear oscillations in the gaze signal for any of the participants we inspected, neither in horizontal, nor vertical, as the correlations in Figure 8 show. We draw the conclusion that the gaze signal of the EWET1 is unaffected by variations in luminance/pupil dilation.

From the DPI data, the majority of individual data exhibited very clear pendulous gaze signal with the same frequency as the luminance manipulation, which also shows up in the correlation plot. It was expected that the DPI would exhibit such a pupil dilation artefact on gaze, because a constricted pupil diameter can interfere with detection of the 4th Purkinje reflection from the rear side of the human lens. Gaze shifts in DPI data ranged over 1°.

The Tobii eye-tracker was heavily affected by data loss (25.8 % on right eyes, 20.2 % on the left eyes) in this condition and displayed

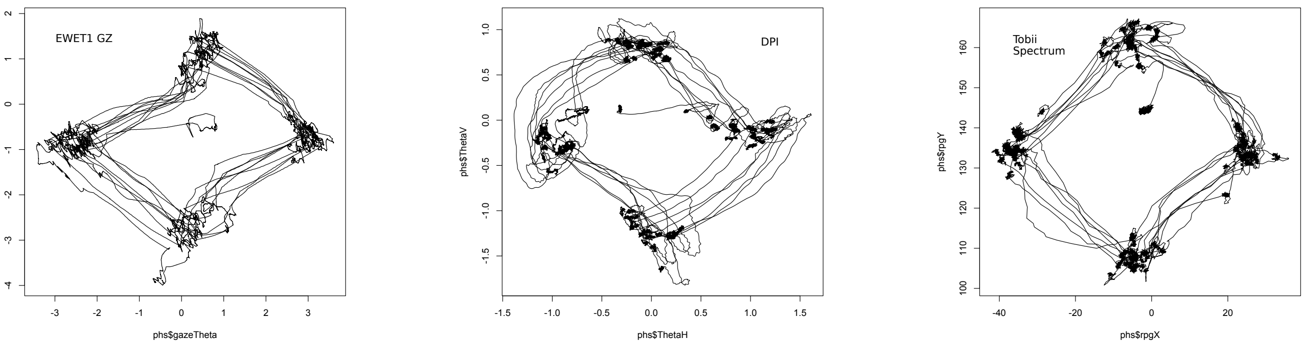

Figure 9: The same participant separately recorded on the three eye-trackers. All rhombs were the same size. Units in the plots are specific for each eye-tracker: degrees, volts and pixels in that order.

an unusually large number of large-amplitude square wave jerks. The Spectrum is a classical video-based P-CR eye-tracker, like the EyeLink 1000, for which the pupil effect on gaze accuracy has previously been demonstrated by Drewes [2014] and others. This artefact is likely to be ubiquitous for all video-based P-CR systems, and make them unsuited for applications where light levels vary. To our knowledge, this is the first time the effect has been shown for the Tobii Spectrum. The gaze shift amplitudes were around 2° on average.

Rhombus data and saccade curvature. In Figure 9, we plotted the full rhombus scanpaths for each eye-tracker, recorded in a typical participant. We can see the characteristic rounded saccades of the DPI in this task, reflecting the damped oscillating movement of the human lens. The Tobii data exhibit somewhat smaller curvatures. In the EWET1 GZ, the post-saccadic oscillations have disappeared.

6 EXPERIMENT 2: CO-RECORDING EWET1 WITH THE DPI

Experiment 2 was conducted in order to examine small differences in concurrently recorded saccade data from the same eye. For Experiment 2, we used the same participants as in Experiment 1, but with the combined eye-tracker shown in Figure 2.

6.1 Stimuli and instruction

Each recording session comprised of 3 trials, presented in a fixed order. Between trials there was a 4-second rest period during which the participants were instructed to blink and rest their eyes, but to keep their head still. The trials were:

1. Calibration: one point at a time was shown for 2 seconds, in a 9-point calibration pattern, covering ±2° range vertically and horizontally.
2. Long fixation: a single fixation point was shown for 10 seconds.
3. Rhombus: a 5° x 5° rhombus (3.55° each side). The participant was instructed to execute saccades from corner to corner at high speed, in clockwise direction, for 20 seconds.

6.2 Results

Rhombus data and saccade curvature. In order to ascertain that the differences in Figure 9 are due to the eye-tracker, we repeated the rhombus task in the combined eye-tracker of Figure 2, where both systems measure the same saccades. Figure 10 shows that EWET1 CR and DPI saccades are very similar, with the somewhat surprising observation that EWET1 CR are at least as curved and have even larger post-saccadic oscillations. However in the EWET GZ signal (Figure 9, and Figure 10), the post-saccadic oscillations no longer appear.

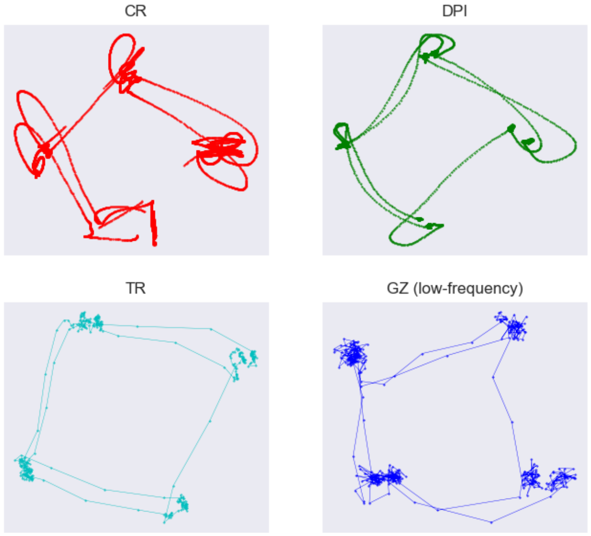

Figure 10: Co-recorded data from simultaneously recorded eye-movements on the EWET1 (left) and the DPI (right), for the rhombus task. All panels are the same size.

Saccade dynamics and translational movements of the eye. In Figure 11, we show a single co-recorded saccade. Note that the dynamic is similar but the post-saccadic oscillations larger in the

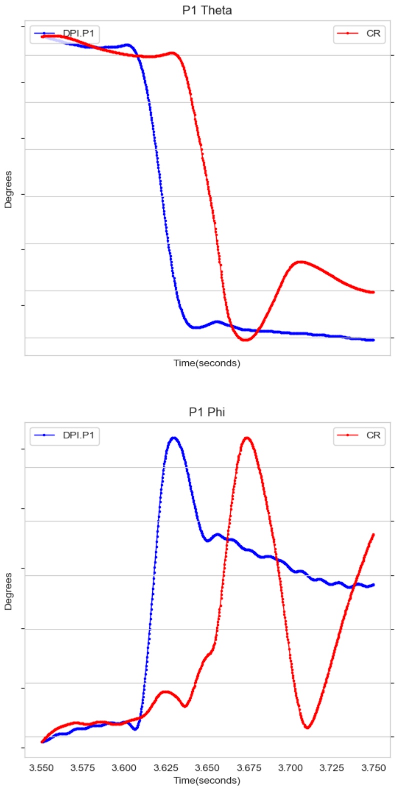

Figure 11: A single 5° saccade co-recorded (Figure 2) simultaneously on the EWET1 and the DPI, horizontal and vertical over time. The two signals are shifted in time for visibility.

EWET1 CR than in the DPI. The wide-spread fixations data of the EWET1 CR are due to small translational movements of the eye, probably breathing and balance muscles, but in the TR and GZ data, they are reduced.

The translational signal represents sideway movements of the eye relative to the eye-tracker, for instance a head-movement, breathing or slippage of a headset. Video-based eye-trackers do not report translation, but this is a unique feature of the DPI and the EWET1, that could be of great importance in future augmented reality headsets. Figure 12 shows the translational signal of the DPI against the TR signal of the EWET1, which also are very similar, except that the DPI does not detect translation in depth (Z).

Figure 12 shows how the lateral movement originates from the saccade itself. The reason is that the measured lateral movement, both in the DPI and the EWET TR, is not of the center of the eye but of a point which is offset by a few millimeters from the center. During fixation, as well as before and after the saccade, we see fluctuations that are actual lateral head/eye movements.

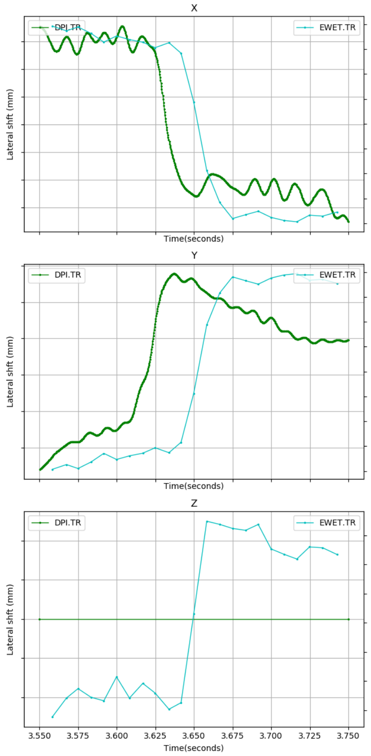

Figure 12: Same saccade as in Figure 2, now comparing the translational component of DPI and the TR signal of the EWET. The DPI does not report translation in depth (Z), which the EWET TR does.

7 EXPERIMENT 3: RESOLUTION AND STEP REPRODUCIBILITY OF THE EWET1 AND DPI, COMPARED TO VIDEO-BASED EYE-TRACKERS

In experiment 3, we used the Stepperbox from Holmqvist and Blignaut [2020] to assess the ideal resolution and step reproducibility of the EWET system, compared to the DPI and the Tobii Spectrum, Here we will also include data from the EyeLink 1000+ taken from Holmqvist and Blignaut [2020].

The Stepperbox is an electromechanical device which rotates two artificial eyes at angles down to 3 seconds of arc. We know from Holmqvist and Blignaut [2020] that all tested video-based P-CR eye-trackers misrepresent the amplitudes of small saccades, up to at least 2°, while the DPI correctly records rotations at all amplitudes.

Figure 13 shows sequences of 2 arcmin steps with 1 second stops in between. Clearly both the EWET1 CR and the DPI can easily resolve eye rotations of this amplitude. The Tobii Spectrum fails at this task, and the EyeLink1000+ only does a marginally better job (data from Holmqvist and Blignaut [2020]).

The resolution of the DPI and the EWET1 are reported in Table 2 as the standard deviation of the recorded step amplitudes. We can see that the variation (the errors) in the measurements of 1-4 arcmin

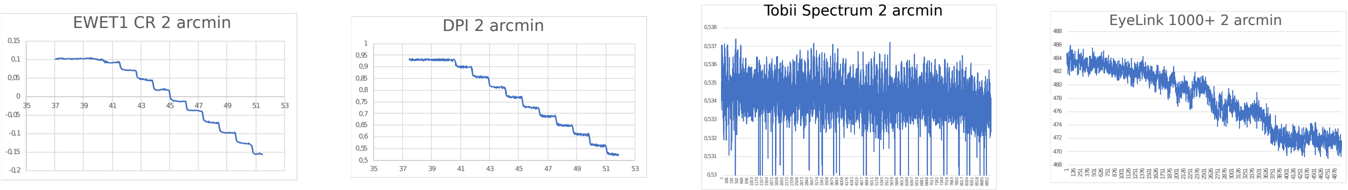

Figure 13: Steps with 2 arcmin amplitude and 1 second stops between steps, recorded with the Stepperbox mechanism from Holmqvist and Blignaut [2020], from which the EyeLink data are taken. All panels are the same size (data for 20 arcmin * 12-13 seconds), and units are specific for each eye-tracker. No filters were applied to the EWET1 and DPI data files, but note that the DPI has several hardware filters that colour its noise. Tobii Spectrum data were unfiltered by default. The EyeLink 1000+ was recorded with default filters on.

Table 2: The STD of measured step amplitudes errors for rotations of the Stepperbox by steps in the range 1-4 arcmin.

	1′	2′	3′	4′
DPI Gaze	0.09	0.17	0.15	0.27
DPI P1	0.13	0.20	0.19	0.14
EWET1 CR	0.35	0.27	0.26	0.32

steps in this prototype version of the EWET1 CR is somewhat larger than for the DPI. The P1 signal of the DPI is added to the table as it is the corresponding feature to the CR signal of the EWET1.

In comparison, the video-based P-CR eye-trackers tested by Holmqvist and Blignaut [2020] had STD resolutions of 9-27 arcmin for 10-100 arcmin steps, while the single video-based eye-tracker tested by Clarke et al. [2002] exhibited 0.17 arcmin STD resolution for 6 arcmin rotations.

8 CONCLUSIONS AND OUTLOOK

We tested a prototype hybrid eye-tracker, consisting of an optoelectronic CR tracker recording at 4000 Hz and a camera-based 120 Hz translational TR movement tracker. It calculates the line of the optical axis (GZ) by combining both signals. A simple calibration routine then determines the visual axis (calibrated GZ).

Our test shows that a hybrid solution to calculate gaze works at least as well as combining signals from the same type of sensors (e.g., P-CR or 1st-4th Purkinje reflection). In particular, the results show that the EWET1 CR tracker exhibits an RMS-S2S precision of one order of magnitude better than the DPI. Even the combined GZ signal shows somewhat better RMS-S2S than the DPI, despite inheriting the poorer precision of the TR. Because the RMS-S2S precision of the DPI is like that of scleral coils [Ko et al. 2016], we indirectly infer that the EWET1 CR has a data quality on par with scleral search coils, dubbed "the golden standard of eye-tracking" by Collewijn [1998].

We have also shown that the EWET1 CR is excellent at detecting microsaccades, that tremor can likely be seen in the data (previously only shown on the DPI and coils), and that the gaze direction data from the EWET1 is unaffected by variation in pupil dilation caused by luminance changes.

Furthermore, the co-recording of the EWET1 with the DPI showed that saccade dynamics of the EWET1 is as good as that of the DPI, and that the GZ signal is free of post saccadic oscillations. The Stepperbox test (Experiment 3) revealed that with artificial eyes, the EWET1 scores as well as the DPI even with extremely small amplitude movements, and that both of these systems easily outperform video-based eye-trackers.

The decoupling of the pupil center from the calculation of gaze makes this hybrid method a good choice for eye-tracking that needs to work in varying light conditions, such as augmented reality systems. The low energy-consumption of the EWET1 compared to established systems could prove to be another benefit for use in AR systems, as it will reduce demands for portable energy-storage.

The EWET1 also promises to become a high-end eye-tracker suitable for the most demanding research, capable of delivering results comparable to those achieved with a DPI or scleral search coils. For event detection, we recommend using the CR signal, while for gaze application, use the GZ.

However, track loss of 18% of the samples is unacceptable (but not unheard of, see Holmqvist and Andersson 2017, p. 165-167). Lower track loss could be achieved by changing details in the physical sensor setup.

Also, accuracy is currently not superior in the EWET1 (GZ median 0.91°). We are improving the quality of the TR data and the TR-CR integration. We are also implementing the user-controlled post-calibration correction method from Poletti and Rucci [2016], which in their application for the DPI yielded an accuracy of 0.08°, and plan to have a binocular EWET during 2020.

9 DISCLOSURE

Author KH is a member of the advisory board of EyeWay Vision Ltd, and consults for the company. Authors AK, MM and MS are employed by EyeWay Vision Ltd. Authors SLÖ and MG have no affiliation to EyeWay Vision Ltd. EyeWay Vision Ltd has submitted patent applications describing the technology of the EWET1 as used in this manuscript.

ACKNOWLEDGMENTS

We are thankful to Florian Hauser at the Ostbayerische Technische Hochschule for lending us the Tobii Spectrum, and to three anonymous reviewers for helpful comments on an earlier version of this manuscript.

REFERENCES

A. H. Clarke, J. Ditterich, K. Drüen, U. Schönfeld, and C. Steineke. 2002. Using high frame rate CMOS sensors for three-dimensional eye tracking. *Behavior Research Methods, Instruments, & Computers* 34, 4 (01 Nov 2002), 549–560. https://doi.org/10.3758/BF03195484

H. Collewijn. 1998. Eye movement recording. In *Vision Research: A Practical Guide to Laboratory Methods*, R. H. S. Carpenter and J. G. Robson (Eds.). Oxford: Oxford University Press, 245–285.

H. D. Crane and C. M. Steele. 1985. Generation-V dual-Purkinje-image eyetracker. *Applied Optics* 24, 4 (Feb 1985), 527–537.

Jan Drewes. 2014. Smaller is better: drift in gaze measurements due to pupil dynamics. *PLoS ONE* (2014).

Jan Drewes, Guillaume S Masson, and Anna Montagnini. 2012. Shifts in reported gaze position due to changes in pupil size: Ground truth and compensation. In *Proceedings of the Symposium on Eye Tracking Research and Applications*. ACM, 209–212.

Yu Fang, Christopher Gill, Martina Poletti, and Michele Rucci. 2018. Monocular microsaccades: Do they really occur? *Journal of vision* 18, 3 (2018), 18–18.

Kenneth Holmqvist. 2015. Common predictors of accuracy, precision and data loss in 12 eye-trackers. (November 2015). Available at ResearchGate.

Kenneth Holmqvist and Richard Andersson. 2017. *Eye tracking: A comprehensive guide to methods, paradigms and measures*. Lund: Lund Eye-Tracking Research Institute.

Kenneth Holmqvist and Pieter Blignaut. 2020. Small eye movements cannot be reliably measured by video-based P-CR eye-trackers. *Behavior Research Methods* (2020).

Ignace Hooge, Kenneth Holmqvist, and Marcus Nyström. 2016. The pupil is faster than the corneal reflection (CR): are video based pupil-CR eye trackers suitable for studying detailed dynamics of eye movements? *Vision research* 128 (2016), 6–18.

Hee-kyoung Ko, D Max Snodderly, and Martina Poletti. 2016. Eye movements between saccades: Measuring ocular drift and tremor. *Vision research* 122 (2016), 93–104.

Eileen Kowler. 2011. Eye movements: The past 25 years. *Vision Research* 51 (2011), 1457–1483. https://doi.org/DOI:10.1016/j.visres.2010.12.014

Michael B McCamy, Jorge Otero-Millan, R John Leigh, Susan A King, Rosalyn M Schneider, Stephen L Macknik, and Susana Martinez-Conde. 2015. Simultaneous Recordings of Human Microsaccades and Drifts with a Contemporary Video Eye Tracker and the Search Coil Technique. *PLOS ONE* 10, 6 (2015), 1–20.

Diederick C Niehorster, Tim HW Cornelissen, Kenneth Holmqvist, Ignace TC Hooge, and Roy S Hessels. 2017. What to expect from your remote eye-tracker when participants are unrestrained. *Behavior Research Methods* (2017), 1–15.

Diederick C Niehorster, Raimondas Zemblys, T. Beelders, and Kenneth Holmqvist. in review. Characterizing gaze position signals and synthesizing noise during fixations in eye-tracking data. *Behavior Research Methods* (in review).

Marcus Nyström, Ignace Hooge, and Kenneth Holmqvist. 2013. Post-saccadic oscillations in eye movement data recorded with pupil-based eye trackers reflect motion of the pupil inside the iris. *Vision Research* 92 (2013), 59–66.

Martina Poletti and Michele Rucci. 2016. A compact field guide to the study of microsaccades: challenges and functions. *Vision research* 118 (2016), 83–97.

Lore Thaler, Alexander C Schütz, Melvyn A Goodale, and Karl R Gegenfurtner. 2013. What is the best fixation target? The effect of target shape on stability of fixational eye movements. *Vision research* 76 (2013), 31–42.

Paolo Viviani, Alain Berthoz, and David Tracey. 1977. The curvature of oblique saccades. *Vision Research* 17, 5 (1977), 661–664. https://doi.org/DOI:10.1016/0042-6989(77)90143-2