Automated static threshold perimetry using a remote eye-tracker

#3908 - D0050

Pete R Jones[1,2], Sarah Kalwarowsky[1], Gary S Rubin[1,2], Marko Nardini[1,3] (*No Commercial Relationships*)

p.r.jones@ucl.ac.uk

[1] UCL Institute of Ophthalmology, London, UK
[2] NIHR Moorfields Biomedical Research Centre, London, UK
[3] Dept of Psychology, Durham University, Durham, UK

1 Summary

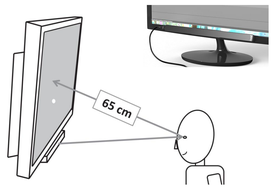
We have demonstrated that automated Static Threshold Perimetry [STP] can be performed using a remote eye-tracker, combined with an ordinary desktop computer. The result is a system that is cheap (hardware: <$1000), user-friendly (no restriction of the head/eyes with chin-rests or fixation-crosses), and could be adapted for use with infants (no button press user-response required). The present version of the test was not as fast or reliable as the existing gold standard (Humphrey Field Analyzer: HFA), and means of improvement are suggested.

2 Background & Methods

Devices such as the Humphrey Field Analyzer [HFA] are the gold standard measure of Static Threshold Perimetry [STP]. Lights of variable intensity presented against a uniform background (10 cd/m^2).

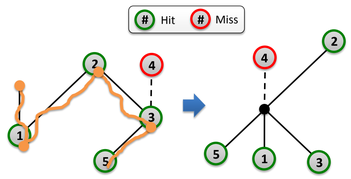


- "Press the button when you see the flash of light"
- Luminance of the light at each of *N* locations adapted to find minimum detection threshold

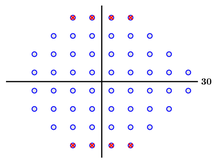


In our test, the (Goldmann III) light-spots were presented on an LCD monitor, and a remote eye-tracker was used to measure the observer's gaze location. We used a Tobii EyeX eye-tracker, which is cheap ($139), and is capable of tracking monocularly.

Instead of having to maintain a constant point of fixation, lights were presented relative to wherever the observer was looking at trial-onset.

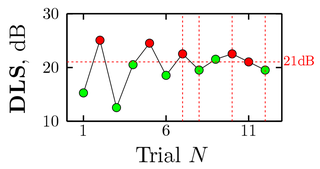


Instead of making an explicit button-press response, the observer was simply asked to look at lights when they saw them



A previous similar test screened for seen/not-seen lights, presented at a single luminance level (Murray et al, 2009). Here, we used an adaptive staircase algorithm was used to find a detection threshold for each location on a 24-2 perimetry grid (similar to a Full Threshold STP algorithm).

Differential Light Sensitivity [DLS] thresholds were found using a 1-up 1-down staircase, with decreasing step sizes. Presentation levels were also informed by neighbouring points. Each staircase terminated after 8 reversals. DLS was the median of the last 4 reversals.



3 Technical challenges

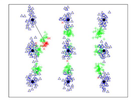
screen calibration

To ensure uniform luminance across the screen, luminance (cd/m^2) was measured for 17 command levels (0…255), at each of 30 screen locations. Tensor product linear interpolation was used to derive spline-fit gamma functions for every screen location (pixel).



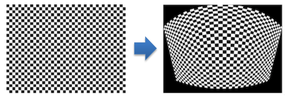
eye-tracking calibration

Observed-to-actual gaze coordinates were mapped using the manufacturer's 9-point calibration routine. Additional calibration was then performed in software. Firstly, via our own 9-point polynomial surface map (an iterative routine, excluding outliers). And then by a trial-by-trial drift-correction mechanism.



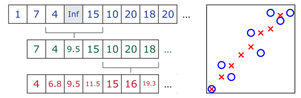
correction for viewing angle

The size and shape of objects on a flat screen vary with retinal eccentricity. We corrected for this in software, by performing an inverse distortion, accounting for angles of azimuth and elevation. This ensured stimuli were a constant size and shape on the retina, at all screen locations.



data cleaning

Incoming eye-tracking data were cleaned in three steps. (1) Outlying values were excluded. (2) Small gaps in data (e.g., due to flickery contact) were filled using linear interpolation (3) Data were smoothed (low-pass filtered) by computing the moving-average.



high-fidelity luminance

Bit-stealing was used to increase the number of testable luminance values. Bit-stealing involves using colour pixels to increase grey-scale resolution. It increased the effective bit-rate from 8 to 10.7, without the need for an expensive graphics card.

response classification

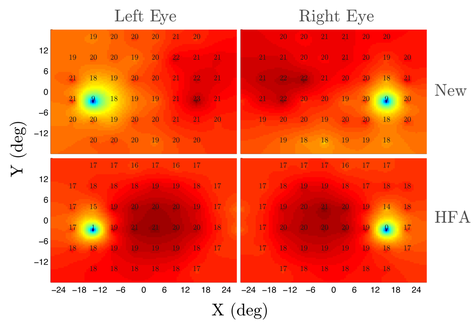
Responses were determined based on a simple algorithm that required *N* gaze points to fall within *X* degrees of the target. These criteria were made more stringent for near points than for far points, in-line with expected error rates (though see § 7)

4 Results: *Accuracy*

Here we show results for 34 naïve, normally sighted adults (18-30 years), each of whom was tested monocularly (17 left, 17 right). The untested eye was patched. To assess accuracy, observers performed both our test ("New") and the HFA. To assess test-retest reliability, each test was performed twice. To simplify comparison, results from both tests were converted to the same decibel scale:

$$dB = 10\log_{10}\left(\frac{155}{\mathrm{DLS}_{cd/m^2}}\right)$$

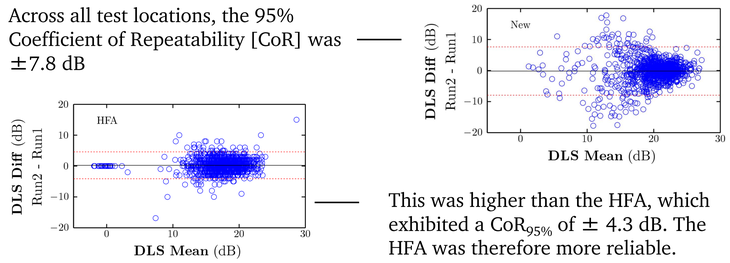
Our new test was able to produce meaningful STP results, which were consistent with those of the HFA. There was a significant correlation thresholds at individual retinal locations [$r_{2436} = 0.27$, $p \ll 0.001$], and between mean sensitivity scores for each observer [$r_{27} = 0.44$, $p = 0.017$].



Note that unlike with the HFA, each point, including the blind spot, was tested in the same way, using the same starting point. The fact that the algorithm could differentiate between the blind-spot and surrounding locations suggests that it should be capable of isolating even relatively small scotomas

5 Results: *Test-Retest Reliability*

Across all test locations, the 95% Coefficient of Repeatability [CoR] was ±7.8 dB

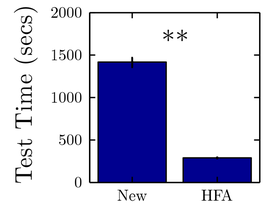


This was higher than the HFA, which exhibited a $\mathrm{CoR}_{95\%}$ of ± 4.3 dB. The HFA was therefore more reliable.

6 Results: *Testability*

5 of 34 (15%) individuals could not be eye-tracked reliably, and so were unable to complete the new test. All individuals completed the HFA.

Test durations were substantially longer than with the HFA (~23mins vs. ~5mins). However, test-times were comparable with previous Full Threshold STP algorithms.

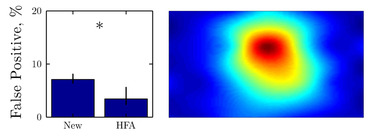


7 Results: *Errors*

Blank and highly-suprathreshold targets were used to asses False Positive and False Negative rates, respectively.

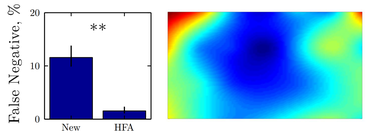
False Positives

The chance of responding correctly by chance was 7%. These errors tended to happen in the centre of the screen, and likely represent observers tending to gravitate to the centre of a blank screen.



False Negatives

The chance of being scored falsely as not seeing the target was 12%. These errors tended to happen in the top corners of the screen, and likely represent limitations of the eye-tracker.



8 Conclusions & Next Steps

We have demonstrated that remote eye-tracking can be used to perform automated STP, using a standard desktop computer and inexpensive hardware.

Our next step will be to improve the efficiency of the test algorithm, by incorporating features such as prior knowledge and a dynamic stopping criterion (McKendrick & Turpin, 2005). Based on pilot data, this will improve test durations and reliability to levels similar to that of the HFA. We will then adapt the test for use in children and infants.

[1] Murray et al. (2009). Ophthalmology, 10, 2017-2026
[2] McKendrick & Turpin (2005). Optom. & Vis. Sci., 82, 981-987

This work was supported by Fight For Sight [#1353], the NIHR Biomedical Research Centre for Ophthalmology (at Moorfields Eye Hospital and the UCL Institute of Ophthalmology), and the Special Trustees of Moorfields Eye Hospital

For reprints or further information please email:
p.r.jones@ucl.ac.uk