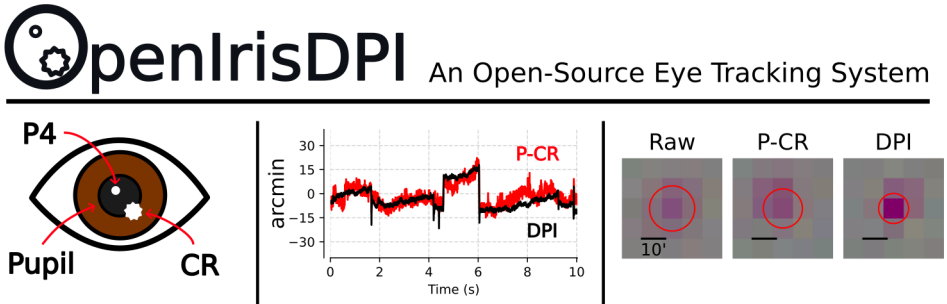
Graphical Abstract

OpenIrisDPI: An Open-Source Digital Dual Purkinje Image Eye Tracker for Visual Neuroscience

Ryan A. Ressmeyer, Jorge Otero-Millan, Gregory D. Horwitz, Jacob L. Yates



- OpenIrisDPI enables both pupil-corneal reflection (P-CR) and dual Purkinje image (DPI) tracking.
- OpenIrisDPI can track both eyes at 500 Hz on consumer hardware.

- P-CR signals have more measurement noise than DPI signals.
- P-CR signals have more fixational drift than DPI signals.

- P-CR and DPI signals improve estimates of LGN receptive fields (RF) during fixation.
- DPI signals improve RF signal-to-noise ratio by more than double compared to P-CR signals.

Highlights

OpenIrisDPI: An Open-Source Digital Dual Purkinje Image Eye Tracker for Visual Neuroscience

Ryan A. Ressmeyer, Jorge Otero-Millan, Gregory D. Horwitz, Jacob L. Yates

- OpenIrisDPI is a new open-source eye tracking system.
- OpenIrisDPI tracks the pupil, corneal reflection, & fourth Purkinje image at 500 Hz.
- Dual Purkinje image-based eye tracking is more precise than pupil-based tracking.
- DPI improves receptive field characterization of LGN neurons in fixating macaques.

OpenIrisDPI: An Open-Source Digital Dual Purkinje Image Eye Tracker for Visual Neuroscience

Ryan A. Ressmeyer[a,*], Jorge Otero-Millan[b], Gregory D. Horwitz[c], Jacob L. Yates[b]

[a] *Bioengineering, University of Washington, 3720 15th Ave NE, Seattle, 98195, WA, USA*
[b] *Herbert Wertheim School of Optometry and Vision Science, University of California, Berkeley, 380 Minor Hall, University of California Berkeley, 94720, CA, USA*
[c] *Neurobiology & Biophysics, University of Washington, 1705 NE Pacific Street, Seattle, 98195, WA, USA*

Abstract

Video-based eye trackers are widely used in vision science, psychology, clinical assessment, and studies of human-computer interaction. Most of these eye trackers use the pupil center and corneal reflection (P-CR) to estimate the direction of gaze. However, P-CR eye trackers are often too imprecise to be used in applications with stringent requirements on eye tracking quality. To address this limitation, we present OpenIrisDPI, an open-source plugin for the OpenIris framework that implements dual Purkinje image (DPI) tracking. The system simultaneously tracks both P-CR and DPI signals, enabling a direct comparison between them. We leveraged this fact to demonstrate that the P-CR method overestimates the amount of eye movement during fixation compared to DPI. We further validated the accuracy of our DPI signals by simultaneously recording from neurons in the macaque lateral geniculate nucleus. These data revealed that reconstructing visual inputs using DPI produced tighter correlations between stimuli and neural responses than using P-CR. Through open-source implementation and detailed documentation, OpenIrisDPI makes high-precision eye tracking readily available to the research community.

Keywords: Digital dual Purkinje image eye tracking, Fixational drift, Pupil

*Corresponding Author
Email address: ryanress@uw.edu (Ryan A. Ressmeyer)

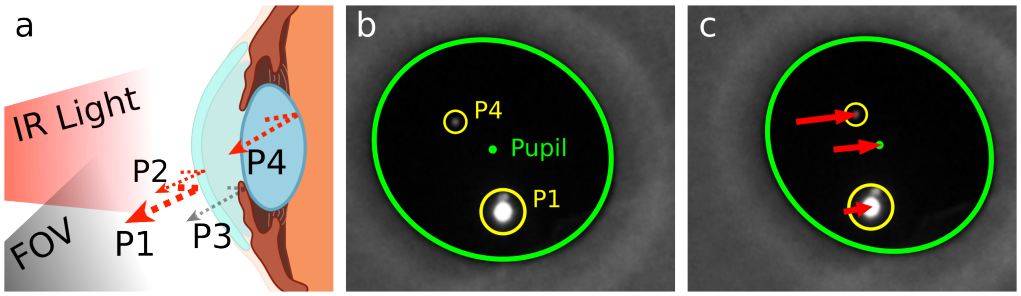


Figure 1: Differential motion of features during eye rotation enables DPI and P-CR tracking. a) Anatomical cross-section showing the four optical surfaces that generate Purkinje reflections: outer cornea (CR or P1), inner cornea (P2), outer lens (P3), and inner lens (P4). The eye is illuminated with an infrared (IR) light, and reflections are within a camera's field of view (FOV). DPI and P-CR eye trackers are distinguished by the usage of either P4 or the pupil center. Both methods track the CR. b) Eye image from the oblique-angle camera showing CR and P4 (marked in yellow), with the pupil contour fitted by a green ellipse. The center of the pupil ellipse is depicted by a green dot and labeled. c) Image captured after a leftward saccade, with red arrows indicating the displacement of CR, P4, and pupil center from their positions in (b). Note that the CR, P4 and the pupil center move by different amounts during eye rotation but would move together during head translation, allowing eye trackers to distinguish between these movements.

decentration, Receptive fields, Lateral geniculate nucleus, Macaca mulatta

1. Introduction

Video-based eye tracking is widely used for basic and applied vision research, but many systems have limitations that make them ill-suited for applications with stringent requirements for precision (see Appendix A for a discussion on precision in the context of eye tracking) [20, 21]. In particular, most video eye trackers estimate the direction of gaze by comparing the locations of the center of the pupil and the corneal reflection (P-CR). Importantly, P-CR systems suffer from two distinct sources of imprecision that affect their ability to reliably measure small eye movements.

The first source is frame-to-frame errors in estimating the center of the pupil and corneal reflection (CR). This noise is temporally uncorrelated in the absence of filtering and can be of sufficient amplitude to obscure the details of small eye movements [38]. The second source of imprecision, known as pupil decentration, is the result of asymmetric dilation and contraction

of the iris. Such asymmetric movements cause the center of the pupil to translate on the order of 0.1 mm between extreme dilation states [58]. P-CR systems misinterpret these small shifts as substantial eye movements of a degree or more [56, 59]. While frame-to-frame measurement noise can be reduced by optimizing a system's optical configuration and feature tracking algorithm, pupillary artifacts are an inherent drawback of P-CR eye trackers. Some studies have addressed pupil-dilation-related imprecision by performing separate calibrations for different luminance conditions [13, 14], but this compensation is both imperfect and rarely implemented in practice.

Dual Purkinje image (DPI) eye tracking offers a more precise alternative to P-CR for measuring miniature eye movements and reliably determining the direction of gaze [28]. DPI and P-CR eye trackers both use the differential movement of the CR and a secondary feature to estimate the direction of gaze (Fig. 1). However, instead of using the pupil center, DPI eye trackers use a reflection from the posterior surface of the lens (fourth Purkinje image or P4). Because they do not rely on the pupil center to estimate the direction of gaze, DPI eye trackers are not affected by pupil decentration.

Considering that DPI eye trackers are both noninvasive and highly precise, one might expect them to be commonplace. Surprisingly, this is not the case. DPI eye trackers are used in only a handful of specialized labs, in part because traditional DPI systems rely on custom optoelectronic feedback loops, which make them notoriously difficult to maintain and operate [8, 57]. Several video-based methods for DPI tracking have recently been developed [1, 32, 53, 57], which may overcome the limitations of these traditional systems.

We present OpenIrisDPI with the goal of making DPI eye tracking widely available to the research community (Fig. 2). OpenIrisDPI is a plugin for the OpenIris eye tracking framework [47] that delivers several novel features: (1) simultaneous pupil and DPI tracking, (2) 500 Hz binocular sampling using modern consumer CPUs, and (3) an optional oblique imaging configuration that eliminates the need for hot mirrors. The source code for the plugin is free and openly available under the GPLv3 license (`https://github.com/ryan-ressmeyer/OpenIrisDPI`). The associated wiki provides an overview of DPI principles and analysis techniques, as well as instructions for building imaging systems, setting critical parameters, integrating OpenIris with experimental systems, and synchronizing recorded timestamps with independent data streams (`https://github.com/ryan-ressmeyer/OpenIrisDPI/wiki`). At the time of publication, a custom digital DPI system with sub-arcminute

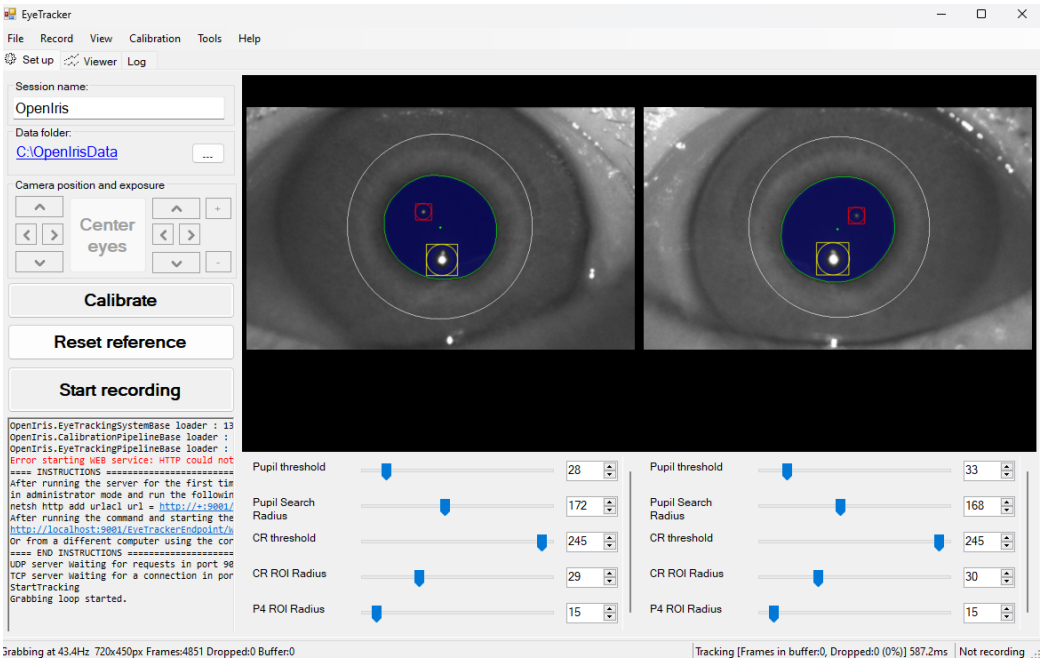


Figure 2: The OpenIris graphical user interface. The upper-right panel shows binocular camera images captured from an oblique angle. The tracking algorithm identifies key features in each eye: the corneal reflection (CR, yellow), the posterior lens reflection (P4, red), and the pupil contour (fitted ellipse, green). The white circle denotes the region where pixels below a threshold are assumed to belong to the pupil, and blue shading indicates pixels below that threshold. The bottom panel provides interactive control of tracking parameters including thresholds, search regions, and feature detection settings. The left panel is used to configure data acquisition parameters—including camera and recording settings—and to show the system log.

noise (see Results) can be assembled for less than $5000.

1.1. DPI Eye Tracking for Visual Neuroscience

Highly precise eye tracking systems are essential for visual neuroscience, where accurate knowledge of the retinal image is critical to characterize neural responses [50]. Neurons in the visual system respond to specific patterns of light falling on the retina [22, 23, 30]. The region of the retina that contributes to a given neuron's response (called the receptive field, or RF) is fixed, but as the eye moves, this region samples different portions of the visual world [16, 18, 60]. The RFs of some neurons subtend just fractions of a degree, which is substantially smaller than both the amplitude of many eye

movements and the noise in some eye trackers.

Experimentalists, therefore, often attempt to minimize eye movements during studies of visual processing, typically by instructing human participants or by training research animals to fixate during stimulus presentations. However, residual eye movements during fixation are inevitable and modulate neural activity throughout the visual hierarchy [3, 29, 51]. Compensating for the effects of these movements on the retinal image, either online [18, 52], or in post-hoc analysis [17, 31, 36, 44, 54], can improve measurements of neuronal responsiveness, stochasticity, and selectivity.

We sought to demonstrate that OpenIrisDPI is suitable for such high-precision applications. We therefore recorded from neurons in the lateral geniculate nucleus (LGN) of fixating macaque monkeys while tracking their eye movements using OpenIrisDPI. We show that DPI-based eye tracking improved the signal-to-noise ratio and spatial resolution of receptive field estimates and outperformed P-CR-based eye tracking. These results highlight the practical impact of high-precision eye tracking for visual neuroscience and demonstrate that compensating for variability introduced by eye movements can lead to more accurate characterization of neural responses in the early visual system, even during steady fixation.

2. Materials and Methods

2.1. OpenIrisDPI Algorithm and Hardware

The OpenIrisDPI tracking algorithm processes video input sequentially: first fitting the pupil boundary, then identifying the CR, and finally localizing P4. Pseudocode for the algorithm is given in the appendix (Appendix B).

The tracking algorithm is implemented as a plugin for OpenIris and utilizes the OpenCV computer vision library through its C# wrapper, Emgu CV. OpenCV is highly optimized for CPU-based parallelism, and enables OpenIrisDPI to process video at high rates. All data presented in this publication were recorded in real-time using a personal computer at a sampling rate of 500 Hz from two 720x450p video streams. Although data were acquired binocularly, data from the left eye only were analyzed. The computer running OpenIrisDPI consisted of consumer-grade components, namely a 12th generation Intel Core i9-12900K CPU, 16 GB of DDR4 RAM, and a 512 GB SSD.

The optical hardware configuration consists of two infrared-sensitive digital cameras (FLIR BFS-U3-16S2M-CS) equipped with 100mm macro lenses

(Laowa). These cameras were mounted below the subject's line of sight (30°) and each was positioned to image a single eye. Infrared illumination was provided by a focused LED source (Thorlabs M940L3) positioned above the cameras (at an oblique angle of 25°) to optimize the visibility of both the CR and P4. Using this configuration, P4 was visible for eye positions up to 10° from the center of the screen in any direction before becoming obscured by the iris.

2.2. Model Eye Experiments

An artificial eye (Ocular Instruments OEMI-7) was used to benchmark measurement error. The model eye was fixed to a stepper motor (STEPPERONLINE Nema 17) connected to a driver (STEPPERONLINE DM542T) that moved in 0.54-arcminute steps. The accuracy of these steps was validated by attaching a laser pointer to the motor and measuring the displacement of the beam over a distance of 43 feet. The artificial eye and stepper motor were then placed on a platform at the approximate location of the animal subject's eyes during experiments. To stimulate standard use, no effort was made to reduce vibrations of the eye or cameras. A third reflection, not present in real eyes, was visible in the model eye and ignored by the tracking algorithm. During experiments, the pupil, CR, and P4 of the model eye were tracked while it was rotated through a series of five equally spaced steps from 2.16' to 10.8', followed by 10 minutes of rest.

2.3. Neurophysiological Experiments

All protocols conformed to the guidelines provided by the US National Institutes of Health and the University of Washington Animal Care and Use Committee. Data were collected from two adult rhesus macaques (Macaca mulatta), one male (Monkey O) and one female (Monkey C). Each monkey was surgically implanted with a titanium headpost and a recording chamber (Crist Instruments) stereotaxically positioned over the LGN. During experiments, the monkeys sat in a primate chair (Crist Instruments) 60 cm from a rear projected digital display (VPixx PROPixx, 51 cm width, 1280x720p resolution) in a dark room while their eyes were tracked using OpenIrisDPI. Monkeys were trained to sit and fixate projected black dots ($0.2° \times 0.2°$) for juice rewards. Data from 11 recording sessions (7 monkey O, 4 Monkey C) were analyzed for this study.

The experiments were conducted in two phases: calibration and fixation. During calibration, saccade targets ($0.2 \times 0.2°$) were presented in random

order on a uniformly spaced 7×7 grid spanning $12 \times 9°$ for Monkey O or in a 5×5 grid spanning $4 \times 4°$ for Monkey C. A target was considered fixated if the monkey's eye position—as reported by a manually-calibrated analog voltage signal between the OpenIrisDPI computer and the experimental control computer—was within $1.4 \times 1.4°$ of its displayed position. After continuous fixation for 600ms, the target was extinguished, and the monkey received a juice reward paired with a tone.

After calibration, the fixation task began. During the task, a black dot was presented continuously at a set location, typically the center of the screen, to ensure that the recorded LGN neurons could be stimulated by the display. The monkey was rewarded at semi-random intervals (range 400-800ms) for maintaining fixation on this dot. While the monkey fixated, a 120×120-element stimulus grid subtending $17.3 \times 17.3°$ was displayed over the receptive fields of recorded neurons. The RGB values of each element were updated at 60 Hz using independent draws from a Gaussian distribution with mean equal to the background. The task continued until 12-18 minutes of stimulus frames had been presented.

2.4. Eye tracker calibration

Raw DPI and P-CR signals for the left eye were computed by taking the difference between the estimated centers of the CR and P4 (for DPI), or the CR and pupil (for P-CR). Data from the calibration task were processed as follows. First, the timestamps of the eye position signals were aligned to the times of task events (e.g. stimulus appearance) using coincident pulses registered by a shared digital synchronization line. This procedure is detailed in the OpenIrisDPI wiki (`https://github.com/ryan-ressmeyer/OpenIrisDPI/wiki/Synchronizing-OpenIris-Data`). Then, the median DPI and P-CR pixel positions 100 ms before the target was extinguished were calculated. Outliers—defined as fixations where the mean distance between a given point and all other points for a given fixation position exceeded $3\times$ the interquartile range above the median—were discarded. The remaining points were then regressed onto the positions of the displayed target positions using a quadratic model [27]. The quality of the fits were visually inspected for accuracy. Where appropriate, calibrated signals are reported in units of visual angle using a small angle approximation.

2.5. High Density Electrophysiological Recordings

We recorded extracellular action potentials from LGN neurons using high-density silicon probes (Neuropixels NHP 1.0 long). The probes were held using a custom adapter that maintained alignment between the probe shank, a guide tube (25mm long 20G hypodermic needle), and microdrive assembly (Narishige). Before experiments began, the guide tube was lowered through the dura mater using a thumb screw. The probe was then advanced using a hydraulic microdrive until it penetrated the LGN. Neural activity was verified as originating from the LGN using a hand-held flickering red light. The action potential (AP) band (30 kHz) and local field potential (LFP) band (2.5 kHz) were recorded using OpenEphys [49]. Action potentials were sorted offline using Kilosort2.5 [42].

2.6. Inter-saccadic drift analyses

Eye position analyses were restricted to epochs when the eye position was within 1.5° of the fixation point for at least 500 ms. Saccades during these epochs were identified using a standard algorithm [39], and intersaccadic epochs were identified as periods between saccades that were $\geq$ 50 ms. Intersaccadic epochs were ignored if the maximum displacement exceeded 1° in a 200 ms window.

The power spectral density (PSD) of horizontal and vertical eye position during fixations was estimated separately for each dataset using Bartlett's method as implemented by the Python library, Scipy. The square root of the PSD was computed to yield the amplitude spectrum. Diffusion constants were estimated by computing the mean squared displacement (MSD) of eye position for time lags of 2 to 200 ms during fixations, then taking the slope of the best-fit line and dividing it by 4 [7, 19, 24].

2.7. Receptive Field Analysis

We analyzed receptive fields using spike-triggered averaging. Eight stimulus frames preceding each spike were reconstructed offline and averaged across spikes [6]. We defined the signal to noise ratio (SNR) of each STA as the peak pixel intensity (RMS across color channels) divided by the median pixel intensity in the first frame (0-13 ms lag). Macaque LGN neurons have minimum response latencies of 16 ms [34], making the first frame of the STA a suitable noise reference. Clusters—the units found by Kilosort—with SNR below 5 were not considered further.

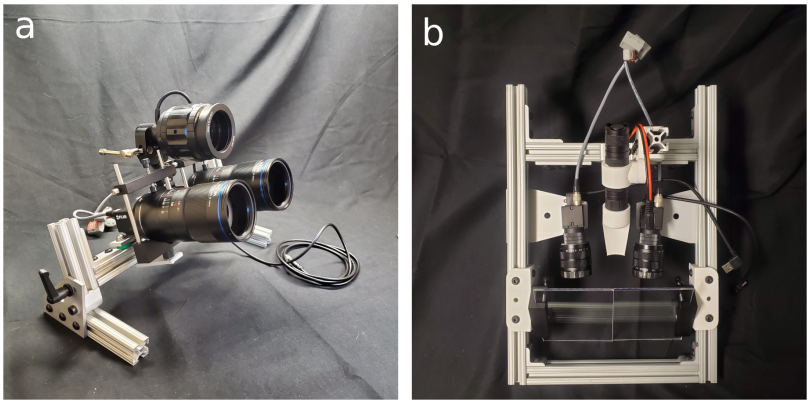


Figure 3: Hardware configurations for DPI eye tracking. a) An oblique imaging system positioned below the subject's line of sight that does not require a hot mirror. b) A conventional, paraxial imaging system that uses hot mirrors to align the optical axes of the cameras and the eyes. This paraxial system is designed to mount to a primate chair (Crist Instruments).

To compensate for fixational eye movements, filtered DPI and P-CR signals (zero-phase 8th order Chebyshev filter, 30 Hz cutoff) were interpolated at the temporal midpoint of each frame. The resulting estimates of gaze were then used to resample the stimulus by shifting each stimulus element to the nearest grid location given the estimated eye position.

Receptive field size was quantified with a 2D Gaussian fit (Scipy curve_fit) to the frame with the maximum pixel intensity. We excluded STAs if the fit failed to converge or yielded implausible radii ($> 20'$). The final radius was calculated as the geometric mean of the standard deviations along the major and minor axes.

3. Results

3.1. Innovations of OpenIrisDPI

Significant contributions of OpenIrisDPI are (1) a CPU-optimized algorithm for simultaneously tracking the pupil, CR, and P4 in real time, (2) documentation of hardware configurations for DPI eye tracking (Fig 3) — including an oblique configuration that is novel for DPI tracking — and (3) integration with the OpenIris eye tracking framework for ease of use and extensibility.

Most previously described DPI eye trackers use a hot mirror to align the optical axes of the eye and camera [1, 8, 57]. This approach minimizes signal non-linearity while still allowing the participant to view experimentally controlled stimuli. However, a hot mirror also adds complexity by filtering the light between the display and the subject. We found that robust DPI eye tracking is achievable from an oblique angle provided that the infrared illuminator is placed at a shallower angle than the imaging cameras. This configuration centers P4 in the pupil for neutral gaze, which maximizes the range of eye positions over which P4 is visible. Placing the imaging system at an oblique angle eliminates the need for a hot mirror and allows a DPI eye tracker to be a drop-in replacement for many commercial systems.

A second innovation of OpenIrisDPI is its CPU-optimized real-time tracking algorithm. Digital DPI eye tracking systems described previously either process video data post-hoc [53], or rely on GPUs to track the Purkinje reflections in real-time [57]. By optimizing the OpenIrisDPI algorithm for the CPU, we aimed to make DPI eye tracking portable and cost-effective, with minimal performance trade-offs. To achieve this goal, we compared the radially symmetric center algorithm [43] used by Wu et al. (2022) [57] to a threshold-rectified centroid algorithm [2] and found the latter to be faster and more accurate for reasonable thresholds in simulations (Fig. B.1). In addition, we simplified the algorithm for locating P4, which previously involved a template matching operation [57]. Excluding the CR, P4 is typically the brightest object in the pupil. We therefore locate P4 by identifying the maximum amplitude pixel within the pupil border, excluding a small region surrounding the CR. The result of these optimizations is a lightweight algorithm that is capable of binocularly tracking the Purkinje reflections and the pupil simultaneously at 500 Hz using consumer-grade hardware. This supports both high-resolution gaze tracking and simultaneous pupillography.

3.2. DPI signals exhibit sub-arcminute precision with a model eye

To assess the baseline performance of the OpenIrisDPI system in our experimental configuration, we used a model eye affixed to a stepper motor (Fig. 4a). The eye was driven through a series of 2.16–10.8' steps for calibration (Fig. 4b), then held steady for 10 minutes to assess signal stability. During these 10 minutes, the DPI signal was less variable than the P-CR signal (fig 4c, P-CR RMS: 1.83', DPI RMS: 0.27'). The amplitude spectrum was flat above 50 Hz of both signals and increased at lower frequencies. The median amplitude of the P-CR signal (1.83×10^{-3}deg / $\sqrt{\text{Hz}}$) was similar to

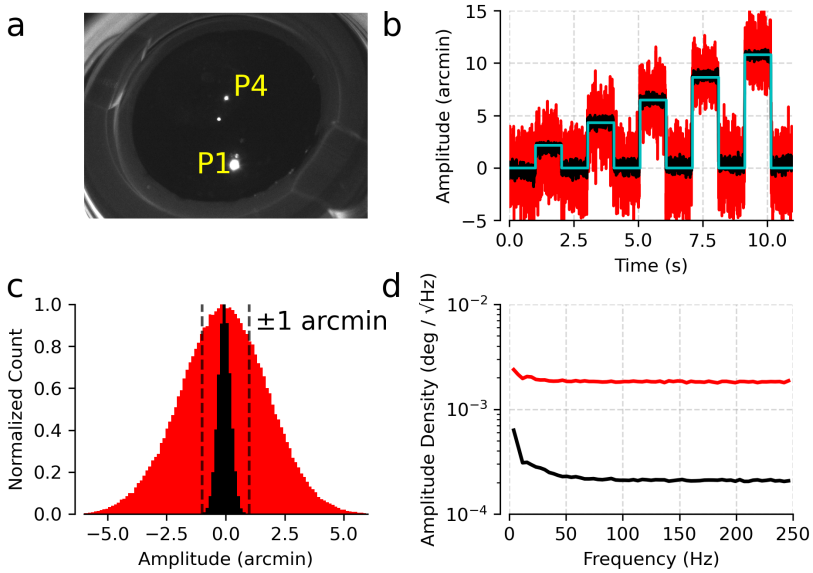


Figure 4: Tracking an artificial eye. a) Single frame from a video of the model eye. CR and P4 are labeled in yellow. b) P-CR (red) and DPI (black) signals in response to a series of steps. The cyan line indicates the requested rotation. c) Distribution of reported orientation while the model eye was stationary. d) Amplitude spectra of P-CR and DPI eye tracker signals.

that obtained by Niehorster et al. (2021) using the SR EyeLink 1000 Plus [38], and the median amplitude of the DPI signal (2.11×10^{-4}deg / $\sqrt{\text{Hz}}$) was substantially lower than any eye tracker reported in that study.

Both P-CR and DPI signals had higher signal amplitude at low than high frequencies (3.91 Hz amplitude minus median amplitude; DPI: 2.90×10^{-4} deg / $\sqrt{\text{Hz}}$; P-CR: 2.35×10^{-4} deg / $\sqrt{\text{Hz}}$). This increase in low frequency signal power was presumably caused by residual mechanical oscillations in the experimental setup.

3.3. Comparing DPI and P-CR signals in vivo

Model eyes, while commonly used for eye tracker validation, fail to capture many aspects of real eyes including pupillary decentration [58], iris kinematics [40], lens wobble [53], and ocular tremor [5]. Consequently, measures of precision obtained with model eyes are likely overestimates. We therefore compared the DPI and P-CR measurements extracted from video data collected from two rhesus macaques (see Methods). Visual inspection of the

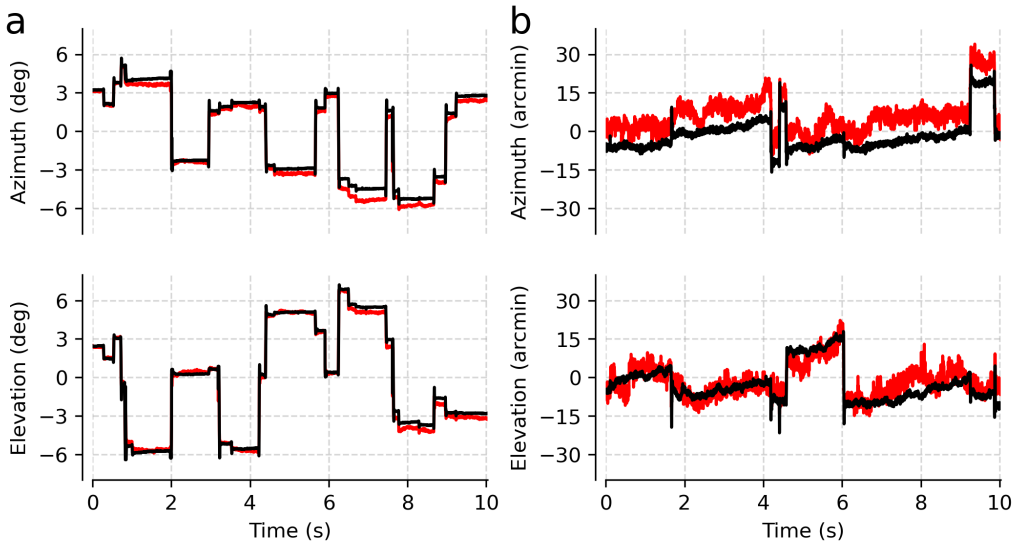


Figure 5: Example of simultaneously recorded DPI (black) and P-CR (red) signals. a) Azimuthal (top) and elevational (bottom) eye movements during calibration. b) Azimuthal (top) and elevational (bottom) eye movements during sustained fixation. Note difference in ordinate scale between (a) and (b).

resultant eye traces (Fig. 5) showed that: (1) the frame-to-frame noise was higher for the P-CR signal than for the DPI signal, and (2) the P-CR signal displayed a slow drift that the DPI signal lacked. The following analyses isolate and quantify these two effects.

3.3.1. DPI signals exhibit less frame-to-frame noise than P-CR signals

To quantify frame-to-frame measurement noise of the DPI and P-CR methods, we calculated broadband noise amplitude between saccades. We assume that eye tracking signals, $x[n]$, are the sum of a noiseless signal, $s[n]$, and a temporally uncorrelated, zero-mean noise source, $\epsilon[n]$. We are interested in estimating the standard deviation of $\epsilon[n]$: $\sigma_\epsilon = \sqrt{\mathbb{E}\{\epsilon^2\}}$. If we had access to $\epsilon[n]$ in isolation, we could use the RMS as an estimate. However, since we only have access to the combined signal, $x[n]$, another method is needed. If $s[n]$ is band-limited below a frequency cut-off, f_c, then an estimate of σ_ϵ may be obtained from the average spectral power above the frequency cut-off [48]. We obtain this estimate using the formula:

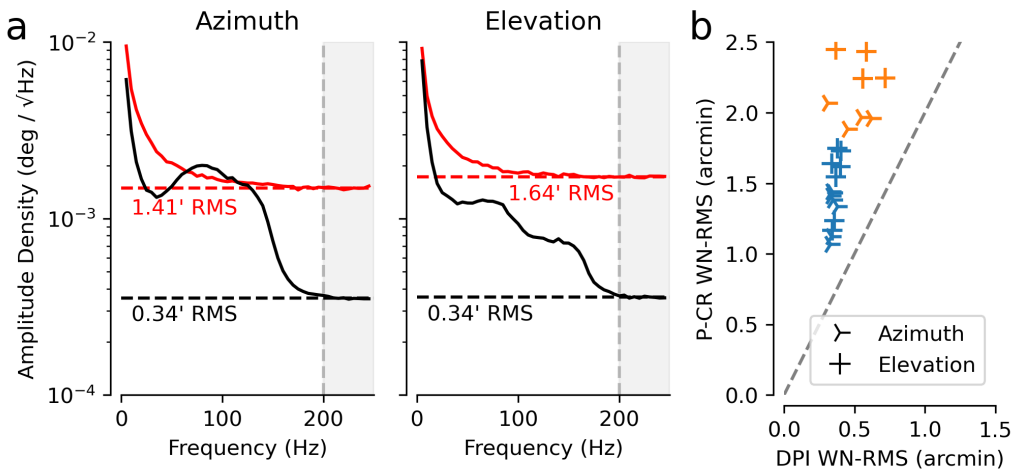


Figure 6: White noise amplitude during fixational intersaccadic drift. a) Amplitude spectrum of DPI (black) and P-CR (red) signals for azimuthal (left) and elevational (right) components of an example session. Dashed lines indicate the mean spectral amplitude density above 200 Hz (shaded gray). b) WN-RMS value computed for DPI (abscissa) and P-CR (ordinate) signals. Azimuthal and elevational signal components are marked with three pointed symbols and crosses, respectively. Blue and orange markers indicate data were collected from Monkey O and C, respectively. The gray dashed line indicates $y = 2x$.

$$\hat{\sigma}_\epsilon = \text{WN-RMS} = \sqrt{\frac{f_s}{2(N - N_0 + 1)} \sum_{k=N_0}^{N} X_s[k]}$$

where $X_s[k]$ is a normalized, one-sided estimate of the PSD of $x[n]$, N is the number of bins in the PSD, N_0 is the first bin above f_c, and f_s is the sampling frequency of the signal. We refer to this measure as the white noise RMS (WN-RMS) amplitude.

We calculated the WN-RMS amplitude of DPI and P-CR signals recorded by OpenIrisDPI to quantify frame-to-frame noise during our electrophysiological experiments (n=11). Our DPI signal exhibited a broad peak in spectral power around 100 Hz, which is consistent with prior reports of ocular tremor [28]. Therefore, f_c was chosen to be 200 Hz to minimize contamination by biological signals (Fig. 6a). Using this threshold, we found that the WN-RMS amplitudes of the DPI signals ($\text{WN-RMS}_{\text{dpi,az}} = 0.39' \pm 0.06'$, $\text{WN-RMS}_{\text{dpi,el}} = 0.44' \pm 0.07'$) were approximately a third of the magnitude of the P-CR signals ($\text{WN-RMS}_{\text{p-cr,az}} = 1.55' \pm 0.20'$, $\text{WN-RMS}_{\text{p-cr,el}} = 1.82' \pm$

0.26′) (Fig. 6b). These differences were statistically significant ($p < 10^{-3}$, Wilcoxon signed-rank test). Furthermore, the noise amplitude of the azimuthal components of the DPI and P-CR signals were consistently lower than that of the elevational components ($p < 10^{-3}$, Wilcoxon signed-rank test), which may be due to the oblique optical configuration. Finally, we observed a significant difference in P-CR WN-RMS amplitude between Monkey C and Monkey O ($p < 10^{-2}$, Mann-Whitney U test on geometric mean of $\text{WN-RMS}_{\text{az}}$ and $\text{WN-RMS}_{\text{el}}$), which may be due to differences in head position between the two animals. There was no statistically significant difference in DPI noise floors between the two monkeys ($p = .16$, Mann-Whitney U test). Thus, the data confirm that DPI signals are less susceptible to temporally uncorrelated measurement noise than P-CR signals when computed from identical video inputs, at least when using the OpenIrisDPI system under the conditions of this study.

3.3.2. P-CR signals report more ocular drift between saccades than DPI signals

In addition to having greater high-frequency noise, P-CR signals also drifted more than DPI signals do (Fig. 7a & b). One possibility is that this excess drift was due to biological factors such as pupil decentration [59]. To address this possibility, we assumed that measurement noise is uncorrelated between successive frames and that biological noise varies slowly over time. To isolate this slow component, we calculated the mean squared displacement (MSD) as a function of lag and compared its magnitude between DPI and P-CR signals (Fig. 7c).

MSD as a function of temporal lag is commonly used to estimate the parameters of a Brownian motion model of intersaccadic drift [15]. Under this model, the slope of the line of best fit is four times the diffusion coefficient of the Brownian motion [24]. We presume that the motion of the eye is identical whether measured by P-CR or by DPI, so we expect that this slope will be the same for both techniques. Note that stationary white noise will increase the MSD for all lags equally, and as a result this prediction holds even if the level of measurement noise differs between techniques. What we observed, however, is that the diffusion coefficient was 17.12 ± 1.09 arcmin^2/s larger ($p < 10^{-4}$, Wilcoxon signed-rank test) when calculated from P-CR data than from DPI data (Fig. 7d).

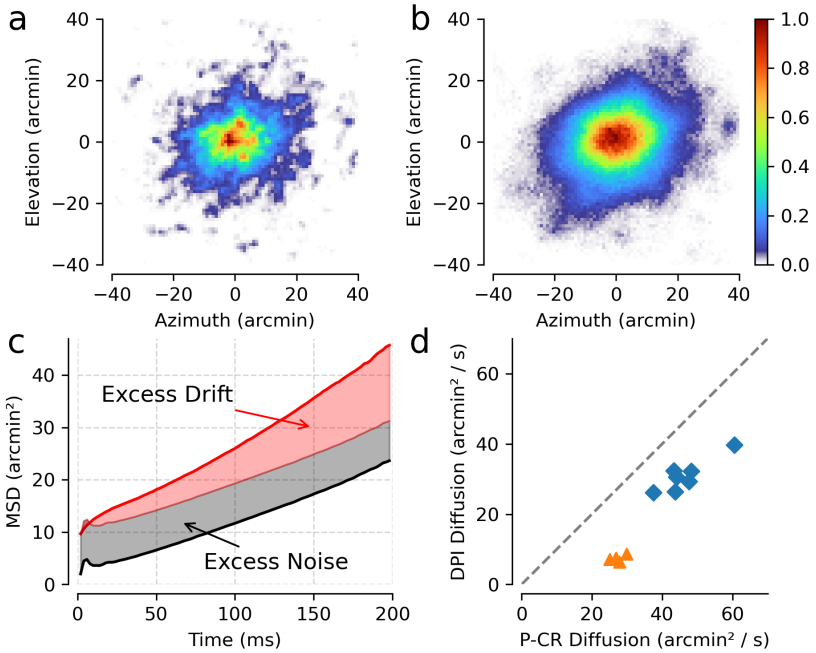


Figure 7: P-CR signals are more dispersed than DPI signals during intersaccadic drift. a) Normalized distribution of gaze direction during fixations measured using the DPI method. b) Same as in a, but using simultaneously collected P-CR data. c) Mean squared displacement (MSD) over various delays for DPI (black) and P-CR (red) signals calculated during intersaccadic drift. Temporally uncorrelated system noise will result in a vertical offset between the two traces (shaded gray) but does not produce the obseved difference in slope (shaded red). d) Diffusion coefficients estimated using the P-CR (abscissa) and DPI (ordinate) methods. Blue squares indicate data from monkey O, orange triangles indicate data from monkey C.

3.4. DPI eye tracking improves receptive field characterization in fixating macaques

We sought to demonstrate the benefits of using a DPI eye tracker for in vivo neurophysiology by analyzing responses to visual stimulation in the LGN of macaque monkeys. Neurons in the LGN receive direct input from the retina and have punctate receptive fields that range from several to tens of arcminutes in diameter [9]. We recorded from populations of LGN neurons while the macaques were tasked with fixating a stationary point (Fig. 8a). During periods of stable fixation, a white noise stimulus was displayed over the recorded population's receptive fields. In offline analysis, we attempted

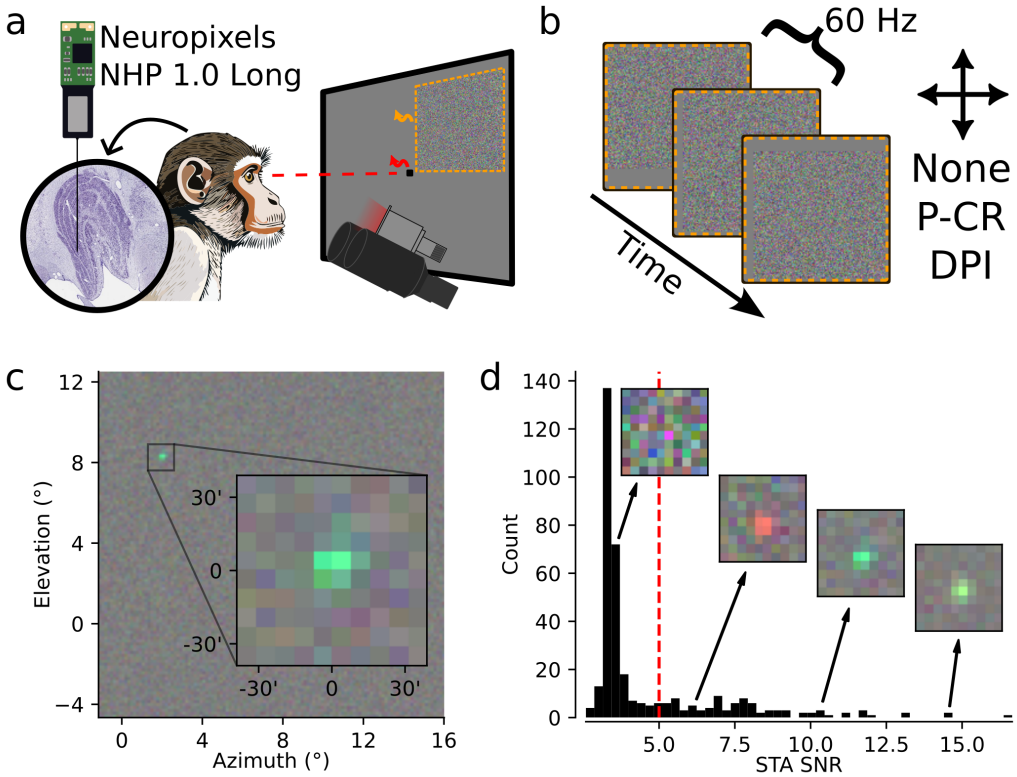


Figure 8: Electrophysiological experimental design and analysis. a) A 45 mm-long Neuropixels NHP 1.0 probe was inserted into the LGN of a macaque fixating a centrally projected dot (gaze drawn in red). Gaze direction was monitored using OpenIrisDPI. During periods of fixation, a white noise stimulus was displayed over the receptive fields of the LGN neurons recorded by the probe (dashed orange square). b) During analysis, the stimulus was reconstructed and the average stimulus preceding every spike of each algorithmically identified spike cluster was calculated. The stimulus was then shifted to account for the effects of fixational eye movements on the retinal movie and the analysis was repeated. c) The peak lag of a representative STA of a presumptive parvocellular neuron without correction for residual eye movements. Inset shows the STA cropped around the RF. d) Histogram of SNR values for each cluster in a representative recording. Insets show cropped STAs, and arrows point to the corresponding SNR values. Only STAs with SNR exceeding a threshold (red dashed line) were included in subsequent analyses. Inset in (a) reproduced from brainmaps.org [37].

to compensate for small eye movements made during fixation using either the DPI or P-CR signals recorded by OpenIrisDPI (Fig. 8b).

Reverse correlating the recorded spikes with these compensated stim-

uli showed that DPI signals greatly improved estimates of neuronal tuning compared to either no compensation—that is, ignoring eye movements—or compensation using P-CR signals. Each recording yielded an average of 405 clusters (range: 176-638), with an average of 68 clusters (range: 33-133) that produced valid receptive fields without gaze compensation (SNR > 5, Fig. 8c & d). We recalculated STAs by translating each stimulus frame by whole-element amounts to account for the gaze positions reported by either DPI or P-CR signals. This revealed that both DPI and P-CR signals improved RF estimates, but that DPI-based compensation significantly outperformed compensation using P-CR signals (Fig. 9). The SNR of STAs relative to uncorrected data improved by 87.4% $\pm$ 10.0% for DPI correction versus just 29.0% $\pm$ 9.2% for P-CR ($p < 10^{-3}$, Wilcoxon signed-rank test). Similarly, DPI correction produced significantly smaller receptive fields in 10 of 11 datasets, with RF radii contracting by 1.66' $\pm$ 0.38' on average for DPI versus 0.56' $\pm$ 0.27' for P-CR correction ($p < 5 \times 10^{-3}$, Wilcoxon signed-rank test). These analyses confirmed that OpenIrisDPI can be used to compensate for fixational eye movements and that DPI signals are superior to P-CR signals for this application.

4. Discussion

We have described a new open-source eye tracking system, OpenIrisDPI, with the goal of increasing access to DPI eye tracking for visual neuroscience. DPI eye tracking is a non-invasive and highly precise technique that faces limited adoption due to a lack of convenient, cost-effective, and widely available systems. OpenIrisDPI enables researchers to construct custom DPI systems using inexpensive, easily obtained consumer electronics. The software required to operate it is free and includes a graphical user interface alongside common features of commercial eye tracking systems [47]. As a result, OpenIrisDPI significantly reduces the barriers to the adoption of DPI eye tracking.

4.1. Validation of OpenIrisDPI's performance

To validate the usefulness of OpenIrisDPI for high-precision applications, we first tested its performance using an artificial eye. The noise amplitude of the P-CR signal was equivalent to that of the Eyelink 1000 Plus reported by Niehorster et al. (2021) [38], and the DPI signal was an order of magnitude less noisy. These results confirmed that the OpenIrisDPI P-CR signal

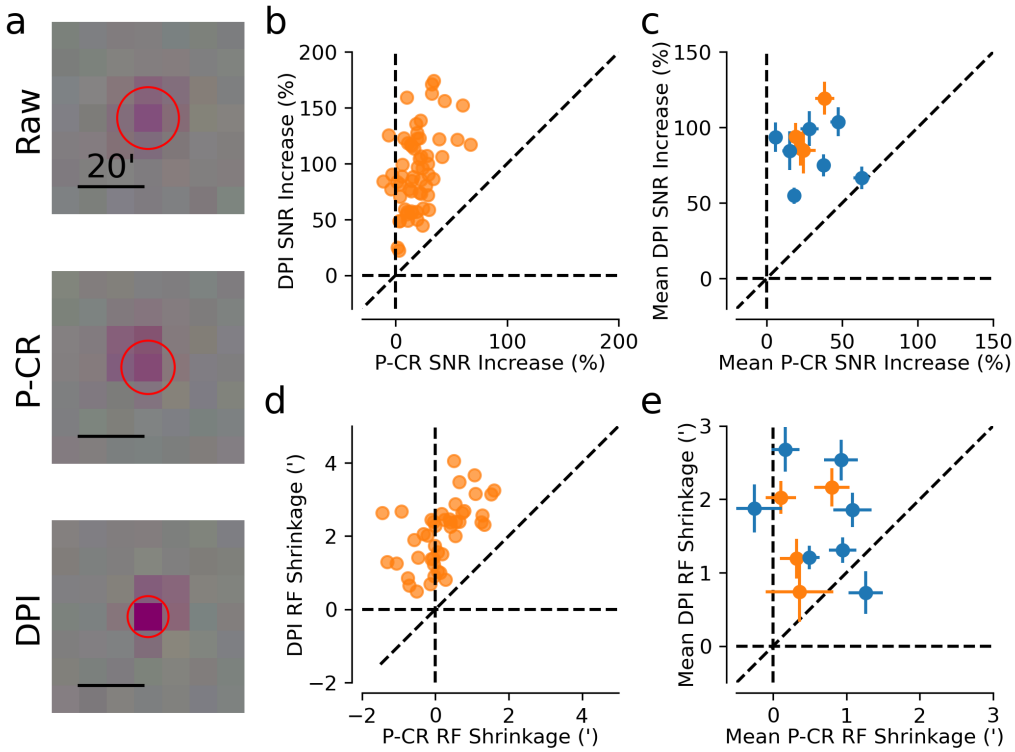


Figure 9: Comparison of DPI- and P-CR-compensated receptive field measurements of LGN neurons in fixating macaques. a) STA of an example cluster with no compensation (top), P-CR compensation (center), and DPI compensation (bottom). Red circle indicates one σ of a Gaussian fit. Scale bar is 20 arcminutes. b) STA SNR improvement relative to baseline for all clusters with valid receptive fields in an exemplar recording using P-CR (abcissa) and DPI (ordinate) signals. c) Mean STA SNR improvement for all datasets. Error bars represent $\pm$ 1.96 SEM. d) Reduction of STA radius compared to baseline for all clusters shown in b) using P-CR (abcissa) and DPI (ordinate) correction. e) Mean reduction of STA RF radius for each dataset. Error bars same as in (c). Color indicates animal identifier as in Fig. 7.

is representative of common eye tracking systems and that the DPI signal contained less measurement noise when tracking model eyes. We note that our DPI eye tracker was approximately four times noisier than the one developed by Wu et al. (2022) [57]. This may have been due to the fact that we used lower magnification to track the pupil, which was necessary for the direct comparison of DPI and P-CR signals.

We then compared DPI and P-CR signals recorded from fixating macaque

monkeys by isolating two distinct components of the eye position signal between saccades: temporally uncorrelated (white) noise and slowly varying drift. The white noise signal amplitude of P-CR signals was three times greater than those of simultaneously recorded DPI signals. Since both signals were calculated from identical video inputs, we can exclude hardware differences as causative. We conclude that DPI tracking is less prone to measurement noise than P-CR tracking when applied to living eyes, but we cannot rule out the possibility that improvements to the pupil fitting algorithm might narrow this gap.

Our analysis of slow variations in eye position showed that P-CR signals drift more than DPI signals. These results are consistent with prior reports comparing simultaneously collected P-CR and scleral search coil signals [27, 35]. The drift we observed cannot be attributed to sensor noise and likely reflects a biological process. Since both DPI and P-CR signals use the CR, the increased fixational instability we observed must result from drift of the pupil center relative to P4. This conclusion is consistent with previous studies in which varying background illumination caused P-CR eye trackers to report artifactual deviations in eye position [14, 27, 56, 59]. Our results extend this understanding by demonstrating that ongoing pupillary dynamics can cause increased drift even when background luminance is constant. Some studies have used P-CR eye trackers to estimate the parameters of intersaccadic drift [11, 12, 19]. It may be necessary to revisit the conclusions of such studies in light of these results.

It is worth noting that our analyses focused on low-amplitude eye movements between saccades because DPI is ill-suited for measuring saccade kinematics [10]. The rapid acceleration and deceleration of the eye during saccades cause the lens to oscillate, which manifests as ringing in the DPI trace at the beginning and end of saccades [10]. This ringing is not representative of the direction of gaze, although lens wobble is expected to affect the image on the retina [53].

4.2. A DPI eye tracker for visual neuroscience

Our neurophysiological recordings from the macaque LGN demonstrate the value of OpenIrisDPI for visual neuroscience research. Awake macaques are a gold-standard animal model for studying primate vision. However, ocular fixation is imperfect even in well-trained macaques, and residual eye movements can affect measurements of neuronal tuning to visual stimuli. We show that OpenIrisDPI can be used to compensate for these fixational

eye movements, as evidenced by the increase in signal-to-noise ratios and contraction of receptive field estimates.

Previous studies have attempted to compensate for fixational eye movements using other eye tracking systems with varying degrees of success. Tang et al. (2007) used search coils sutured to the sclera [54]. They reported a reduction in RF widths of 4 arcminutes in V1 and the LGN, which is consistent with our results. However, Read and Cumming used a similar technique to study V1 and discovered significant inaccuracies in the coil system that counteracted any beneficial effects [44]. A similar result was reported by McFarland et al. (2014), who inferred the eye's position using neural activity and found it to be inconsistent with scleral search coils signals [36]. Using search coils to compensate for fixational eye movements is therefore possible but unreliable, which is concerning given the invasive nature of this technique. OpenIrisDPI, in contrast, is both noninvasive and straightforward to implement, as proven by its successful deployment in five laboratories at the time of publication.

OpenIrisDPI may also enable new paradigms for studying neural function in awake animals. Yates et al. (2023) showed that DPI eye tracking can be used to study foveal receptive fields in V1 of freely-viewing marmosets [61]. Studies that build on this fixation-free methodology are poised to shed light on phenomena at the intersection of natural oculomotor behavior and visual processing, including saccadic suppression [33] and oculomotor-related visual illusions [41, 55]. Furthermore, cognitive processes, like visual attention, impact the statistics of fixational eye movements [15, 24] and modulate neural activity in early visual areas [45, 46]. High-precision eye tracking can be used to disambiguate the feed-forward effects of eye movements from modulatory cognitive feedback, which may benefit the study of such cognitive process.

Finally, OpenIrisDPI may also be a useful tool for studying binocular coordination and vergence eye movements. Recent research by Jaschinski (2016) and Hooge et al. (2019) has demonstrated that artifacts caused by pupil-decentration are particularly troublesome when measuring vergence eye movements using pupil-based video eye trackers [21, 26]. DPI eye trackers overcome this limitation by being insensitive to unwanted pupillary motion. OpenIrisDPI's ability to track both eyes simultaneously and to serve as a convenient, cost-effective replacement for existing pupil-based systems make it an appealing option for studying binocular eye movements with high precision.

Acknowledgments

This work was funded by NIH EY032900 to Gregory D Horwitz, NIH EY032179 to Jacob L Yates, and an NSF GRFP fellowship to Ryan Ressmeyer.

Declaration of generative AI an AI-assisted technologies in the writing process

During the preparation of this work, the authors used Claude.ai for revisions. After using this tools, the authors reviewed and edited the content as need and take full responsibility for the content of the published article.

References

[1] Evgeniy Abdulin, Lee Friedman, and Oleg Komogortsev. "Custom video-oculography device and its application to fourth purkinje image detection during saccades". In: *arXiv preprint arXiv:1904.07361* (2019).

[2] Jorge Ares and Justo Arines. "Influence of thresholding on centroid statistics: full analytical description". In: *Applied optics* 43.31 (2004), pp. 5796–5805.

[3] Wyeth Bair and Lawrence P O'Keefe. "The influence of fixational eye movements on the response of neurons in area MT of the macaque". In: *Visual neuroscience* 15.4 (1998), pp. 779–786.

[4] Pieter Blignaut and Tanya Beelders. "The precision of eye-trackers: a case for a new measure". In: *Proceedings of the symposium on eye tracking research and applications*. 2012, pp. 289–292.

[5] Norick R Bowers, Alexandra E Boehm, and Austin Roorda. "The effects of fixational tremor on the retinal image". In: *Journal of vision* 19.11 (2019), pp. 8–8.

[6] EJ Chichilnisky. "A simple white noise analysis of neuronal lightresponses". In: *Network: computation in neural systems* 12.2 (2001), p. 199.

[7] Ashley M Clark et al. "Eye drift during fixation predicts visual acuity". In: *Proceedings of the National Academy of Sciences* 119.49 (2022), e2200256119.

[8] Hewitt D Crane and Carroll M Steele. "Generation-V dual-Purkinje-image eyetracker". In: *Applied optics* 24.4 (1985), pp. 527–537.

[9] Lisa J Croner and Ehud Kaplan. "Receptive fields of P and M ganglion cells across the primate retina". In: *Vision research* 35.1 (1995), pp. 7–24.

[10] Heiner Deubel and Bruce Bridgeman. "Fourth Purkinje image signals reveal eye-lens deviations and retinal image distortions during saccades". In: *Vision research* 35.4 (1995), pp. 529–538.

[11] Leandro L Di Stasi et al. "Intersaccadic drift velocity is sensitive to short-term hypobaric hypoxia". In: *European Journal of Neuroscience* 39.8 (2014), pp. 1384–1390.

[12] Leandro L Di Stasi et al. "Microsaccade and drift dynamics reflect mental fatigue". In: *European Journal of Neuroscience* 38.3 (2013), pp. 2389–2398.

[13] Jan Drewes, Guillaume S Masson, and Anna Montagnini. "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*. 2012, pp. 209–212.

[14] Jan Drewes et al. "Smaller is better: Drift in gaze measurements due to pupil dynamics". In: *PloS one* 9.10 (2014), e111197.

[15] Ralf Engbert and Reinhold Kliegl. "Microsaccades keep the eyes' balance during fixation". In: *Psychological science* 15.6 (2004), pp. 431–431.

[16] Moshe Gur, Alexander Beylin, and D Max Snodderly. "Response variability of neurons in primary visual cortex (V1) of alert monkeys". In: *Journal of Neuroscience* 17.8 (1997), pp. 2914–2920.

[17] Moshe Gur and D Max Snodderly. "High response reliability of neurons in primary visual cortex (V1) of alert, trained monkeys". In: *Cerebral cortex* 16.6 (2006), pp. 888–895.

[18] Moshe Gur and D Max Snodderly. "Studying striate cortex neurons in behaving monkeys: benefits of image stabilization". In: *Vision research* 27.12 (1987), pp. 2081–2087.

[19] Carl JJ Herrmann, Ralf Metzler, and Ralf Engbert. "A self-avoiding walk with neural delays as a model of fixational eye movements". In: *Scientific reports* 7.1 (2017), p. 12958.

[20] Kenneth Holmqvist and Pieter Blignaut. "Small eye movements cannot be reliably measured by video-based P-CR eye-trackers". In: *Behavior research methods* 52 (2020), pp. 2098–2121.

[21] Ignace TC Hooge, Roy S Hessels, and Marcus Nyström. "Do pupil-based binocular video eye trackers reliably measure vergence?" In: *Vision Research* 156 (2019), pp. 1–9.

[22] David H Hubel and Torsten N Wiesel. "Integrative action in the cat's lateral geniculate body". In: *The Journal of physiology* 155.2 (1961), p. 385.

[23] David H Hubel and Torsten N Wiesel. "Receptive fields, binocular interaction and functional architecture in the cat's visual cortex". In: *The Journal of physiology* 160.1 (1962), p. 106.

[24] Janis Intoy and Michele Rucci. "Finely tuned eye movements enhance visual acuity". In: *Nature communications* 11.1 (2020), p. 795.

[25] *Accuracy (trueness and precision) of measurement methods and results — Part 1: General principles and definitions*. Standard. Geneva, CH: International Organization for Standardization, July 2023.

[26] Wolfgang Jaschinski. "Pupil size affects measures of eye position in video eye tracking: implications for recording vergence accuracy". In: *Journal of Eye Movement Research* 9.4 (2016).

[27] Daniel L Kimmel, Dagem Mammo, and William T Newsome. "Tracking the eye non-invasively: simultaneous comparison of the scleral search coil and optical tracking techniques in the macaque monkey". In: *Frontiers in behavioral neuroscience* 6 (2012), p. 49.

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

[29] David A Leopold and Nikos K Logothetis. "Microsaccades differentially modulate neural activity in the striate and extrastriate visual cortex". In: *Experimental Brain Research* 123 (1998), pp. 341–345.

[30] Jonathan B Levitt, Daniel C Kiper, and J Anthony Movshon. "Receptive fields and functional architecture of macaque V2". In: *Journal of neurophysiology* 71.6 (1994), pp. 2517–2542.

[31] Margaret S Livingstone and Doris Y Tsao. "Receptive fields of disparity-selective neurons in macaque striate cortex". In: *Nature neuroscience* 2.9 (1999), pp. 825–832.

[32] Conny Lu et al. "Improved vergence and accommodation via purkinje image tracking with multiple cameras for ar glasses". In: *2020 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*. IEEE. 2020, pp. 320–331.

[33] Ethel Matin. "Saccadic suppression: a review and an analysis." In: *Psychological bulletin* 81.12 (1974), p. 899.

[34] John HR Maunsell et al. "Visual response latencies of magnocellular and parvocellular LGN neurons in macaque monkeys". In: *Visual neuroscience* 16.1 (1999), pp. 1–14.

[35] Michael B McCamy et al. "Simultaneous recordings of human microsaccades and drifts with a contemporary video eye tracker and the search coil technique". In: *PLoS One* 10.6 (2015), e0128428.

[36] James M McFarland et al. "High-resolution eye tracking using V1 neuron activity". In: *Nature communications* 5.1 (2014), p. 4605.

[37] Shawn Mikula et al. "Internet-enabled high-resolution brain mapping and virtual microscopy". In: *Neuroimage* 35.1 (2007), pp. 9–15.

[38] Diederick C Niehorster, Raimondas Zemblys, and Kenneth Holmqvist. "Is apparent fixational drift in eye-tracking data due to filters or eyeball rotation?" In: *Behavior Research Methods* 53 (2021), pp. 311–324.

[39] Marcus Nyström and Kenneth Holmqvist. "An adaptive algorithm for fixation, saccade, and glissade detection in eyetracking data". In: *Behavior research methods* 42.1 (2010), pp. 188–204.

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

[41] Jorge Otero-Millan, Stephen L Macknik, and Susana Martinez-Conde. "Microsaccades and blinks trigger illusory rotation in the "rotating snakes" illusion". In: *Journal of Neuroscience* 32.17 (2012), pp. 6043–6051.

[42] Marius Pachitariu et al. "Spike sorting with Kilosort4". In: *Nature Methods* (2024), pp. 1–8.

[43] Raghuveer Parthasarathy. "Rapid, accurate particle tracking by calculation of radial symmetry centers". In: *Nature methods* 9.7 (2012), pp. 724–726.

[44] Jenny CA Read and Bruce G Cumming. "Measuring V1 receptive fields despite eye movements in awake monkeys". In: *Journal of neurophysiology* 90.2 (2003), pp. 946–960.

[45] John H Reynolds, Tatiana Pasternak, and Robert Desimone. "Attention increases sensitivity of V4 neurons". In: *Neuron* 26.3 (2000), pp. 703–714.

[46] Pieter R Roelfsema, Victor AF Lamme, and Henk Spekreijse. "Object-based attention in the primary visual cortex of the macaque monkey". In: *Nature* 395.6700 (1998), pp. 376–381.

[47] Roksana Sadeghi et al. "Open Iris-An Open Source Framework for Video-Based Eye-Tracking Research and Development". In: *bioRxiv* (2024).

[48] Stefan Schuster et al. "Noise Variance and Signal-to-Noise Ratio Estimation from Spectral Data". In: *2019 IEEE International Instrumentation and Measurement Technology Conference (I2MTC)*. IEEE. 2019, pp. 1–6.

[49] Joshua H Siegle et al. "Open Ephys: an open-source, plugin-based platform for multichannel electrophysiology". In: *Journal of neural engineering* 14.4 (2017), p. 045003.

[50] D Max Snodderly. "A physiological perspective on fixational eye movements". In: *Vision research* 118 (2016), pp. 31–47.

[51] D Max Snodderly, Igor Kagan, and Moshe Gur. "Selective activation of visual cortex neurons by fixational eye movements: implications for neural coding". In: *Visual neuroscience* 18.2 (2001), pp. 259–277.

[52] DM Snodderly and Moshe Gur. "Organization of striate cortex of alert, trained monkeys (Macaca fascicularis): ongoing activity, stimulus selectivity, and widths of receptive field activating regions". In: *Journal of Neurophysiology* 74.5 (1995), pp. 2100–2125.

[53] Juan Tabernero and Pablo Artal. "Lens oscillations in the human eye. Implications for post-saccadic suppression of vision". In: *PloS one* 9.4 (2014), e95764.

[54] Yamei Tang et al. "Eye position compensation improves estimates of response magnitude and receptive field geometry in alert monkeys". In: *Journal of neurophysiology* 97.5 (2007), pp. 3439–3448.

[55] Xoana G Troncoso et al. "Microsaccades drive illusory motion in the Enigma illusion". In: *Proceedings of the National Academy of Sciences* 105.41 (2008), pp. 16033–16038.

[56] Ulrich Wildenmann and Frank Schaeffel. "Variations of pupil centration and their effects on video eye tracking". In: *Ophthalmic and Physiological Optics* 33.6 (2013), pp. 634–641.

[57] Ruei-Jr Wu et al. "High-resolution oculomotor measurements via a digital Dual Purkinje Image eye-tracker". In: *Journal of Vision* 22.14 (2022), pp. 4203–4203.

[58] Harry J Wyatt. "The form of the human pupil". In: *Vision research* 35.14 (1995), pp. 2021–2036.

[59] Harry J Wyatt. "The human pupil and the use of video-based eyetrackers". In: *Vision research* 50.19 (2010), pp. 1982–1988.

[60] Will Xiao et al. "Feature-selective responses in macaque visual cortex follow eye movements during natural vision". In: *Nature neuroscience* (2024), pp. 1–10.

[61] Jacob L Yates et al. "Detailed characterization of neural selectivity in free viewing primates". In: *Nature Communications* 14.1 (2023), p. 3656.

Appendix A. Defining Accuracy, Precision, and Trueness for Eye Tracking.

Throughout this paper, we adopt the definitions for accuracy, precision, and trueness outlined by the International Organization for Standardiza-

tion[25]. Under this definition, the accuracy of a measurement refers to its total deviation from a true value of interest. Accuracy is then decomposed into two parts: precision and trueness. Precision refers to the tendency for repeated measurements to be similar to one another, whereas trueness refers to the closeness of the expectation of any given measurement to the true value of interest. With these definitions in hand, we define how precision and trueness relate to eye tracking systems as follows.

The true value of interest for head-centric eye tracking is typically either the orientation of the globe or the direction of gaze. Though related, these two values are not identical. A full description of the globe's orientation requires knowing three rotational components (pitch, yaw, and roll), whereas just two components are required to uniquely define the ray projecting from the preferred retinal locus (elevation and azimuth). In this study, we ignore torsional eye motion and only estimate the direction of gaze. 'Trueness' in this context describes the accuracy of calibration, or, equivalently, the accuracy of the function mapping from eye tracker readings to gaze vectors. 'Precision' then refers to the ability of an eye tracker to produce similar readings given a fixed true gaze vector.

Consideration should be given to the challenges present when attempting to quantify either the trueness or precision of an eye tracking system. The eyes of alert subjects are never still, even when a subject is instructed to maintain fixation. Thus, during typical experiments there is inherent uncontrolled variability in the true value of interest. Because the eye cannot be held perfectly steady, many metrics have been developed that aim to quantify imprecision during experiments [4]. However, these methods often conflate fixational motion with imprecision (as in the bivariate contour ellipse), or may ignore low-frequency sources of imprecision (as is true for the sample-to-sample RMS). In this study, we address these issues by explicitly stating our assumption that measurement noise is additive and white to isolate this noise from other sources.

Appendix B. The OpenIrisDPI Feature Tracking Algorithm.

Pseudocode for OpenIrisDPI's tracking algorithm is given below in Algorithm 1. The algorithm first detects the pupil boundary by computing the convex hull of all points on the edge of the region that contains pixels the values of which are below user-defined threshold. An ellipse is then fit to the identified points to determine the pupil center, width, height, and major

axis angle. Next, the approximate location of the CR, which is usually the brightest point in the image, is computed as the centroid of pixels above a threshold set close to the sensor's saturation level. This coarse estimate is then refined by recomputing the thresholded-rectified centroid of a local region around the estimated center. Finally, the approximate position of P4 is found by identifying the highest amplitude pixel within the pupil, excluding the CR. The sub-pixel location of P4 is then found by computing the threshold-rectified centroid in a region surrounding the approximate center. We use the threshold-rectified centroid because we found it to be faster and more accurate in simulations than the radially symmetric center algorithm for reasonable thresholds (Fig. B.1).

Algorithm 1 OpenIrisDPI Algorithm.

Require: Image I, crop dimensions D, blur radius R_b, pupil threshold T_{pup}, pupil search radius R_{pup}, CR threshold T_{cr}, CR ROI radius R_{cr}, pupil erode radius R_{erode}, P4 ROI radius R_{P4}

Ensure: Pupil parameters, CR position, P4 position

function ThresholdedCenterOfMass(patch, threshold)
  $patch_{sub} \leftarrow patch - threshold$
  $patch_{sub}[patch_{sub} < 0] \leftarrow 0$
  return Centroid($patch_{sub}$)
end function

$I_{crop} \leftarrow \text{Crop}(I, D)$ ▷ Stage 1: Preprocess
$I_{blur} \leftarrow \text{GaussianBlur}(I_{crop}, R_b)$

$I_{thresh} \leftarrow I_{blur} < T_{pup}$ ▷ Stage 2: Detect Pupil Boundary
$center_{approx} \leftarrow \text{Centroid}(I_{thresh})$
$mask_{pup} \leftarrow \text{CircularMask}(center_{approx}, R_{pup})$
$I_{lap} \leftarrow \text{Laplacian}(I_{thresh} \wedge mask_{pup})$
$I_{edge} \leftarrow I_{lap} > 0$
$contour \leftarrow \text{ConvexHull}(I_{edge})$

$pupil \leftarrow \text{FitEllipse}(contour)$ ▷ Stage 3: Fit Pupil

$I_{cr} \leftarrow I_{blur} \wedge mask_{pup} > T_{cr}$ ▷ Stage 4: Localize CR
$CR_{approx} \leftarrow \text{Centroid}(I_{cr})$
$roi_{cr} \leftarrow \text{ExtractROI}(I_{blur}, CR_{approx}, R_{cr})$
$CR \leftarrow \text{ThresholdedCenterOfMass}(roi_{cr}, T_{cr})$

$I_{masked} \leftarrow \text{MaskRegion}(I_{blur}, CR, R_{cr})$ ▷ Stage 5: Localize P4
$mask_{pup} \leftarrow \text{FillContour}(contour)$
$mask_{pup} \leftarrow \text{Erode}(mask_{pup}, R_{erode})$
$I_{p4} \leftarrow I_{masked} \wedge mask_{pup}$
$P4_{approx} \leftarrow \text{ArgMax}(I_{p4})$
$roi_{p4} \leftarrow \text{ExtractROI}(I_{blur}, P4_{approx}, R_{P4})$
$P4 \leftarrow \text{ThresholdedCenterOfMass}(roi_{p4}, T_{pup})$
return $pupil$, CR, $P4$

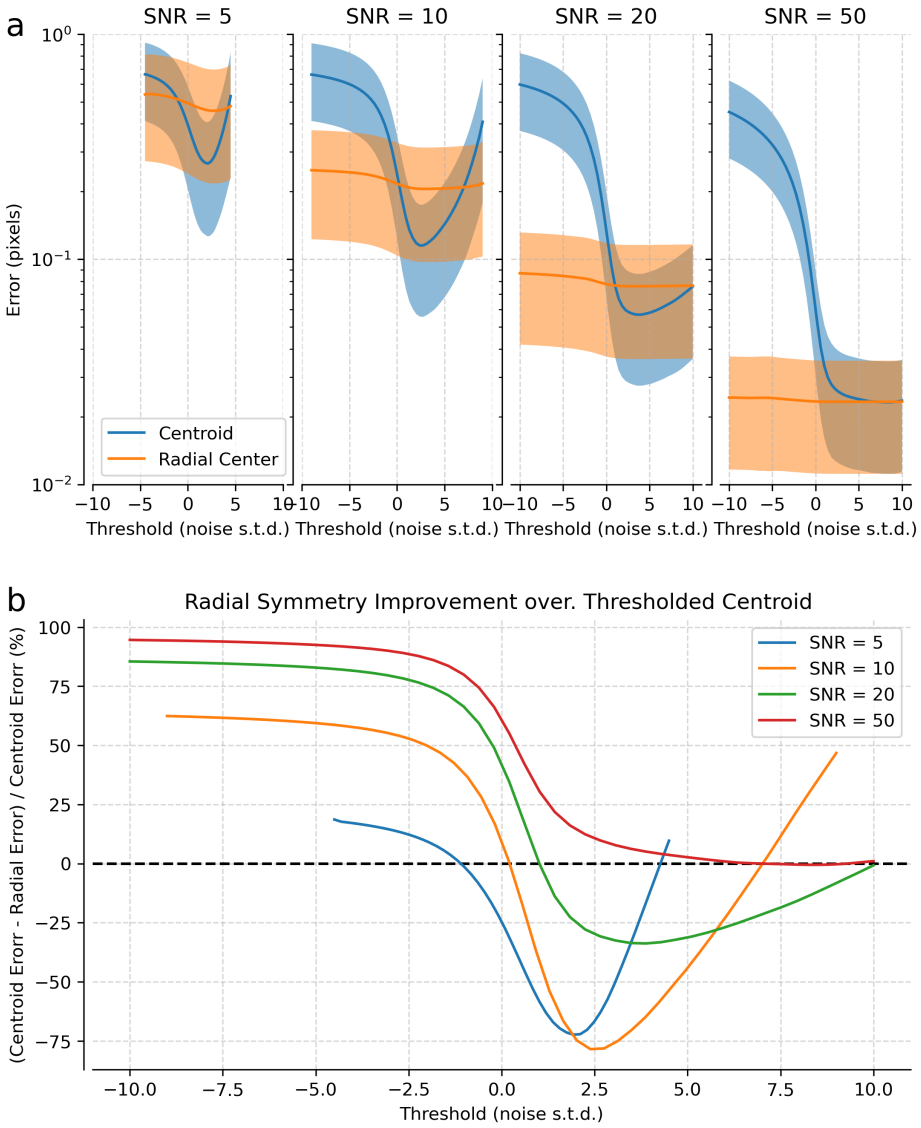


Figure B.1: Simulations of localization error using thresholded center of mass and radially symmetric center algorithms. a) Localization error across different signal-to-noise ratios (SNR = 5, 10, 20, and 50). Images were simulated as 20 x 20 pixel Gaussian blobs with additive noise ($I[i,j] = G(i,j) + \eta$, $\eta \sim \mathcal{N}(0,1)$, $G(i,j) = \text{SNR} \times \exp(-\frac{(i-10)^2+(j-10)^2}{6})$). Localization error was defined as the Euclidean distance between the estimated and true center of $G(i,j)$. Lines show mean $\pm$ std of localization error across 10,000 simulated images as a function of threshold value for the thresholded center of mass (blue) and radially symmetric center (orange). b) Relative performance difference between methods, normalized by thresholded center of mass error. At optimal threshold values, the radially symmetric center is as or more error prone than the thresholded center of mass.