

RESEARCH ARTICLE

Modeling visual performance differences ‘around’ the visual field: A computational observer approach

Eline R. Kupers^{1*}, Marisa Carrasco^{1,2}, Jonathan Winawer^{1,2}

1 Department of Psychology, New York University, New York, New York, United States of America, **2** Center for Neural Science, New York University, New York, New York, United States of America

* eline.kupers@nyu.edu



OPEN ACCESS

Citation: Kupers ER, Carrasco M, Winawer J (2019) Modeling visual performance differences ‘around’ the visual field: A computational observer approach. *PLoS Comput Biol* 15(5): e1007063. <https://doi.org/10.1371/journal.pcbi.1007063>

Editor: Wolfgang Einhäuser, Technische Universität Chemnitz, GERMANY

Received: October 4, 2018

Accepted: May 2, 2019

Published: May 24, 2019

Copyright: © 2019 Kupers et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: Both simulation and analysis, is publicly available via GitHub (<http://github.com/isetbio/JWLOrientedGabor>). In addition, the data structures created by the simulation and analyses are permanently archived on the Open Science Framework URL: <https://osf.io/mygvu/>.

Funding: This research was supported by the US National Eye Institute R01-EY027401 (MC and JW). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Abstract

Visual performance depends on polar angle, even when eccentricity is held constant; on many psychophysical tasks observers perform best when stimuli are presented on the horizontal meridian, worst on the upper vertical, and intermediate on the lower vertical meridian. This variation in performance ‘around’ the visual field can be as pronounced as that of doubling the stimulus eccentricity. The causes of these asymmetries in performance are largely unknown. Some factors in the eye, e.g. cone density, are positively correlated with the reported variations in visual performance with polar angle. However, the question remains whether these correlations can quantitatively explain the perceptual differences observed ‘around’ the visual field. To investigate the extent to which the earliest stages of vision—optical quality and cone density—contribute to performance differences with polar angle, we created a computational observer model. The model uses the open-source software package ISETBIO to simulate an orientation discrimination task for which visual performance differs with polar angle. The model starts from the photons emitted by a display, which pass through simulated human optics with fixational eye movements, followed by cone isomerizations in the retina. Finally, we classify stimulus orientation using a support vector machine to learn a linear classifier on the photon absorptions. To account for the 30% increase in contrast thresholds for upper vertical compared to horizontal meridian, as observed psychophysically on the same task, our computational observer model would require either an increase of ~7 diopters of defocus or a reduction of 500% in cone density. These values far exceed the actual variations as a function of polar angle observed in human eyes. Therefore, we conclude that these factors in the eye only account for a small fraction of differences in visual performance with polar angle. Substantial additional asymmetries must arise in later retinal and/or cortical processing.

Author summary

A fundamental goal in computational neuroscience is to link known facts from biology with behavior. Here, we considered visual behavior, specifically the fact that people are

Competing interests: The authors have declared that no competing interests exist.

better at visual tasks performed to the left or right of the center of gaze, compared to above or below at the same distance from gaze. We sought to understand what aspects of biology govern this fundamental pattern in visual behavior. To do so, we implemented a computational observer model that incorporates known facts about the front end of the human visual system, including optics, eye movements, and the photoreceptor array in the retina. We found that even though some of these properties are *correlated* with performance, they fall far short of *quantitatively explaining it*. We conclude that later stages of processing in the nervous system greatly amplify small differences in the way the eye samples the visual world, resulting in strikingly different performance around the visual field.

Introduction

Psychophysical performance differs with visual field position

Psychophysical performance is not uniform across the visual field. The largest source of this non-uniformity is eccentricity: acuity is much higher in the central visual field (fovea), limiting many recognition tasks such as reading and face recognition to only a relatively small portion of the retina. As a result, central visual field loss, such as macular degeneration, can be debilitating. Even a modest difference in eccentricity can have substantial effects on performance. For example, contrast thresholds on an orientation discrimination task approximately triple at 8° compared to 4° eccentricity [1, 2]. Similar effects are found for a wide range of tasks (for a review on peripheral vision, see [3]).

Interestingly, visual performance differs not only as a function of distance from the fovea (eccentricity), but also around the visual field (polar angle). The polar angle effects can be quite systematic. At a fixed eccentricity, contrast sensitivity and spatial resolution are better along the horizontal than the vertical meridian, and better along the lower than the upper vertical meridian. The two effects have been described by Carrasco and colleagues and called the “horizontal-vertical anisotropy” and the “vertical meridian asymmetry” [2, 4–8]. Such effects, often called *performance fields*, are found in numerous tasks, including contrast sensitivity and spatial resolution [5, 6, 9–23] (Fig 1), visual search [24–31], crowding [32–34], motion perception [35], visual short-term memory [36], contrast appearance [7] and spatial frequency appearance [37]. These effects can be large. For example, contrast thresholds at 8° eccentricity can be 5 times lower on the horizontal meridian compared to the vertical [5], a larger effect than doubling the eccentricity, from 4° to 8° [2].

The causes of performance differences with polar angle are not known. Some eye factors may contribute to them as they do to differences across eccentricity. For instance, the drop-off in density of cones and retinal ganglion cells with eccentricity contributes to decreased acuity [38, 39]. In this paper, we take a modeling approach to quantify the extent to which optics and photoreceptor sampling could plausibly contribute to the reported performance differences with polar angle.

Cone density differs with visual field position

In the human eye, cone density varies with eccentricity and polar angle. Foveal cones have relatively small diameters and are tightly packed, becoming sparser in the periphery due to both increased size and larger gaps between them [40, 41].

Cone density also differs as a function of polar angle at a fixed eccentricity. From ~2° to 7° eccentricity, density is about 30% greater on the horizontal than the vertical meridian [40, 41]

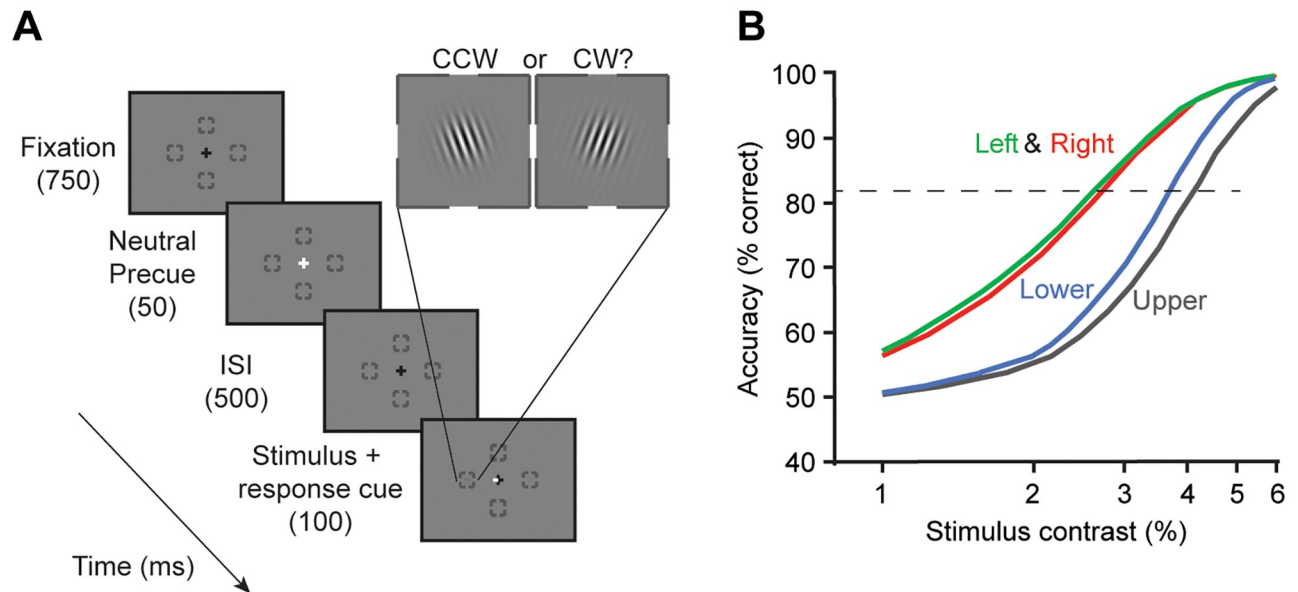


Fig 1. Example of psychophysical task and performance differences ‘around’ the visual field. (A) Experimental design of a two-alternative forced choice (2-AFC) orientation discrimination task. While observers maintained fixation, a brief cue appeared, and after a 500 ms ISI, a Gabor stimulus (4 cycles/°) was presented at one of four possible iso-eccentric locations (4.5° eccentricity). A response cue indicated to the observer the target location they were asked to make an orientation judgment on (clockwise/CW or counter-clockwise/CCW relative to vertical). Contrast varied across trials. (B) Contrast-dependent psychometric functions of an example observer show best performance for the two horizontal locations (Left and Right), and poorest performance for upper vertical. The increase in contrast thresholds (contrast at which performance reached 82% correct, dashed line) ranged from about 2.5% (Left and Right) to 4% (Upper). Data from figure 6 in Cameron, Tai, and Carrasco [5].

<https://doi.org/10.1371/journal.pcbi.1007063.g001>

(Fig 2). This 30% difference is about the same as the cone density decrease from 3° to 4° eccentricity along a single meridian. As a result, iso-density contours are elongated by about 30% in the vertical axis compared to horizontal. Because cone density is higher on the meridian where performance is better (horizontal compared to vertical), one might be tempted to conclude that cone density explains the performance difference. We return to this question in subsequent sections.

Optical quality differs with visual field position

Before light hits the retina, it has already been transformed by refraction from passing through different media (cornea, vitreous and aqueous humors), by diffraction from the pupil, as well as by optical aberrations (chromatic and achromatic) of the lens and intraocular light scattering (for an overview see [42]). These transformations reduce the optical quality of the image projected onto the retina.

Optical quality is not uniform across the retina [43, 44]. A clear, systematic effect is that both defocus and higher-order aberrations become worse with eccentricity. We assume that myopes and hyperopes wear corrective lenses to achieve good focus at the fovea. Defocus is the largest contribution to image quality [42]. The effects of defocus are largest in the far periphery, but still evident when comparing fovea to parafovea (e.g., 0 vs. 5°, Fig 3).

Most measurements of optical quality in human are either at the fovea or along the horizontal meridian. These measurements show that in addition to the decline in optical quality with eccentricity, there are also hemifield effects: For example, in the periphery, the temporal retina tends to have poorer optics than the nasal retina [44]. There are some [43], but many fewer measurements along the vertical meridian compared to the horizontal meridian. To our

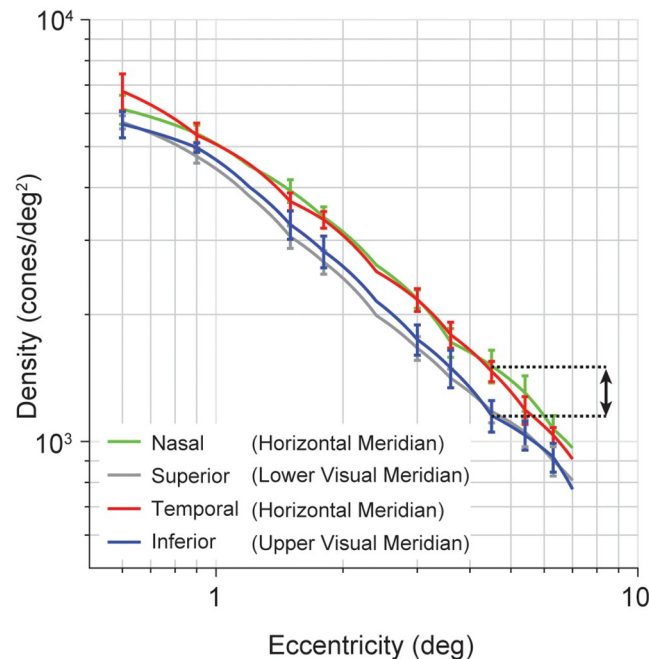


Fig 2. Variations in human cone density as a function of eccentricity and polar angle. Cone density falls off sharply with eccentricity, and also varies between horizontal and vertical retinal axes. Data are pooled across both eyes. Error bars represent standard error across 10 observers. Arrow indicates difference in cone density between horizontal and upper visual meridian at 4.5° eccentricity (matching Fig 1). Figure recomputed with data from observers between 22–35 years old (‘Group 1’) reported by Song et al. [41].

<https://doi.org/10.1371/journal.pcbi.1007063.g002>

knowledge, it is not yet firmly established whether there are systematic differences in optical quality between the vertical and horizontal meridians. However, the fact that optical quality varies with eccentricity as well as between nasal and temporal retina suggests that one should at least consider optics as a possible explanatory factor for performance differences around the visual field.

Quantifying the contribution of components in the eye to behavioral performance using a computational observer model

The meridian differences in cone density are correlated with meridian differences in psychophysical performance, with higher cone density and better performance on the horizontal axis (adjusted $r^2 = 0.88$, Fig 4) [40, 41]. However, a correlation does not necessarily imply an explanation. Without an explicit linking hypothesis or model that can predict how a difference in cone density should affect visual performance on a given task, we cannot know whether this correlation is meaningful for explaining this behavior. Should a decrease in cone density increase contrast thresholds? And if so, should a decrease of about 25% in cone density (difference between horizontal and vertical at 4.5° eccentricity) lead to an increase in contrast threshold of about 25%, as observed psychophysically?

Answering these questions requires a computational model. A computational model of the eye can quantify the extent of each component’s contribution on visual performance and potentially reveal which components limit performance on a given task.

In this study, we quantify the contribution of cone density and optical quality on visual performance according to a computational observer model. We then compare the modeled contributions to the observed quantities, and ask whether the observed differences in cone

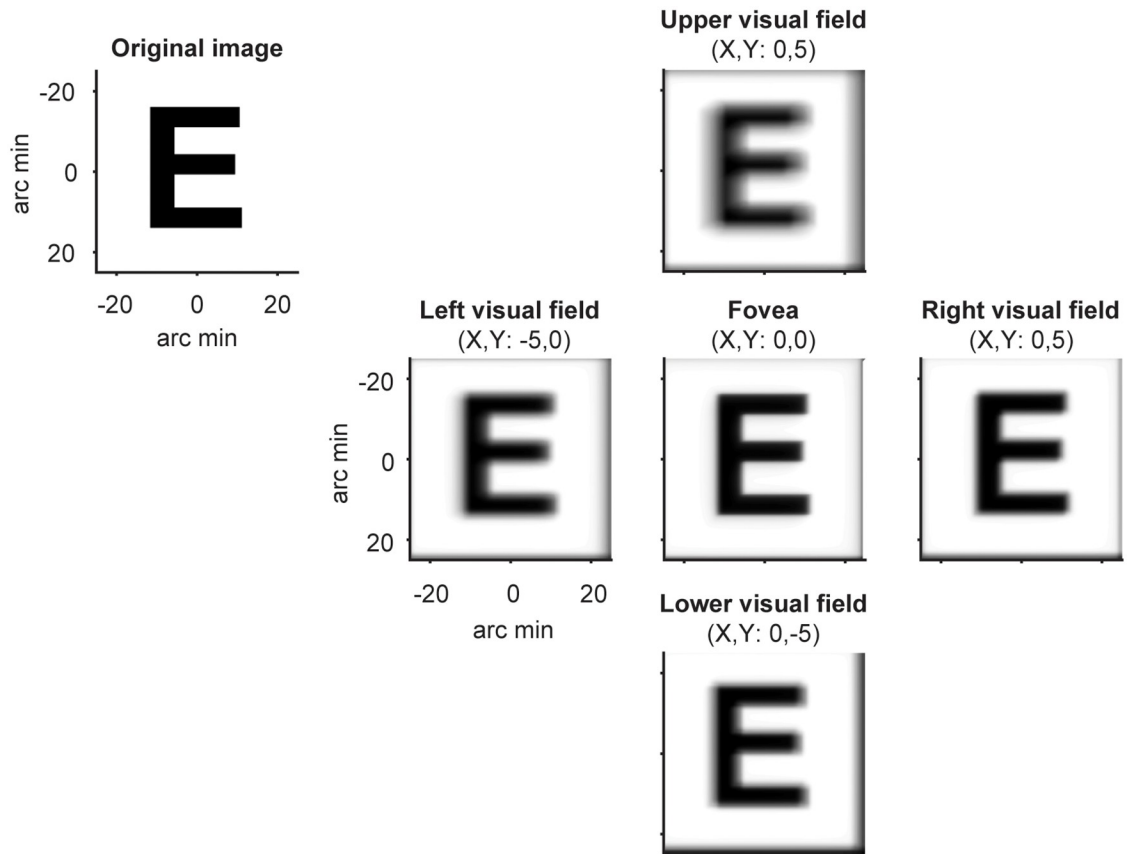


Fig 3. Variations in optical quality as a function of visual field location for example observer. The letter E was convolved with 5 location-specific point spread functions (PSFs) from an example observer, using wavefront measurements of image quality and correcting for the central refractive error. The wavefront measures are based on a prior study [43], and provided courtesy of Pablo Artal.

<https://doi.org/10.1371/journal.pcbi.1007063.g003>

density and optical quality as a function of polar angle can explain the observed differences in performance.

Several studies have proposed ideal observer models of visual tasks limited only by optics and photoreceptor sampling. When such ideal observer models are compared to human performance as a function of eccentricity, results show that these factors alone (optics and photoreceptor properties) cannot completely account for the differences in psychophysical performance (e.g. [45–48]). Generally, these studies find that human performance falls off more rapidly with eccentricity than would be predicted based only on optical and photoreceptor limits, indicating that there are additional downstream processes whose efficiency varies with eccentricity. These studies have not assessed the degree to which differences in performance as a function of polar angle are explained by front-end factors.

To implement the computational observer model of the human eye, we used the Image Systems Engineering Toolbox for Biology (ISETBIO [49–51]), a publicly available toolbox, to simulate encoding stages in the front-end of the human visual system (available at <http://isetbio.org/>). We used this model to simulate a 2-AFC orientation discrimination task using Gabor stimuli matched in parameters as reported by Cameron, Tai, and Carrasco [5]. Our computational observer model consists of multiple stages representing the front-end of the visual system: the spectral radiance of the experimental visual stimuli, optical quality of the lens and

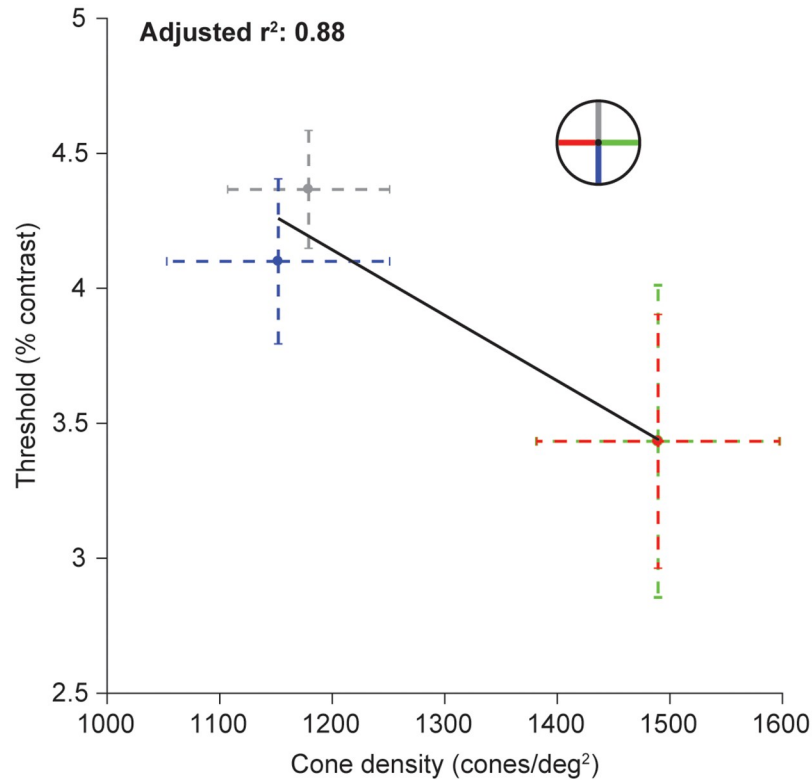


Fig 4. Correlation between performance and cone density across polar angles. Contrast thresholds (y-axis) were averaged across three observers reported in Cameron, Tai, and Carrasco [5]. Thresholds were obtained for stimuli at 4.5° eccentricity, above (gray), below (blue), left (red), and right (green) of fixation as indicated by the legend. Cone density along the four meridians are the average values reported by Song et al. [41] at 4.5° eccentricity. (For right and left, cone density was averaged for the nasal and temporal retina, since observers performed the psychophysics task binocularly). Error bars indicate one standard error across 10 observers (cone density) and 3 observers (contrast thresholds).

<https://doi.org/10.1371/journal.pcbi.1007063.g004>

cornea, fixational eye movements, the cone mosaic with different types of photoreceptors and their isomerization rate for a given stimulus presentation. Finally, the observer model has an inference engine to classify the stimulus as belonging to one of two possible classes (clockwise or counter-clockwise). Our inference engine employs a linear support vector machine (SVM) classifier to learn the two stimulus classes from the cone outputs on training data, and then applies the SVM to left out test data. Hence the classifier, unlike an ideal observer model, does not have explicit information about the stimulus, noise distributions, or the system (optics, cone types).

With a computational observer model, one can show where along the visual pathway information loss happens and how this loss of information is inherited or potentially compensated for in later encoding stages of the visual pathway. Here, we investigate to what extent visual performance (contrast threshold) depends on variations in cone density and optical quality. By systematically varying cone density and optical quality (defocus) independently, we can compare the computational observer model performance to reported differences in the literature on the human eye to quantify the individual contribution of each of these two factors to differences in visual performance across the visual field.

By studying polar angle effects, this study complements prior observer model studies of eccentricity effects; thus, this study provides novel information to our understanding of visual

perception across the visual field. Moreover, our simulations include stimulus uncertainty (phase randomization) and fixational eye movements. These factors generate trial-to-trial noise that is far more complex than simple independent Poisson noise at each receptor, making the closed-form solution of an ideal observer model more difficult to implement. Hence, we use a computational (but not ideal) observer model. This brings model performance closer to human performance than typical ideal observer models, and more generally, is an important tool for situations in which ideal observer models cannot be implemented.

Results

Overview of computational observer model

To investigate to what extent performance differences in the visual field depend on variations in cone density and optical quality, we developed a computational observer model of the first stages of the human visual pathway. The computational observer was presented with oriented Gabor stimuli, tilted either clockwise or counter-clockwise from vertical, with a phase of 90° or 270° (randomized across trials) to simulate a 2-AFC orientation discrimination task. (The phase was not relevant to the judgment). To compare the performance of the computational observer to human observers, we matched the stimulus parameters to a psychophysics study [5].

Scene radiance. The model starts from the photons emitted by a visual display, defined as the scene radiance. Three frames of two example time-varying achromatic Gabor stimuli are in panel A of Fig 5. The left column shows the result of each stage for a Gabor with high contrast (100% black outline) and middle column for a low contrast Gabor (10%, orange outline). Both of these Gabor stimuli have a counter-clock-wise orientation from vertical, contain a spatial frequency of 4 cycles/ $^\circ$ and are presented at 4.5° eccentricity. The high and low contrast Gabor stimuli contain the same mean radiance, but the high contrast Gabor has an amplitude that is 10x larger than the low contrast Gabor (1D representation at dotted line, right column of panel A).

Retinal irradiance. The second stage of the model simulates the retinal irradiance: the result of the time-varying radiance passing through the simulated human optics (including refraction and aberrations caused by the pupil, cornea, and lens). The retinal irradiance is the light image just before the photons are captured by the photopigment in the retina. Panel B of Fig 5 shows the retinal irradiance summed across all wavelengths. The effect of the optics is to blur the Gabor stimuli and to reduce the fraction of short wavelength light. The mean irradiance is the same for the two stimuli.

Isomerization. The third stage implements a cone mosaic and computes photon absorptions for each cone at each time sample (panel C of Fig 5). The cone mosaic is a rectangular patch with a field of view of $2 \times 2^\circ$ at 4.5° eccentricity. Each cone type absorbs a percentage of the emitted photons, depending on the wavelength of the light and the efficiency of the cone type.

The model implements two sources of noise. The first source comes from small fixational eye movements. These eye movements cause shifts of the stimulus on the cone mosaic during the trial. The second noise source is from photons, which are inherently noisy and follow a Poisson distribution.

During the onset of a high contrast Gabor, the L-, M-, and S-cones increase their absorptions on average by ~ 50 , ~ 30 , and ~ 4 photons/ms respectively. After stimulus offset, the absorptions return to baseline at ~ 110 , ~ 75 , and ~ 12 photons/ms. The absorption rates for a mean luminance display (~ 110 photons/ms for the L-cones) are validated by an independent computation of isomerization given the luminance by Wyszecki and Stiles [52], implemented

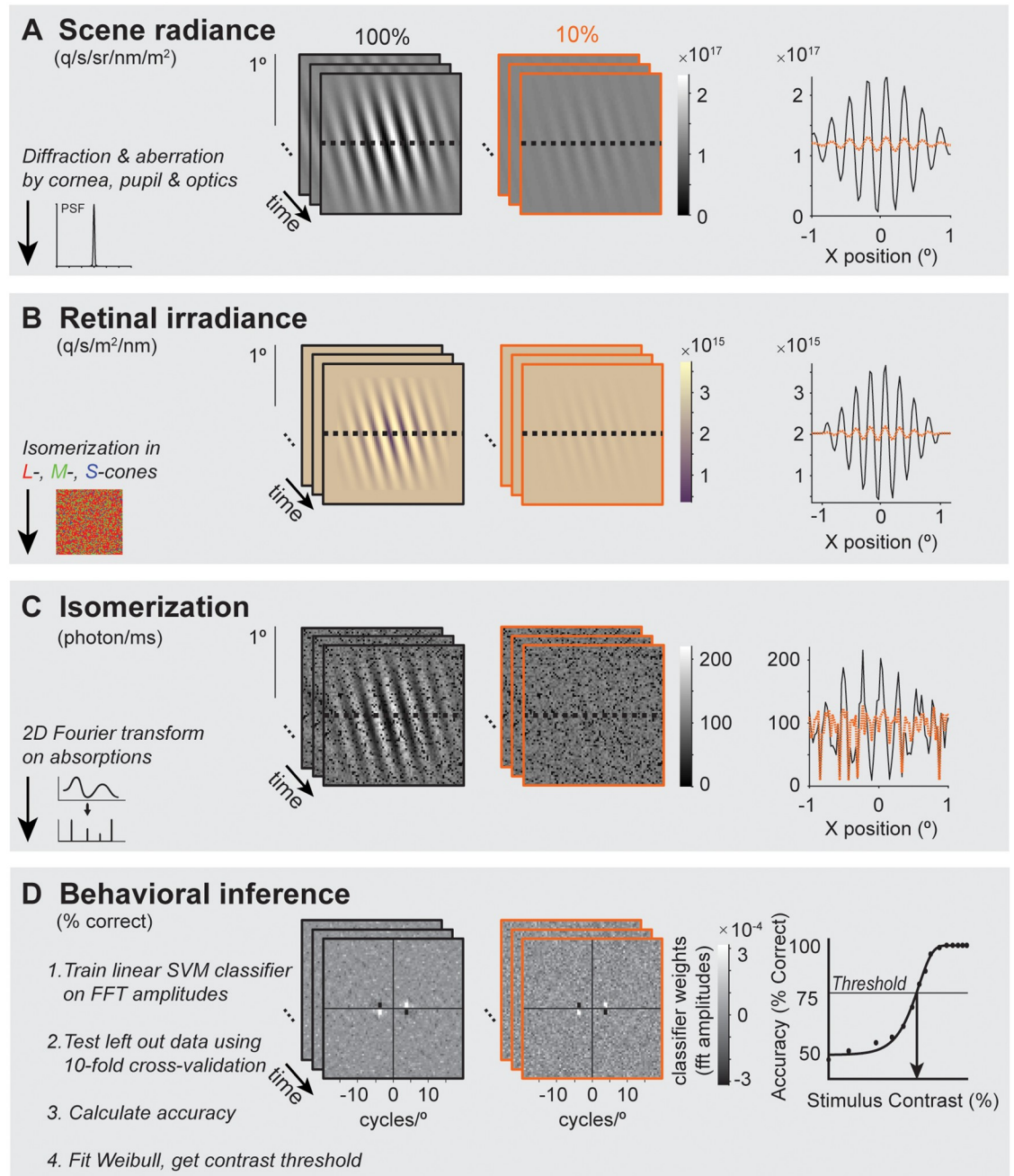


Fig 5. Overview of computations in observer model. Computations are shown for a 100% (black outline) or 10% (orange outline) contrast Gabor stimulus. **(A) Scene radiance.** The computation starts with the scene spectral radiance of a time-varying scene. The 2D images show the radiance summed across wavelengths (400–700 nm) during 1-ms windows. 1D slices through the images are plotted on the right. **(B) Retinal irradiance.** The retinal irradiance is the image that has passed through the eye's optics, before being processed by the retina. The optics are modeled as a typical human wavefront with a 3-mm pupil, indicated by the schematic point spread function (PSF) on the lower left of panel A. For illustration purposes, wavelengths of the 2D representation were converted into RGB values. The yellowing is a result of the spectral filtering by the eye's optics. **(C) Isomerization.** The time-varying retinal irradiance is transformed into photon absorptions using a rectangular grid of L-, M- and S-cones. The dark pixels correspond to S-cones (see text for more details). **(D) Behavioral inference.** The last stage of the computational observer is a simplified decision stage, performing a 2-AFC orientation discrimination task. First, the absorption image at each time point was transformed into the amplitude domain by the Fourier transform. The amplitude images were used to train a linear support vector machine (SVM) classifier using 10-fold cross-validation. The images show an example of the trained classifier weights. The weights are high at locations corresponding to the stimulus frequency (4 cycles/ $^\circ$) and orientation ($\pm 15^\circ$). Once trained, the classifier predicted the left-out data. The cross-validated accuracy was fit by a Weibull function (right part of panel D) to determine the contrast threshold.

<https://doi.org/10.1371/journal.pcbi.1007063.g005>

in the ideal observer model by Geisler (equation 2, p.776, [53]), where the average L-cone absorption under 100 cd/m^2 with a 3-mm pupil is predicted as ~ 108 photons/ms. The S-cones absorb fewer photons than the L- and M-cones. This is because inert pigments in the lens and macula absorb more light at short wavelengths, and because the photopigment density is lower in the S-cones. As a result, the locations of S-cones in the absorption array (Fig 5, panel C) are dark.

Behavioral inference. The last stage computes a 2-AFC inference of the stimulus orientation from the cone absorptions. An ideal observer would have full knowledge of the response statistics (mean and distributions of responses of each cone at each time point from each of the possible stimuli, including the higher order statistics across the cones). The only information it does not have access to is the sampled noise in the particular trial. It uses the full knowledge to make the optimal decision given a particular measurement. It is unlikely that human decision-making has access to this full knowledge.

Here we implement a computational observer that learns patterns from the data, as a human observer might do during an experiment (or in real life). Our computational observer uses a linear support vector machine (SVM) classifier. The classifier uses the weights learned from training data to classify the stimulus orientation of the left-out data.

Because the within-class stimuli differ in phase and because the eyes move during the trial, the outputs of individual cones are not informative about the decision. This would cause a linear classifier trained directly on the cone outputs to fail. Similarly, a simple template match would fail because each time the eye moves the template changes. We therefore decided to transform the cone outputs prior to training the classifier by computing the 2D Fourier transform on the cone array at each time point. We retain the amplitudes and discard the phase information. Because the Fourier transform separates the phase and amplitude for each spatial frequency and orientation, with sufficient signal to noise it is possible to infer the stimulus orientation (irrespective of phase) from the amplitude spectrum. Transforming the outputs of the cone array in this way can be thought of as giving the observer model partial information about the task: namely, that orientation and spatial frequency (but not phase) might be relevant.

As a proof of concept, the classifier shows two expected patterns. First, the largest weights of the classifier are centered on the peak spatial frequency (4 cycles/°) and orientations ($\pm 15^\circ$) of the stimuli (Fig 5, panel D). Second, the classifier accuracy increases with stimulus contrast (Fig 5, panel D, right). To summarize the data for a given simulated experiment, we computed the contrast threshold by fitting a Weibull function to the cross-validated accuracy as a function of stimulus contrast. The contrast threshold for the computational observer with typical human optics and a cone mosaic matched to $\sim 4.5^\circ$ eccentricity was 2.7%. This is slightly lower (thus better performance) than thresholds reported in the psychophysics experiment with the same stimulus parameters [5], which ranged from 3.6–9% contrast for human observers.

The effect of stimulus uncertainty and small fixational eye movements in the computational observer model

Two aspects of our simulated experiments give rise to stimulus uncertainty, small fixational eye movements and phase randomization of the stimulus (90° or 270°). The effects of these are similar. We consider eye movements first.

We model eye movements using an algorithm developed by Cottaris et al. [54], which is included in ISETBIO. This algorithm combines a statistical model of fixational drift by Mergenthaler and Engbert [55] with a model of microsaccades based on statistics reported by Martinez-Conde et al. [56, 57]. The displacement of the stimulus due to eye movements in our

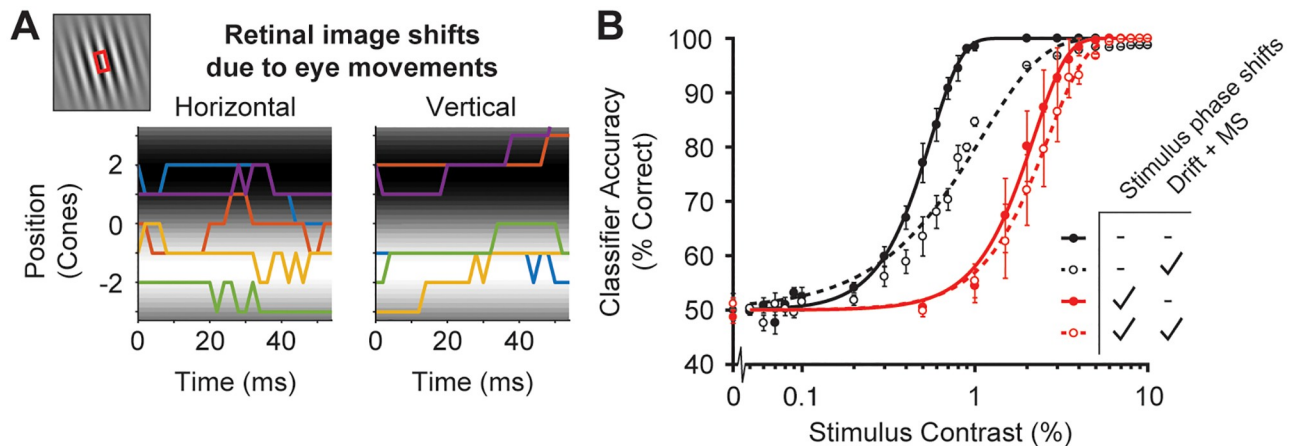


Fig 6. The effect of fixational eye movements on model performance. (A) Horizontal and vertical and 2D displacements in units of cones of the retinal image over time. Colors indicate eye movement paths of 5 example trials. A 1D slice of a single cycle of the Gabor stimulus is shown as the background (red box). (B) Computational observer performance with eye movements and stimulus phase shifts (red dashed), phase shifts only (red solid), eye movements only (black dashed) and neither (black solid). Accuracy is averaged across 5 simulated experiments which each have 400 trials per stimulus contrast. Error bars represent standard error of the mean across the 5 simulated experiments for each stimulus contrast.

<https://doi.org/10.1371/journal.pcbi.1007063.g006>

simulations is relatively small: within one trial (a single colored line), the retinal displacement tends to be about 2–4 cones or less (Fig 6A). This is small compared to the spatial scale of our stimulus, for which a full cycle corresponds to ~6 cones at 4.5° eccentricity. Given that the trials last only 54 ms, the probability of a microsaccade is low. Hence when both microsaccades and drift are present, eye movements are dominated by drift.

The fixational eye movements have a large effect on the computational observer model. Compared to a model with no phase uncertainty and no eye movements, adding fixational eye movements causes the contrast threshold to about double (Fig 6B, black solid line versus black dashed line). It might be surprising that eye movements have any effect on the model performance: An image translation is equivalent to a phase shift in the Fourier domain and the model discards phase information. However, because the retinal mosaic contains multiple cone types with different sensitivities, a shift in the stimulus causes a change in both the amplitude and phase spectra of the absorption images, affecting the information available to the classifier. Put simply, the possible set of cone responses is more variable in the presence of eye movements (whether or not the cone outputs are transformed into the Fourier domain), making it more difficult for the classifier to cleanly separate the two stimulus classes. Note that in addition to increasing the threshold, the presence of eye movements also flattens the slope of the psychometric function. This is because adding eye movements is not equivalent to lowering the contrast: even at the highest contrasts, the classifier may have some uncertainty because the eye movements can result in a set of cone responses that is very different from the average set of training responses.

The second source of stimulus uncertainty is phase randomization. On each trial, one of two phases 180° apart was randomly selected. Again, for a retina with only one cone type, the change in stimulus phase would translate to a change in the phases of the cone outputs, and would not impair a classifier operating on the Fourier amplitudes. But with a mixed cone array, the two phases result in a pattern of cone responses that differ in ways beyond a simple phases change. Like eye movements, the phase uncertainty makes performance worse (a 5x increase in threshold compared to the condition with no eye movements and no phase

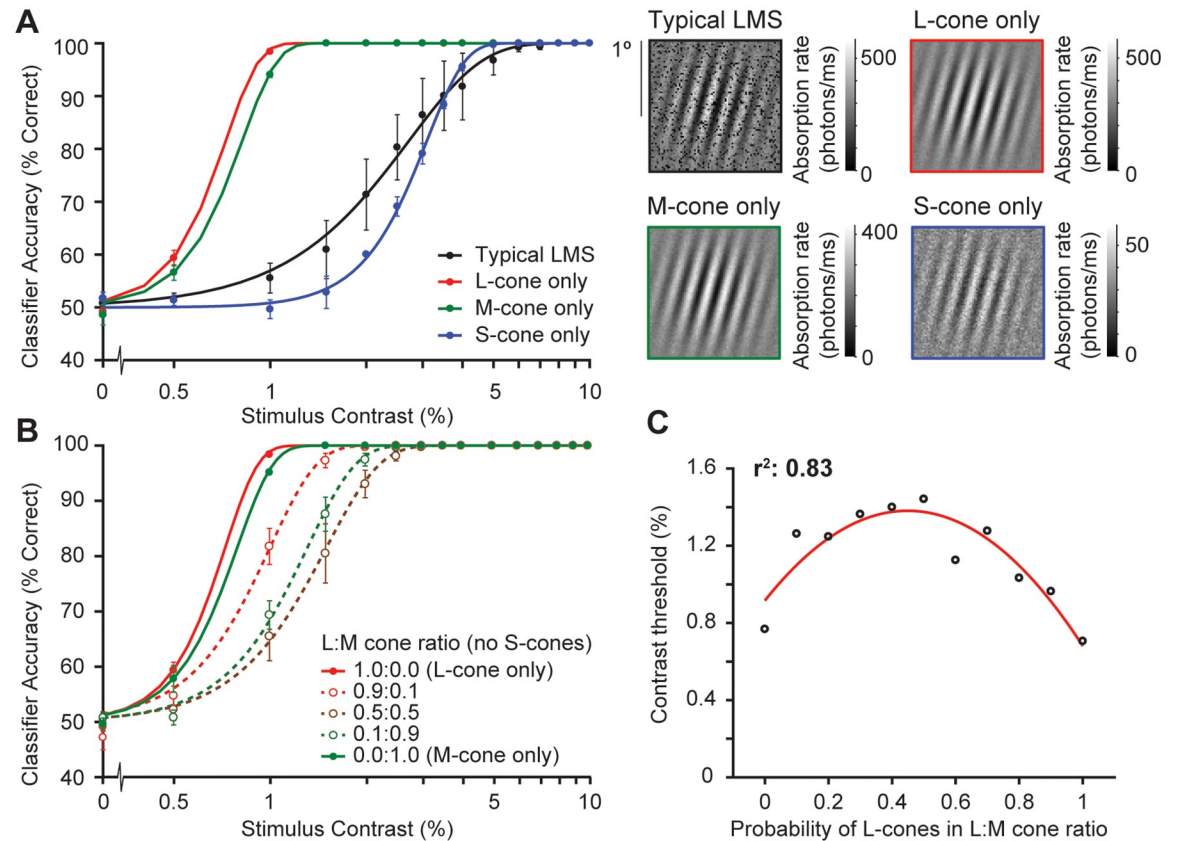


Fig 7. The effect of cone type on model performance. (A) Classifier accuracy for absorption rates as a function of stimulus contrast for the 4 different cone mosaics. Example isomerization images simulated by our computational observer model are shown on the right. A typical LMS cone mosaic has 60% L-cones, 30% M-cones, and 10% S-cones (black outline). L- and M-cone mosaics (red and green outline) have very similar absorption rates, whereas the absorption rate for S-cones is 4 times lower (blue outline). Each psychometric function is the average across performance of 2,000 trials per stimulus contrast (30,000 trials in total). Error bars represent standard error of the mean across 5 simulated experiments with 400 trials for each stimulus contrast. (B) Same layout as (A) but comparing simulated experiments with LM-only cone mosaics. We systematically varied the L:M ratio from L-cone only to M-cone only mosaics in steps of 10%. This panel shows L-cone only (solid red line) and M-cone only (solid green line); the same as in panel A. The three mixture mosaics contain 90% L-cone and 10% M-cone (dashed red line), 50% L- and M-cone (dashed brown line), and 10% L-cone and 90% M-cone (dashed green line). No S-cones were included in these mosaics. (C) Contrast thresholds as a function of probability of L-cones in L:M cone ratio, fitted with a quadratic function.

<https://doi.org/10.1371/journal.pcbi.1007063.g007>

uncertainty, solid black vs. solid red line, Fig 6B). Combining the two sources of uncertainty makes performance a little worse than either one alone (red dashed line).

The effect of cone type on orientation discrimination

Human observers drastically differ in their ratio of L-, M-cones, even at a fixed retinal location [58, 59]. Our computational observer model assumed a ratio of 0.6:0.3:0.1 L:M:S cones. In separate simulations we found that varying the ratio between cone types within a mosaic affects model performance.

To quantify the effect of cone type on our computational model performance, we varied the ratio of cone types in two ways. First, we tested the effect of cone types using uniform cone mosaics: L-cone only, M-cone only, or S-cone only, compared to a mixed cone array with a typical L:M:S ratio of 0.6:0.3:0.1 (Fig 7A).

There is a very big effect of cone type. A uniform mosaic of only S-cones has a threshold about 4 times higher than one of only L- or M- cones. The M-cone only and L-cone only mosaics result in similar performance (a 10% higher threshold for the M-cones). The large decrease in threshold for S-cone only mosaics is partly explained by chromatic aberration: our model assumes that the eye is in focus at 550 nm (in between L- and M-cone peak spectral sensitivity), causing large amounts of blur for shorter wavelengths that the S-cones are sensitive to. Blurring the image is similar to lowering stimulus contrast and therefore increases the contrast threshold. Additionally, the S-cones have lower efficiency because of the yellowing of the lens and the filtering of the macular pigment. The slopes of the psychometric functions for all three uniform mosaics are similar. This is because the changes in efficiency and in focus are similar to a change in contrast, meaning a remapping along the x-axis, or a horizontal shift in the curve.

Interestingly, the trichromatic mosaic results in a shallower psychometric function than any of the uniform cone arrays, and a threshold comparable to that found for the S-cone only mosaic. That may be surprising given that our trichromatic retina contained only 10% S-cones. Why is its performance so poor? This is because of stimulus uncertainty and eye movements: when the array has only one type of receptor, a phase difference in the stimulus or a position change caused by an eye movement translates to a simple phase difference in the absorption images, which does not impair the classifier. On the other hand, for a trichromatic retina, a phase difference or position difference in the stimulus has a more complex effect on the absorptions, resulting in performance that is worse than expected from simply computing a weighted average of the performance by the three separate cone types.

Second, to better understand the effect of arrays with cone mixtures, we simulated experiments with retinas with only L- and M-cones in various ratios. Best performance was found for uniform retinas (100% L-cones, and then almost as good, 100% M-cones) (Fig 7B). When introducing a mixture of cones in the mosaic, even a small fraction, thresholds increase. For example, a mosaic with 10% M-cones and 90% L-cones increases the threshold from 0.71% (L-cone only) to 1.0% stimulus contrast. The worst performance is found for an L:M ratio of 50:50. The computational observer model shows an approximately quadratic relation between contrast threshold and probability of L-cones in the LM mixture cone mosaic ($r^2 = 0.83$, Fig 7C). These results indicate that our model is sensitive to the variability caused by differences in mean absorption rates across cone types, even with small differences in peak spectral sensitivity and efficiency between L- and M-cones. This pattern, whereby the best performance occurs for uniform retinas, depends on there being stimulus uncertainty (phase differences in the stimulus and/or small fixational eye movements). With no uncertainty, there would be little effect of mixing L- and M-cones.

For individual differences in L:M:S cone ratios to explain the meridional effect in human performance between the horizontal and upper visual field, one would need to have a typical ratio of L:M:S on the inferior retina (upper visual field), but a retina with exclusively L-cones or M-cones along the horizontal meridian. Given that the arrangement of L-, M- and S-cones in human retina is approximately random [60], such a scenario is biologically implausible.

The effect of optical quality on orientation discrimination

Large levels of defocus worsen visual acuity [61], where defocus levels larger than 0.75 diopters (corresponding to 20/40 vision on the Snellen acuity chart for near sightedness) are usually compensated for with visual aids. Here, we tested the effect of defocus on the 2-AFC orientation discrimination task reported by Cameron, Tai, and Carrasco [5] to investigate whether variations in defocus could explain the decrease in performance with polar angle. If the task is

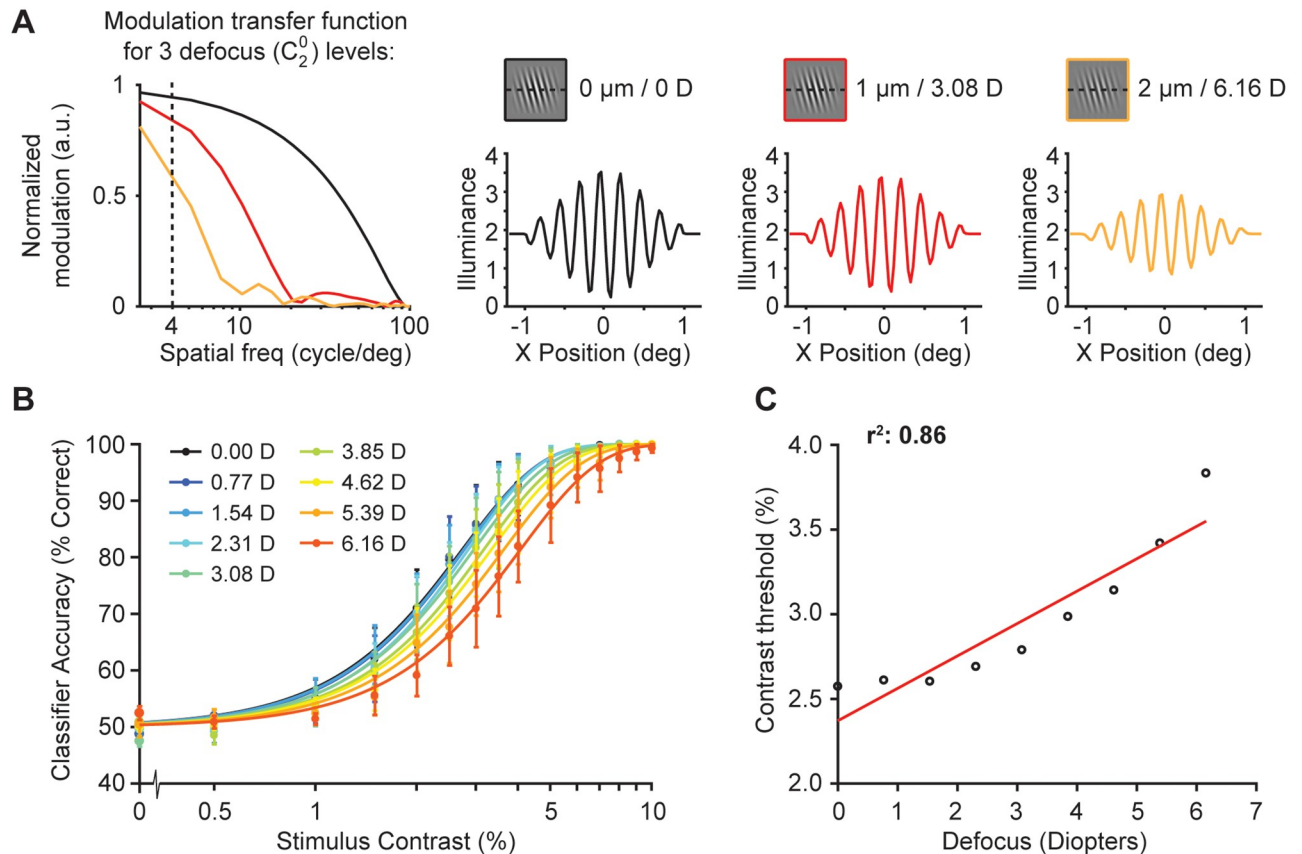


Fig 8. The effect of optical quality on model performance. (A) The modulation transfer function (MTF) of our computational observer model is shown using different levels of defocus (an addition to diffraction and higher order aberrations). The MTFs are based on a typical human wavefronts using the statistical model provided by Thibos [62] based on data from Thibos et al. [63]. The three example MTFs in black, red and yellow lines indicate a defocus Zernike coefficient of 0, 1 or 2 μm respectively (equivalent to 0, 3.08 or 6.16 diopters for a 3-mm pupil). Dotted line represents 4 cycles/°. A 1D slice through the stimulus shows the effect of the three levels of defocus. (B) Classifier accuracy for absorption rates as a function of stimulus contrast for 9 different defocus levels for a cone mosaic simulated at 4.5° eccentricity. Accuracy is averaged across performance of 2,000 trials per stimulus contrast (30,000 trials in total). Error bars represent standard error of the mean across 5 simulated experiments with 400 trials for each stimulus contrast. (C) Contrast thresholds from panel B as a function of defocus in diopters, fitted with a linear function.

<https://doi.org/10.1371/journal.pcbi.1007063.g008>

very sensitive to the level of defocus, then small differences in optical quality as a function of polar angle might explain the observed differences in performance.

Defocus affects the modulation transfer function of a typical human wavefront by attenuating the high frequencies (Fig 8A). The Gabor patches in our experiment had a peak spatial frequency of 4 cycles/° (dashed line). For this spatial frequency, the simulated levels of defocus in the observer model cause a modest reduction in contrast.

As expected, large increases in defocus cause the computational observer model to perform worse, evidenced by a rightward shift of the psychometric curve (Fig 8B). When comparing contrast thresholds as a function of defocus level, the computational observer model shows a monotonic relation with defocus, which, for simplicity, we approximate with a linear fit ($r^2 = 0.86$, Fig 8C).

The effect of defocus on model performance is small. To explain an increase of 1.5% in contrast threshold, similar to what is observed psychophysically as a function of polar angle, the computational observer model would require an additional 7 diopters of defocus. This is far higher than any plausible difference in defocus as a function of polar angle at 4.5°. Typically,

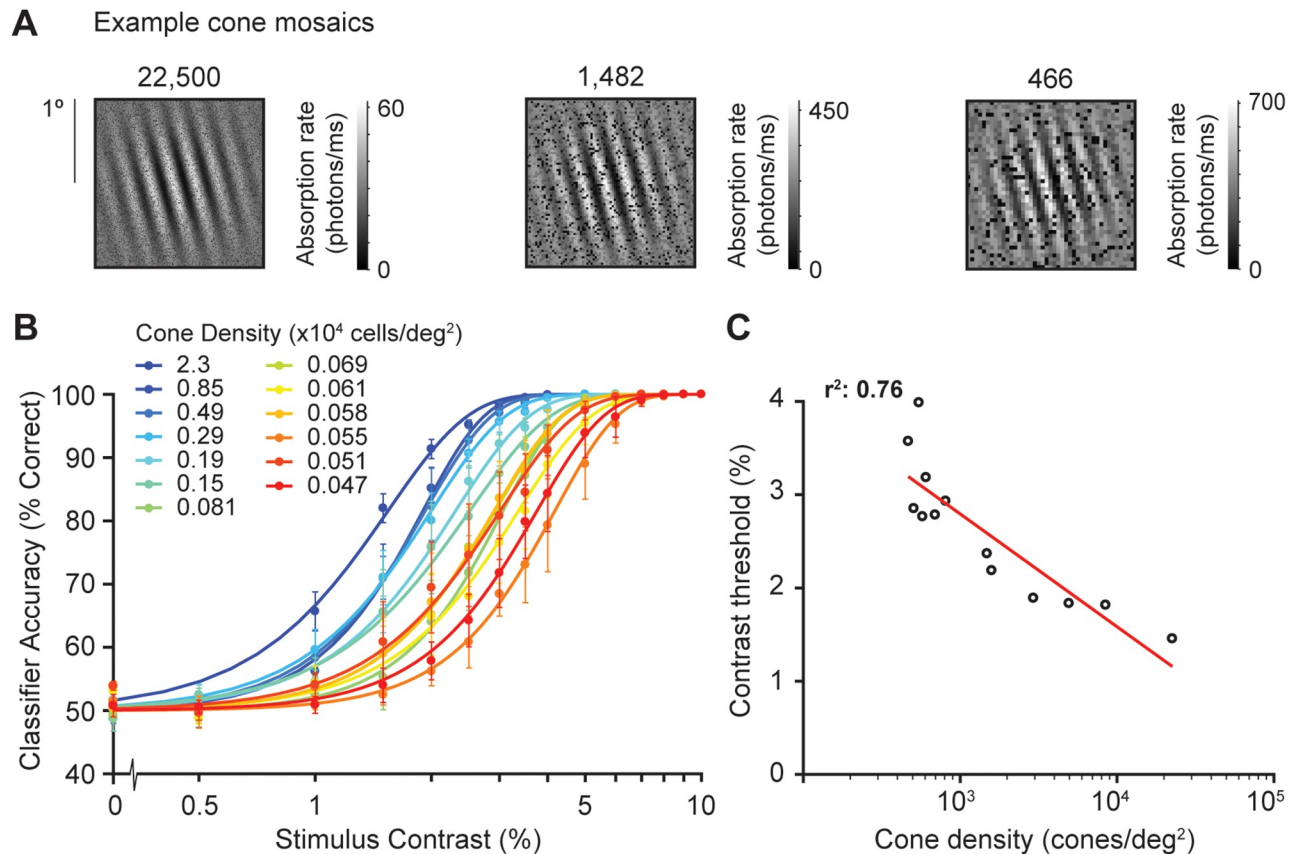


Fig 9. The effect of cone density on model performance. (A) Example cone mosaics simulated by our computational observer model. The highest tested cone density (22,500 cones/deg²) is equivalent to the cone density in the fovea (0°) whereas the lowest (466 cones/deg²) is equivalent to a cone density at 40° eccentricity, according to data from Curcio et al. [40]. In our computational observer model, lower cone density implicitly results in larger absorption rates, because the cone area increases and therefore captures more photons. (B) Classifier accuracy for absorption rates as a function of stimulus contrast for 13 different cone density levels, each psychometric function is the average across performance of 2,000 trials per stimulus contrast (30,000 trials in total). Error bars represent standard error of the mean across 5 simulated experiments with 400 trials for each stimulus contrast. (C) Contrast thresholds from panel B as a function of log cone density (cones/deg²), fitted with a log-linear function.

<https://doi.org/10.1371/journal.pcbi.1007063.g009>

defocus at 4.5° along the horizontal or vertical meridian is within ~0.2 diopters of defocus at the fovea [43, 44]. The difference between the vertical and horizontal locations at 4.5° would be even less. Assuming a difference in defocus of 0.2 diopters, the optical quality would explain only about 3% of the effect of visual performance as a function of polar angle for this task.

The effect of cone density on orientation discrimination

The cone mosaic varies substantially with retinal location. As eccentricity increases, cone diameter increases, as does spacing between the cones, resulting in lower density. We used our computational observer model to quantify the extent to which variations in the cone mosaic could explain the changes in performance with polar angle. We simulated a large range of cone densities, from about 3 times lower to 15 times greater than the typical density at 4.5° eccentricity (i.e., the eccentricity of the psychophysical experiment in [5]). As we varied the cone density, we also varied the cone size and spacing between cones according to the reported relation between density and coverage [40]. The denser mosaics sample the stimulus more finely, with fewer absorptions per cone, because as the cone area decreases it captures fewer photons. (Fig 9A).

Our computational observer model shows a decrease in contrast threshold as a function of cone density (Fig 9B and 9C). However, the effect is relatively small. For every 6-fold increase in cone density, the computational model contrast threshold reduces by 1 percentage point (e.g., from 4% to 3%). The meridional effect on human performance is ~4.4% (upper vertical meridian) vs. 3.4% (horizontal). For cone density to account for this observed difference in human contrast thresholds, there would need to be more than a 500% meridional difference in cone density. This is far greater than the 20–30% reported difference in cone density at 4.5° eccentricity [40, 41, 64]. This indicates that, according to our computational observer model, cone density accounts for less than 10% of the differences in visual performance with polar angle on the orientation discrimination task reported by Cameron et al. [5].

Discussion

An explicit model is needed to link biological measurements with psychophysical performance

Our goal was to assess the degree to which front-end properties of the visual system explain well-established psychophysical performance differences around the visual field. In particular, we quantified the contribution of two factors in the eye—cone density and optical quality (defocus)—to contrast thresholds measured at different polar angles in an orientation discrimination task as reported by [5]. These front-end factors have been reported to vary with polar angle, and in principle, the observed performance differences could be a consequence of the way the first stages of vision process images. For instance, cone density is higher on the horizontal meridian compared to the vertical meridian (up to 20° eccentricity [40, 41, 64]). Nonetheless, without a model to link these factors to performance, how much explanatory power they have cannot be assessed. We therefore developed a computational observer model to test these potential links. The underlying software we used, ISETBIO, has recently been used to model a number of basic psychophysical tasks, including contrast sensitivity [50], Vernier acuity [65], illumination discrimination [66], color perception [67], chromatic aberration [68], visual perception with retinal prosthesis [69], and spatial summation in Ricco’s area [70].

Optics and cone density can explain only a small part of performance fields

Although cone density along the cardinal meridians correlates with behavior, our model showed that this correlation has little explanatory power: Differences in cone density can only account for a small fraction of the variation in visual performance as a function of meridian. Similarly, variation in optical quality within a plausible biological range has only a very small effect on contrast thresholds in the model of our task. Our observer model puts a ceiling on these two factors at less than 10% of the observed psychophysical effects. To fully explain these visual performance differences with polar angle, our computational model would require a difference of more than 7 diopters in defocus and a difference of more than 500% in cone density for the horizontal compared to the upper vertical meridian. Such large differences are far outside the range of plausible biological variation; defocus at 4.5° eccentricity is typically within 0.1–0.2 diopters of the fovea [44] and cone density at the horizontal meridian is ~20–30% more than the vertical at this eccentricity [40].

Downstream processing contributes to performance fields

The fact that neither optics nor the cone sampling array can explain more than a small fraction of the effect of polar angle on contrast thresholds indicates that downstream mechanisms must explain the majority of this effect. This conclusion is in line with similar work using an ideal

observer model to study eccentricity-dependent effects (along the horizontal meridian) in which optics and photoreceptor properties cannot completely account for the eccentricity-dependent performance seen in human contrast sensitivity or spatial resolution [46, 48]. In particular, human performance falls off with eccentricity more sharply than would be predicted by the optics or photoreceptor properties alone. One reason for this is that downstream processing further accentuates the loss of neural resources with eccentricity. For example, the ratio of retinal ganglion cells per cone declines with eccentricity [71]. Accounting for this effect brings predicted performance closer to human in terms of the rate of decline with eccentricity, although ideal observer models that account for RGC density still perform much better than human [45, 46]. More accurate models of human detection performance are achieved by incorporating cortical computations in addition to optical and retinal factors (e.g. [47]).

Could performance differences with polar angle be explained by variation in retinal ganglion cell density? Midget retinal ganglion cells are the most prevalent class across the retina (~80% of the retinal ganglion cell population), have small cell bodies and small dendritic trees, and are hypothesized to set a limit to achromatic spatial acuity [72]. Like photoreceptors, midget retinal ganglion cells sample the visual field asymmetrically. For example, at 4.5° eccentricity, the density of midget retinal ganglion cells on the horizontal meridian is reported to be 1.4 times greater than on the vertical meridian (~1,330 vs. ~950 cells/deg² on the horizontal vs. inferior retina) [71, 73]. This 40% meridional effect is larger than the 20–30% effect at the level of the cones, indicating that polar angle asymmetries in cone density are accentuated in further retinal processing. We have not included retinal ganglion cells in our model, but given our observer model with the cone array, we speculate that this further meridional difference in ganglion cell density alone will not be sufficient to explain the reported meridional psychophysical effects.

Properties of retinal ganglion cells other than density might also vary with polar angle, such as receptive field size. If so, this too, could contribute to performance differences. For example, retinal ganglion cell receptive field size increases with eccentricity [38]. This increase, combined with random inputs of cone type to ganglion cell, has been proposed to explain the precipitous fall-off in chromatic acuity with eccentricity [74]. Little is known about retinal ganglion cell receptive field size as a function of polar angle in human. As estimated from post-mortem dendritic size, midget retinal ganglion cell receptive field sizes are smaller in the nasal quadrant than other quadrants [75, 76]. In macaque, Croner and Kaplan [77] report differences in midget cell density (e.g., nasal vs. temporal) but do not find receptive field size differences with polar angle. Without clear reports of meridional differences in midget ganglion cell receptive field and dendritic field sizes, it is unlikely that properties of retinal ganglion cells would fully account for the reported meridional effects in visual performance.

A second potential factor is visual cortex. Some aspects of performance fields manifest as amplitude differences in the BOLD fMRI signal in V1. Liu, Heeger and Carrasco [78] reported a 40% larger BOLD amplitude in V1 for stimuli on the lower than the upper vertical meridian. This asymmetry was found for high but not low spatial frequency stimuli, matching psychophysical results. They did not report differences between stimuli on the vertical versus horizontal meridians. Performance fields may also be reflected in the geometry of visual cortex. For example, a template of the V1 map fit to a population of 25 observers showed more cortical area devoted to the horizontal than the vertical meridian, although the authors acknowledged that this could be a fundamental fact about V1 or an artifact of the flattening process used in their analyses [79]. This areal difference has been confirmed in an independent data set [80]. These data also showed that population receptive fields (pRFs) in V1 and V2 are ~10% smaller when comparing horizontal to vertical quadrants. The geometry and the pRF size effects are complementary: greater area and smaller pRFs along the horizontal meridian are both

consistent with this part of visual cortex analyzing the visual field in greater detail. However, there are also psychophysical differences between the upper and the lower vertical meridian [2, 6, 7], for which no pRF differences have been reported [80]. Our ongoing work suggests that there are in fact differences in cortical magnification between the upper and lower meridian. It is not yet known whether any of these meridional effects in V1 (such as greater area for the horizontal than vertical meridian) are inherited properties from the retina, or whether V1 further amplifies the polar angle differences.

Just as visual cortex does not sample all locations in the visual field in the same manner, it also does not sample spatial frequencies and orientations perfectly uniformly (e.g. [81]). Interestingly, psychophysical asymmetries with polar angle are reported to differ as a function of spatial frequency (larger asymmetries at higher spatial frequencies), but not as a function of stimulus orientation [4–8]. However, it is unknown whether and how cortical sampling differences would account for the observed behavioral patterns. A model is needed to explicitly link non-uniform coverage across V1 to behavioral performance with regard to stimulus spatial frequency.

In addition to factors in early visual cortex, cognitive factors will also be important to consider in developing a full understanding of visual performance across polar angles. Exogenous covert visual attention does not compensate for discriminability differences across polar angles [4–6, 19], but endogenous covert attention may do so. We are currently investigating this possibility.

Limitations of the model

Our goal in building a computational observer model was to explicitly link known facts about the biology of the visual system with psychophysical performance. The value of the model is evidenced by the difference in the inference one might have drawn from a purely correlational approach (performance is best where cone density is highest) and the inference drawn from the model (little relation between cone density and performance). Nonetheless, all models are simplifications, and ours is no exception.

First, our model contained only one eye, whereas most of the psychophysical evidence in support of performance fields comes from binocular experiments. But the few studies with monocular stimulus presentation confirm differences in performance across polar angle and show a similar magnitude of the effect as for binocular stimulus presentation [4, 8]. Hence this limitation is unlikely to affect our conclusions. The effects of binocular viewing on our model performance are likely to be complicated. On the one hand, by doubling the number of photoreceptors, the signal to noise ratio would increase, consistent with the fact that contrast thresholds are lower with binocular viewing [82]. On the other hand, differences between the eyes, such as in the optics, the cone mosaic, or eye movements, could result in impaired performance. To quantify performance differences between binocular and monocular experiments, one would need to combine the information from the two eyes at some stage. In the human visual system, signals from the two eyes converge in V1 [83–85], although where exactly in V1 is still debated (e.g. [86, 87]). Because our model does not explicitly model visual processing beyond the photoreceptors, we leave the implementation of a biologically accurate binocular viewing condition for future studies.

Second, we modeled the cone mosaic as a rectangular patch with uniform density for each simulation, whereas the photoreceptors in human retina are organized in a hexagonal grid with a gradual change in density as a function of eccentricity. The uniformly spaced rectangular grid was implemented to save computational resources. The difference between an eccentricity-dependent mosaic and a uniform mosaic can be important for modeling performance

near the fovea [67], as density declines rapidly over a short distance [40]. However, further in the periphery, the density changes are modest across a small patch. And given that our model showed that very large differences in the cone array were needed to explain variation in psychophysical performance, it is unlikely that using a hexagonal, eccentricity-dependent array would have altered our conclusions.

Third, we did not model differences in photopigment density or macular pigment density as a function of retinal position. Pigment density has an effect on wavelength sensitivity and overall efficiency [52]. Although our model did not vary pigment density, it did include position-dependent efficiency, implemented by varying the cone coverage, which ranged from close to 1 (no gaps between cones) near the fovea to ~0.25 in the far periphery. Hence, additional variation in efficiency arising from pigment density would be unlikely to have a substantial impact on model performance. Moreover, macular pigment density does not vary systematically with polar angle at iso-eccentric locations [88].

Finally, our computational model only deals with visual processes up to photon absorptions by the cones. Processes up to this point, optics, photon noise, and cone sampling, are well characterized and can be accurately modeled. In future work, we will build on our computational observer model to investigate the contribution of downstream factors, such as post-receptor retinal circuitry and pooling of signals by retinal ganglion cells and visual cortex.

The inference engine

The performance of a classifier depends, in part, on how much knowledge of the task the classifier has access to. Our observer model had far less information than an ideal observer model. By definition, ideal observer models have complete knowledge about the relationship between inputs and outputs (except for the trial-to-trial stochastic noise) and use this knowledge to make optimal decisions, thereby setting an upper limit on performance [28, 29, 46, 53, 89–91]. When ideal observer models are applied to very early signals in the visual system such as cone responses, they typically outperform human observers by a large margin, e.g., by a factor of 10 to 100 [46, 92]. In contrast, our computational observer model performs similar to human observers (~2–4% contrast thresholds in our task). The similarity to human performance should be interpreted with caution, however, since small variations in our model, such as the number of training trials, affect the performance of the model.

Our inference engine has two types of knowledge about the task, one more general about visual processing and one more specific to the particular experiment we simulated. The general (and implicit) knowledge arises from transforming the 2D time-varying cone absorption images to amplitude spectra. Transforming the data in this way preserves most of its representation and it effectively gives the observer model knowledge that spatial frequency and orientation (but not phase) might be relevant for the task. This transform does not indicate which spatial frequencies or orientations are relevant. Although the visual system does not literally compute a Fourier transform of the cone responses, cells in visual cortex are tuned to orientation and spatial frequency in local patches of the image [93, 94], and complex cells in V1 are relatively insensitive to phase [84, 95]. Hence the implicit knowledge we provide to the classifier via transform to the amplitude spectra is an approximation that is conceptually inspired by general processing strategies in the visual system, but is not a specific implementation of a V1 stage with complex cells or knowledge about our particular task. Pilot simulations in which the classifier operated directly on the absorption images resulted in near-chance performance. This is expected, because the phase randomization of the stimuli causes the number of absorptions for any particular cone to be uninformative as to the stimulus orientation.

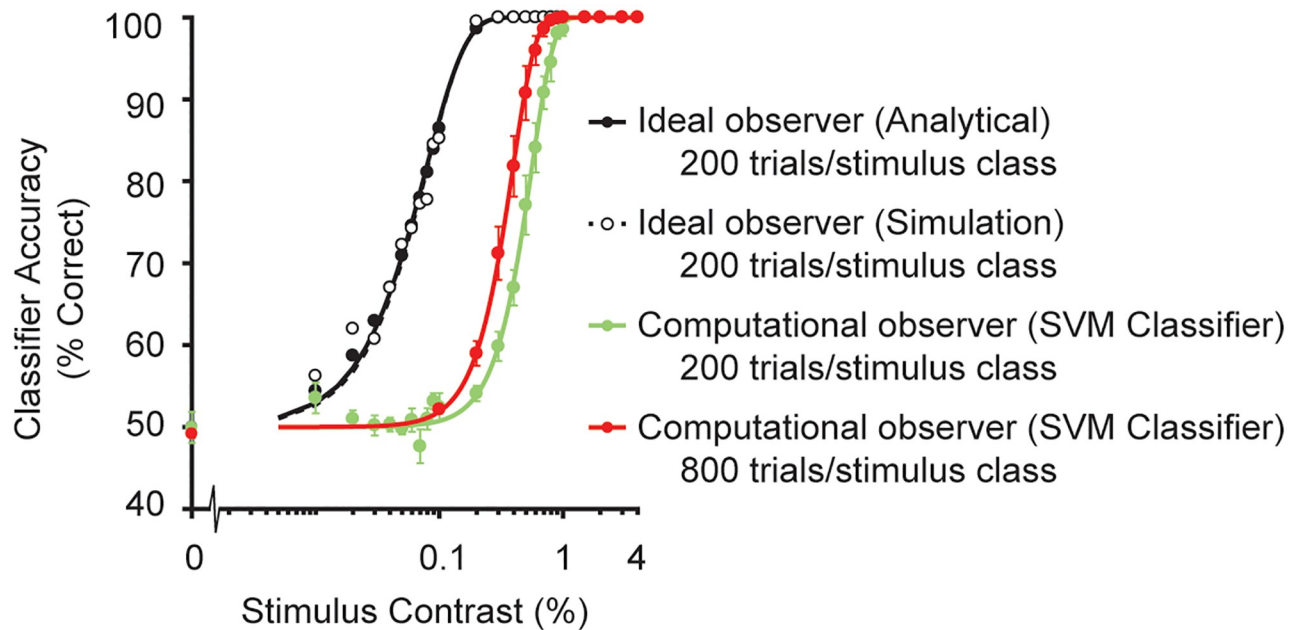


Fig 10. Ideal versus computational observer model performance. Comparison of model performance for simulations in which there were only one stimulus phase and no eye movements, presented at 4.5° eccentricity and typical human optics without defocus. Ideal observer was implemented using the derivation and formula from Geisler [53]. The analytical computation (filled black circles) is the closed form solution given the two (noiseless) templates, assuming the only source of noise is photon noise (Poisson). The ideal observer by simulation (unfilled black circles) applies the ideal decision for 200 trials per stimulus class, where each trial has no uncertainty other than photon noise. The two ideal observer calculations will converge to the same values with sufficiently large simulations. Computational observer performance is replotted from Fig 6B, where a linear SVM classifier learns the mean and variance of cone absorptions for the two stimulus classes for 200 trials per stimulus class (green line). The same computational observer model performs slightly better with more trials (800 per stimulus class, red line).

<https://doi.org/10.1371/journal.pcbi.1007063.g010>

More specific knowledge in the computational observer model comes from the training trials, which are used to learn the best linear separation (hyperplane) between the two stimulus classes. The plane is defined by a weighted sum of the classifier inputs (amplitude spectra in our case), which can be thought of as an approximation to receptive field analysis by downstream neurons. The high weights learned by the classifier for this task correspond to oriented, band-pass filters, which match properties of the stimuli (Fig 5, panel D). Because the model has incomplete knowledge, values far from the stimulus (very high or very low spatial frequency, and orientations far from the stimulus orientations) have non-zero weights, which are learned during training on a finite number of noisy trials.

Our model differs from ideal observer models. To understand how the differences affect performance, we implemented an ideal observer model as described by Geisler [53]. An ideal observer model for simulations with multiple cone types and fixational eye movements is extremely complex, and hence for this comparison, we examined the case of no eye movements and only a single stimulus phase for each of the two stimulus classes. The ideal observer performs far better than our computational observer, with a threshold about 10x lower (Fig 10, black vs. green line). This difference is caused by the fact that the linear SVM classifier needs to learn the stimulus classes and the noise distributions, rather than starting with such knowledge. If we increase the number of trials fourfold, the SVM computational observer performance improves a little (red vs. green line in Fig 10), because it has more samples to estimate learn the classes. As shown by Cottaris et al. [50] for a related task, the SVM may require thousands to millions of trials to reach ideal performance levels. With the additional uncertainty from multiple stimulus phases and fixational eye movements, the computational observer

model performs even worse, as shown in Fig 6. It is this performance—the SVM classifier trained with 200 trials operating on simulations that include fixational eye movements and two stimulus phases—that is similar to human performance. The type of computational observer model implemented here is useful when the ideal observer model is unwieldy or intractable and has the benefit of being potentially more similar to how the human learns the task.

Conclusion

Overall, our model includes a relatively detailed, biologically plausible front-end, which incorporates realistic details about the optics, photon noise, small fixational eye movements, and wavelength- and position-sampling by photoreceptors. This front-end processing was combined with a linear classifier that performs at levels comparable to the human without providing explicit knowledge about the tasks. Future work will incorporate more biologically explicit models of downstream processing, including retinal and cortical circuitry. Such models are likely to reveal that later processing in the nervous system inherits, and possibly amplifies, asymmetries in processing around the visual field that begin in the earliest stages of vision, and thus, to explain a larger portion of the psychophysical asymmetries found in many visual tasks.

Methods

Computational observer model software overview

The computational observer model relies on the publicly available, MATLAB-based Image Systems Engineering Toolbox for Biology (ISETBIO [49–51]), available at <http://isetbio.org/>. The ISETBIO toolbox incorporates the image formation process, wavelength-dependent filtering, optical quality, and the spatial arrangement and biophysical properties of cones. We used the ISETBIO toolbox for the core model architecture and supplemented it with experiment-specific custom MATLAB code. The experiment-specific code implements stimulus parameters matched to a prior psychophysical study [5], manipulation of biological parameters to assess their impact on performance, and a 2-AFC linear support vector machine classifier. In the interest of reproducible computational methods, the experiment-specific code, for both simulation and analysis, is publicly available via GitHub (<http://github.com/isetbio/JWLOrientedGabor>). In addition, the data structures created by the simulation and analyses are permanently archived on the Open Science Framework URL: <https://osf.io/mygvu/>.

Psychophysical experiment

Our simulations were created to match a previous psychophysical study [5]. In that study, stimuli were achromatic oriented Gabor patches. The Gabors were comprised of harmonics of 4 cycles/°, windowed by a Gaussian with a standard deviation of 0.5°, presented at 4.5° eccentricity, at one of 8 locations equally spaced around the visual field (see also Fig 1A). Gabor patches were tilted either 15° clockwise or counter-clockwise from vertical, and presented for 54 ms on each trial. The contrast of the Gabor patches varied from trial to trial. The contrast levels were selected for each observer based on pre-experiment testing, and usually ranged from about 1% to 10% Michelson contrast using a method of constant stimuli. The observer's task was to indicate the orientation of the Gabor stimulus relative to vertical (clockwise or counter-clockwise) with a button press. Data were analyzed by fitting a Weibull function to the mean performance (% correct) at each contrast level, independently at different locations around the visual field.

Stimuli (scene spectral radiance)

The observer model starts with a description of the stimulus, called a ‘scene’ in ISETBIO. The scene is defined by the spectral radiance at each location in space and time (the ‘light field’). The spectral radiance contained wavelengths ranging from 400–700 nm, discretized to 10 nm steps, with equal photons at each wavelength (3.8×10^{15} quanta/s/sr/nm/m²). The stimulus was discretized into 2-ms time steps and 1.8-arcminute spatial steps (32 samples per degree). The scene comprised Gabor stimuli with parameters described above (Methods section ‘Psychophysical experiment’), oriented either clockwise or counter-clockwise, represented within a field of view of 2° diameter, and presented for 54 ms per trial. The dimensions of the scene were therefore 64 x 64 x 31 x 28 (height x width x wavelength x time). Gabor patches varied in Michelson contrast between 0.05% and 10% (or 0.01% and 10% for simulations with one stimulus phase and no fixational eye movements). We also incorporated a stimulus with 0% contrast stimulus as a sanity check whether our model would perform at chance level. For all stimuli, the mean luminance was 100 cd/m². Because photon noise and eye movement noise are added later (see Methods sections ‘Optics’ and ‘Cone mosaic’), and because we do not model the scene before or after the stimulus onset/offset, the scene is in fact identical at all 28 time points.

Machine learning algorithms can exploit sources of information that a human observer would be unlikely to use. For example, if the value of a single image pixel happened to correlate with the stimulus class, a classifier could succeed based on only the value of this pixel. We wanted to prevent our classifier from succeeding in this way. In our simulations (unlike the Cameron et al. paper [5]), the phase of the Gabor patches was selected from two values 180° apart ($\varphi = 90^\circ$ and $\varphi = 270^\circ$), randomized across trials. A 180° phase difference means that the two possible stimuli within a class were identical except for a sign reversal. As a result, the expected value of each pixel in each stimulus class was 0 (relative to the background). Similarly, the expected value of the cone absorption rates at each location on the retina within a stimulus class was 0 (relative to the background). Therefore, the linear classifier could not succeed using the absorption level from any single cone. We believe human observers do not perform the task this way either, hence randomizing the phase is likely to make the observer performance more similar to the human performance.

While most of the simulations contained two possible stimulus phases and fixational eye movements, we also explicitly manipulated these factors in two simulations. In the Results section describing the effect of stimulus uncertainty and small fixational eye movements (Fig 6B), we compare the effect of one vs. two stimulus phases and the presence vs. absence of small fixational eye movements on model performance. In the Discussion section ‘The inference engine’ (Fig 10), we compare ideal observer model performance to our computational observer model performance for a simulated experiment with only one stimulus phase and no eye movements.

Optics (retinal irradiance)

The optics transform the scene into a retinal image. We first describe the optics used for the simulations in Results sections on the effect of stimulus uncertainty and eye movements, cone type and cone density on model performance (Figs 6, 7 and 9). For these simulations, the optics were matched to a typical human eye with a 3-mm pupil (diameter) in focus at 550 nm using a statistical model of wavefront aberrations [62]. This statistical model is based on measurements from healthy eyes of 100 observers [63], and described by a basis set of Zernike polynomials [96]. The statistical model by Thibos contained the first 15 Zernike coefficients (Z0–Z14, using OSA standard indexing). The simulated human wavefront was used to construct a point spread function (PSF). This PSF was convolved with the scene at every time

point to generate the retinal image. After this spatial blurring, the optical image was further transformed by spectral filtering (light absorption by inert pigments in the lens and macula), which primarily reduce the intensity of short-wavelength light. Finally, the optical images were padded by 0.25° on each side with the mean intensity at each wavelength. The padding is needed to handle eye movements, so that cones near the edge of the simulated retinal patch have a defined input even when these cones are moved outside the scene boundaries. The dimensions of the optical image are the same as the dimensions of the scene, except for the spatial padding: $80 \times 80 \times 31 \times 28$ (height \times width \times wavelength \times time), which was discretized the same way as the scene.

To investigate the effect of optical quality on visual performance of our task, we systematically added further defocus to the model of human optics (Fig 8). We did this by increasing the Z4 Zernike coefficient (defocus) from $0\text{--}2 \mu\text{m}$ in steps of $0.25 \mu\text{m}$ (corresponding to $0\text{--}6.16$ diopters for a 3-mm pupil), while keeping all other Zernike coefficients from Thibos' statistical model unchanged. Note that using a defocus coefficient of 0 does not result in perfect diffraction limited optics, given that the other aberrations are still non-zero. We manipulated defocus rather than all the higher-order aberrations because at the stimulus eccentricity we simulated (4.5°), defocus is the largest contributor to optical quality [44].

Cone mosaic: Spatial sampling and isomerization

We constructed the cone mosaic as a uniformly spaced rectangular patch with a field of view matched to the stimulus ($2 \times 2^\circ$). Each cone mosaic contained a random distribution of L-, M- and S- cones with a ratio of 0.6:0.3:0.1. We used the Stockman-Sharpe [97] cone fundamentals to estimate cone photopigment spectral sensitivity, assuming 50% optical density for L- and M-cones, and 40% for S-cones. Peak efficiency was assumed to be equal for each cone class, 66.67% multiplied by the retinal coverage (the fraction of local retina occupied by cones).

For the simulations on the effect of stimulus uncertainty and eye movements, cone type and optical quality on model performance (Figs 6–8), the cone density was $1,560 \text{ cells/deg}^2$, approximately matched to the density at 4.5° on the horizontal retina as reported by Curcio et al. [40]. This results in an array of 79×79 cones for our 2° patch. The positions of the L-, M-, and S-cones were randomized within the array (but held to fixed ratio). For these simulations, we assumed a coverage proportion of 0.49, meaning that the cone inner segments sampled from about half of the optical image, and missed about half due to the spaces between cones. A coverage of less than 1 acts like a reduction in efficiency, since photons are lost to the gaps between cones. In general, cone coverage decreases with eccentricity as the density of rods increases, filling the spaces between cones.

For the Results section on the effect of cone types (Fig 7), we quantified model performance when the L:M:S cone ratios varied. First, we simulated experiments with mosaics containing a single cone type, i.e. only L-cones, only M-cones, or only S-cones. Second, we systematically varied the L:M ratio in a cone mosaic without S-cones. We quantified model performance for 11 different mosaics, varying from 100% L-cones to 100% M-cones discretized in steps of 10%.

For the set of experiments investigating cone density (Fig 9), we systematically varied cone density spanning $22,500$ to 466 cones/deg^2 (corresponding to cone arrays ranging from 297×297 to 43×43). For each cone density, we determined an equivalent eccentricity based on the relation between eccentricity and density on the nasal meridian from Curcio et al. [40]. We then adjusted the cone coverage according to this eccentricity, assuming that coverage declines exponentially as a function of eccentricity, from 1 (fovea, no gaps between cones) to 0.25 at

40°. This approximation is similar to that used by Banks et al. [46], which was based on data from Curcio et al. [40].

The number of absorptions was computed for each cone in two steps. First, the noiseless number of absorptions was computed by multiplying the appropriate cone sensitivity function (L-, M-, or S-cone) by the corresponding location in the optical image (hyperspectral), and scaling this value by the peak efficiency (66.67%). The cone coverage was accounted for by only sampling the optical image at the locations within the cone inner segments. Second, the noiseless values were converted to noisy samples by assuming a Poisson distribution.

The dimensions of the cone array absorptions were 79 x 79 x 28 (rows x columns x time) for the simulations in the Results section on the effect of stimulus uncertainty and eye movements, cone type and optical quality (Figs 6–8). When the cone density varied (Fig 9), the first two dimensions of the cone array size also changed.

Eye movements

We added small fixational eye movements (drift and microsaccades), before computing the isomerization rate for each cone at each time sample. The ISETBIO toolbox provides an algorithm developed by Cottaris et al. [54] that generates eye movement samples based on Mergenthaler and Engbert's drift model [55] and microsaccade statistics reported by Martinez-Conde et al. [56, 57].

The drift model computes eye movement paths for a single trial with modified Brownian motion process. The eye movement paths were generated in units of arc minutes and then converted to discrete cone shifts in the horizontal and vertical direction. If the amplitude of an eye movement was smaller than the distance between two cones, the displacement was accumulated over multiple time samples, until the threshold was reached, before a new shift was added to the eye movement path.

The drift model was implemented by adding a displacement vector to the current position at each time point. The displacement vector was determined by combining 3 inputs: 2D Gaussian noise, an autoregressive term for persistent dynamics at short time scales, and a delayed negative feedback for antipersistent dynamics at longer time scales. The parameters we used for this model were the ISETBIO defaults, which contained a horizontal and vertical delay defined as $X = 0.07$ s and $Y = 0.04$ s, feedback steepness of 1.1, and feedback gain of 0.15. The control function had a mean of 0 and standard deviation of 0.075 and the gamma parameter was set to 0.25. The mean noise position and standard deviation were set to 0 and 0.35, respectively. Before computing the velocity of a drift period, the drift model applied a temporal smoothing filter to the eye movement paths using a 3rd order Savitzky-Golay filter over a velocity interval of 41 ms.

For periods where the drift was stabilized, the eye movement code checked for microsaccade jumps to add to the eye movement path. Whether or not a microsaccade was added depended on when the last microsaccade was. In our experiment, we used the ISETBIO default where the interval between microsaccades followed a gamma function with a mean of 450 ms, with a minimum duration of 2 ms. A microsaccade was defined as a vector where the mean amplitude of a microsaccade was 8 arc minutes. Each vector contained an additional endpoint jitter of 0.3 arc min in length and 15° in direction. The microsaccade jumps were either 'corrective' (towards the center of the mosaic) or 'random' (any direction). The microsaccade mean speed was defined as 39°/s, with a standard deviation of 2°/s. Given that the defined interval between microsaccades was long compared to the stimulus duration (54 ms), most trials did not contain microsaccades.

A 216 ms warmup period was implemented before the trials began. Eye movements during this period affected the eye position at the start of the trial but were not otherwise included in the analysis.

Simulated experiments and behavioral inference

A simulated experiment comprised 6,000 trials, with 400 trials at each contrast. Most simulated experiments contained 15 contrast levels within the range of 0–10%. More contrast levels were used in a few simulations in which the thresholds were very different between conditions (29 contrast levels). In all experiments there were 400 trials per contrast level. These 400 trials included 200 clockwise and 200 counter-clockwise stimuli, each of which was further subdivided into 100 trials at each of 2 phases. The data from a single contrast level within a single experiment were represented as a 4D array (m rows \times m columns \times 28 time-points \times 400 trials), in which m is the number of cones along one side of the retinal patch (79 in the experiments for Figs 6–8, variable for the experiments in Fig 9).

Within the 6,000 trials of an experiment, all parameters other than the stimulus orientation (clockwise or counter-clockwise) and phase (90° or 270°) were held constant, including the spatial distribution of L-, M-, and S-cones, the optics, the cone density, the cone coverage, and the presence or absence of fixational eye movements. Each simulated experiment was repeated 5 times, so that a single psychometric function summarized 2,000 trials per contrast level. The arrangement of L-, M-, and S-cones was regenerated randomly for each of the 5 repeated experiments. Error bars in Figs 6–9 indicate standard errors of the mean across the 5 experiments. Some stimulus contrasts' error bars were very small (usually for high stimulus contrasts), which resulted in error bars being masked by the data point.

Classification (clockwise vs. counter-clockwise) via cross-validation was performed separately for each stimulus contrast level (set of 400 trials) in each experiment as follows. First, each $m \times m$ image of cone absorptions was transformed into an $m \times m$ amplitude spectrum using the 2D fast Fourier transform and discarding the phase information. This left the dimensionality unchanged (m rows \times m columns \times 28 time-points \times 400 trials within a single contrast level). Second, the amplitudes were concatenated across space and time into a 2D matrix (400 trials \times $28m^2$ values per trial). A 400-element vector labeled the trials by stimulus class (1 for clockwise and -1 for counter-clockwise). Third, the 2D matrix and 400-element vector with labels was used for training and testing a linear support vector machine (SVM) classifier on the amplitude images using MATLAB's *fitsvm* with 10-fold cross-validation, kernel function set to 'linear', and the built-in standardization option (to z-score each row of the data matrix). The learned classifier weights represented the best linear separation (hyperplane) between the two stimulus classes. With these trained weights, the classifier predicted the stimulus class label for the left-out trials in a given data fold. We used MATLAB's *kfoldLoss* function to average the accuracy across the 10-folds, which yielded one accuracy measure (% correct) per contrast level per experiment.

The implementation of an ideal observer model was based on Geisler [53]. First, we computed the ideal observer outputs based on Geisler's closed-form solution in equation 3 (p. 777). This solution computed the ideal observer discriminability between two stimuli at the level of the cones given only the expected number of absorptions per cone per stimulus. To compute these values, we used a noiseless version of our computational observer in ISETBIO (no photon noise, typical human optics without defocus, no fixational eye movements, single stimulus phase, and a cone mosaic at 4.5° eccentricity). The d' -prime for each contrast level was converted to percent correct assuming an unbiased criterion. Second, we also computed the ideal observer percent correct using the sampled (noisy) absorptions, rather than the closed

form solution (equation 5, p. 781). As expected, with a reasonably large number of trials, the two implementations converge.

Quantifying the contribution of cone density and optics on the computational observer performance

To quantify the contribution of a given factor in the eye, we averaged the classifier accuracy for each stimulus contrast level across the 5 experiments to fit with a Weibull function (Eq 1). This resulted in full psychometric functions for each cone density, cone type, or optical quality level. We calculated error bars for each contrast level as the standard error of the mean across the 5 iterations. For each psychometric function, we defined the contrast threshold as the power of 1 over the slope of the Weibull function β , in our case $\beta = 3$, of the performance level expected at chance (0.5 for a 2-AFC task). This results in a contrast threshold taken at $\sim 80\%$ ($0.5^{1/3} = 0.7937$, defined as α in Eq 2).

$$y = 1 - (1 - g) \cdot e^{-\left(\frac{kt}{\alpha}\right)^\beta} \quad (1)$$

Where g is the performance expected at chance (0.5), t is the threshold, β is the slope of the Weibull function, and k is defined as:

$$k = \left(-\log\left(\frac{1-g}{1-g}\right)\right)^{\frac{1}{\beta}} \quad (2)$$

The contrast thresholds, t , were summarized as a quadratic function as a function of L-cone probability (Fig 7), as linear function of defocus (Fig 8), or an exponential function of cone density (Fig 9, represented as a straight-line on a semi-log axis). The square of Pearson correlation coefficient was used to report proportion variance explained (r^2) by the fit. These quadratic, linear, or log-linear fits enabled us to compute the change in cone density or the change in defocus needed to achieve a 1% increase in contrast threshold—similar to the meridional effect observed in human performance ($\sim 4.4\%$ at the upper vertical meridian vs. $\sim 3.4\%$ at the horizontal meridian as seen in [5]).

Acknowledgments

We would like to thank Brian Wandell, David Brainard and Nicolas Cottaris for their advice and encouragement. We thank Pablo Artal for kindly providing data on human optical quality.

Author Contributions

Conceptualization: Eline R. Kupers, Marisa Carrasco, Jonathan Winawer.

Formal analysis: Eline R. Kupers, Jonathan Winawer.

Funding acquisition: Marisa Carrasco, Jonathan Winawer.

Investigation: Eline R. Kupers, Jonathan Winawer.

Supervision: Marisa Carrasco, Jonathan Winawer.

Visualization: Eline R. Kupers.

Writing – original draft: Eline R. Kupers, Jonathan Winawer.

Writing – review & editing: Eline R. Kupers, Marisa Carrasco, Jonathan Winawer.

References

1. Rovamo J, Virsu V, Nasanen R. Cortical magnification factor predicts the photopic contrast sensitivity of peripheral vision. *Nature*. 1978; 271(5640):54–6. PMID: [625324](#).
2. Abrams J, Nizam A, Carrasco M. Isoeccentric locations are not equivalent: the extent of the vertical meridian asymmetry. *Vision Res*. 2012; 52(1):70–8. <https://doi.org/10.1016/j.visres.2011.10.016> PMID: [22086075](#)
3. Strasburger H, Rentschler I, Jüttner M. Peripheral vision and pattern recognition: a review. *J Vis*. 2011; 11(5):13. <https://doi.org/10.1167/11.5.13> PMID: [22207654](#).
4. Carrasco M, Talgar CP, Cameron EL. Characterizing visual performance fields: effects of transient covert attention, spatial frequency, eccentricity, task and set size. *Spat Vis*. 2001; 15(1):61–75. PMID: [11893125](#)
5. Cameron EL, Tai JC, Carrasco M. Covert attention affects the psychometric function of contrast sensitivity. *Vision Res*. 2002; 42(8):949–67. [https://doi.org/10.1016/S0042-6989\(02\)00039-1](https://doi.org/10.1016/S0042-6989(02)00039-1) PMID: [11934448](#).
6. Talgar CP, Carrasco M. Vertical meridian asymmetry in spatial resolution: visual and attentional factors. *Psychon Bull Rev*. 2002; 9(4):714–22. PMID: [12613674](#).
7. Fuller S, Rodriguez RZ, Carrasco M. Apparent contrast differs across the vertical meridian: visual and attentional factors. *J Vis*. 2008; 8(1):16. <https://doi.org/10.1167/8.1.16> PMID: [18318619](#)
8. Corbett JE, Carrasco M. Visual performance fields: frames of reference. *PLoS One*. 2011; 6(9):e24470. <https://doi.org/10.1371/journal.pone.0024470> PMID: [21931727](#)
9. Rovamo J, Virsu V. An estimation and application of the human cortical magnification factor. *Exp Brain Res*. 1979; 37(3):495–510. <https://doi.org/10.1007/BF00236819> PMID: [520439](#).
10. Virsu V, Rovamo J. Visual resolution, contrast sensitivity, and the cortical magnification factor. *Exp Brain Res*. 1979; 37(3):475–94. PMID: [520438](#).
11. Rijdsdijk JP, Kroon JN, van der Wildt GJ. Contrast sensitivity as a function of position on the retina. *Vision Res*. 1980; 20(3):235–41. [https://doi.org/10.1016/0042-6989\(80\)90108-X](https://doi.org/10.1016/0042-6989(80)90108-X) PMID: [7385597](#).
12. Robson JG, Graham N. Probability summation and regional variation in contrast sensitivity across the visual field. *Vision Res*. 1981; 21(3):409–18. [https://doi.org/10.1016/0042-6989\(81\)90169-3](https://doi.org/10.1016/0042-6989(81)90169-3) PMID: [7269319](#).
13. Lundh BL, Lennerstrand G, Derfeldt G. Central and peripheral normal contrast sensitivity for static and dynamic sinusoidal gratings. *Acta Ophthalmol (Copenh)*. 1983; 61(2):171–82. PMID: [6880630](#).
14. Regan D, Beverley KI. Visual fields for frontal plane motion and for changing size. *Vision Res*. 1983; 23(7):673–6. [https://doi.org/10.1016/0042-6989\(83\)90208-0](https://doi.org/10.1016/0042-6989(83)90208-0) PMID: [6613009](#).
15. Skrandies W. Human contrast sensitivity: regional retinal differences. *Hum Neurobiol*. 1985; 4(2):97–9. PMID: [4030428](#).
16. Anderson RS, Wilkinson MO, Thibos LN. Psychophysical localization of the human visual streak. *Optom Vis Sci*. 1992; 69(3):171–4. PMID: [1565412](#).
17. Mackeben M. Sustained focal attention and peripheral letter recognition. *Spat Vis*. 1999; 12(1):51–72. PMID: [10195388](#).
18. Altpeter E, Mackeben M, Trauzettel-Klosinski S. The importance of sustained attention for patients with maculopathies. *Vision Res*. 2000; 40(10–12):1539–47. [https://doi.org/10.1016/S0042-6989\(00\)00059-6](https://doi.org/10.1016/S0042-6989(00)00059-6) PMID: [10788657](#).
19. Carrasco M, Williams PE, Yeshurun Y. Covert attention increases spatial resolution with or without masks: support for signal enhancement. *J Vis*. 2002; 2(6):467–79. <https://doi.org/10.1167/2.6.4> PMID: [12678645](#).
20. Seiple W, Holopigian K, Szlyk JP, Wu C. Multidimensional visual field maps: relationships among local psychophysical and local electrophysiological measures. *J Rehabil Res Dev*. 2004; 41(3A):359–72. PMID: [15543452](#).
21. Silva MF, Maia-Lopes S, Mateus C, Guerreiro M, Sampaio J, Faria P, et al. Retinal and cortical patterns of spatial anisotropy in contrast sensitivity tasks. *Vision Res*. 2008; 48(1):127–35. <https://doi.org/10.1016/j.visres.2007.10.018> PMID: [18067943](#).
22. Silva MF, Mateus C, Reis A, Nunes S, Fonseca P, Castelo-Branco M. Asymmetry of visual sensory mechanisms: electrophysiological, structural, and psychophysical evidences. *J Vis*. 2010; 10(6):26. <https://doi.org/10.1167/10.6.26> PMID: [20884575](#).
23. Silva MF, d'Almeida OC, Oliveiros B, Mateus C, Castelo-Branco M. Development and aging of visual hemifield asymmetries in contrast sensitivity. *J Vis*. 2014; 14(12). <https://doi.org/10.1167/14.12.19> PMID: [25326605](#).

24. Carrasco M, Giordano AM, McElree B. Temporal performance fields: visual and attentional factors. *Vision Res.* 2004; 44(12):1351–65. <https://doi.org/10.1016/j.visres.2003.11.026> PMID: 15066395.
25. Chaikin JD, Corbin HH, Volkmann J. Mapping a field of short-time visual search. *Science.* 1962; 138(3547):1327–8. <https://doi.org/10.1126/science.138.3547.1327> PMID: 14019856.
26. Kristjansson A, Sigurdardottir HM. On the benefits of transient attention across the visual field. *Perception.* 2008; 37(5):747–64. <https://doi.org/10.1068/p5922> PMID: 18605148.
27. Krose BJ, Julesz B. The control and speed of shifts of attention. *Vision Res.* 1989; 29(11):1607–19. PMID: 2635484.
28. Najemnik J, Geisler WS. Eye movement statistics in humans are consistent with an optimal search strategy. *J Vis.* 2008; 8(3):4 1–14. <https://doi.org/10.1167/8.3.4> PMID: 18484810
29. Najemnik J, Geisler WS. Simple summation rule for optimal fixation selection in visual search. *Vision Res.* 2009; 49(10):1286–94. <https://doi.org/10.1016/j.visres.2008.12.005> PMID: 19138697.
30. Pretorius LL, Hanekom JJ. An accurate method for determining the conspicuity area associated with visual targets. *Hum Factors.* 2006; 48(4):774–84. <https://doi.org/10.1518/001872006779166370> PMID: 17240724.
31. Rezec AA, Dobkins KR. Attentional weighting: a possible account of visual field asymmetries in visual search? *Spat Vis.* 2004; 17(4–5):269–93. PMID: 15559106.
32. Fortenbaugh FC, Silver MA, Robertson LC. Individual differences in visual field shape modulate the effects of attention on the lower visual field advantage in crowding. *J Vis.* 2015; 15(2). <https://doi.org/10.1167/15.2.19> PMID: 25761337
33. He S, Cavanagh P, Intriligator J. Attentional resolution and the locus of visual awareness. *Nature.* 1996; 383(6598):334–7. <https://doi.org/10.1038/383334a0> PMID: 8848045.
34. Toet A, Levi DM. The two-dimensional shape of spatial interaction zones in the parafovea. *Vision Res.* 1992; 32(7):1349–57. PMID: 1455707.
35. Fuller S, Carrasco M. Perceptual consequences of visual performance fields: the case of the line motion illusion. *J Vis.* 2009; 9(4):13 1–7. <https://doi.org/10.1167/9.4.13> PMID: 19757922
36. Montaser-Kouhsari L, Carrasco M. Perceptual asymmetries are preserved in short-term memory tasks. *Atten Percept Psychophys.* 2009; 71(8):1782–92. <https://doi.org/10.3758/APP.71.8.1782> PMID: 19933562
37. Edgar GK, Smith AT. Hemifield differences in perceived spatial frequency. *Perception.* 1990; 19(6):759–66. <https://doi.org/10.1068/p190759> PMID: 2130373.
38. Rodieck RW. *The First Steps of Seeing.* Oxford: Oxford University Press; 1998.
39. Wandell BA. *Foundations of Vision:* Sinauer Associates; 1995.
40. Curcio CA, Sloan KR, Kalina RE, Hendrickson AE. Human photoreceptor topography. *J Comp Neurol.* 1990; 292(4):497–523. <https://doi.org/10.1002/cne.902920402> PMID: 2324310.
41. Song H, Chui TY, Zhong Z, Elsner AE, Burns SA. Variation of cone photoreceptor packing density with retinal eccentricity and age. *Invest Ophthalmol Vis Sci.* 2011; 52(10):7376–84. <https://doi.org/10.1167/iovs.11-7199> PMID: 21724911
42. Artal P. Image Formation in the Living Human Eye. *Annu Rev Vis Sci.* 2015; 1:1–17. <https://doi.org/10.1146/annurev-vision-082114-035905> PMID: 28532382.
43. Polans J, Jaeken B, McNabb RP, Artal P, Izatt JA. Wide-field optical model of the human eye with asymmetrically tilted and decentered lens that reproduces measured ocular aberrations. *Optica.* 2015; 2(2):124. <https://doi.org/10.1364/optica.2.000124>
44. Jaeken B, Artal P. Optical quality of emmetropic and myopic eyes in the periphery measured with high-angular resolution. *Invest Ophthalmol Vis Sci.* 2012; 53(7):3405–13. <https://doi.org/10.1167/iovs.11-8993> PMID: 22511633.
45. Anderson SJ, Mullen KT, Hess RF. Human peripheral spatial resolution for achromatic and chromatic stimuli: limits imposed by optical and retinal factors. *J Physiol.* 1991; 442:47–64. <https://doi.org/10.1113/jphysiol.1991.sp018781> PMID: 1798037
46. Banks MS, Sekuler AB, Anderson SJ. Peripheral spatial vision: limits imposed by optics, photoreceptors, and receptor pooling. *J Opt Soc Am A.* 1991; 8(11):1775–87. PMID: 1744774.
47. Bradley C, Abrams J, Geisler WS. Retina-V1 model of detectability across the visual field. *J Vis.* 2014; 14(12). <https://doi.org/10.1167/14.12.22> PMID: 25336179
48. Williams DR, Coletta NJ. Cone spacing and the visual resolution limit. *J Opt Soc Am A.* 1987; 4(8):1514–23. PMID: 3625331.
49. Farrell JE, Winawer J, Brainard DH, Wandell B. 27.2: Distinguished Paper: Modeling visible differences: The computational observer model. *SID Symposium Digest of Technical Papers2014.* p. 352–6.

50. Cottaris NP, Jiang H, Ding X, Wandell BA, Brainard DH. A computational-observer model of spatial contrast sensitivity: Effects of wave-front-based optics, cone-mosaic structure, and inference engine. *J Vis*. 2019; 19(4):8. <https://doi.org/10.1167/19.4.8> PMID: 30943530.
51. Brainard DH, Jiang H, Cottaris NP, Rieke F, Chichilnisky EJ, Farrell JE, et al. ISETBIO: Computational tools for modeling early human vision. *Imaging and Applied Optics 2015*; 2015/06/07; Arlington, Virginia: Optical Society of America; 2015.
52. Wyszecki G, Stiles WS. *Color science*. 2nd ed. New York: Wiley; 1982.
53. Geisler WS. Physical limits of acuity and hyperacuity. *J Opt Soc Am A*. 1984; 1(7):775–82. <https://doi.org/10.1364/JOSAA.1.000775> PMID: 6747742.
54. Cottaris NP, Rieke F, Wandell BA, Brainard DH, editors. *Computational observer modeling of the limits of human pattern resolution*. OSA Fall Vision Meeting; 2018; Reno, Nevada.
55. Mergenthaler K, Engbert R. Modeling the control of fixational eye movements with neurophysiological delays. *Phys Rev Lett*. 2007; 98(13):138104. <https://doi.org/10.1103/PhysRevLett.98.138104> PMID: 17501244.
56. Martinez-Conde S, Macknik SL, Hubel DH. The role of fixational eye movements in visual perception. *Nature reviews Neuroscience*. 2004; 5(3):229–40. <https://doi.org/10.1038/nrn1348> PMID: 14976522.
57. Martinez-Conde S, Macknik SL, Troncoso XG, Hubel DH. Microsaccades: a neurophysiological analysis. *Trends Neurosci*. 2009; 32(9):463–75. <https://doi.org/10.1016/j.tins.2009.05.006> PMID: 19716186.
58. Hofer H, Carroll J, Neitz J, Neitz M, Williams DR. Organization of the human trichromatic cone mosaic. *J Neurosci*. 2005; 25(42):9669–79. <https://doi.org/10.1523/JNEUROSCI.2414-05.2005> PMID: 16237171.
59. Roorda A, Williams DR. The arrangement of the three cone classes in the living human eye. *Nature*. 1999; 397(6719):520–2. <https://doi.org/10.1038/17383> PMID: 10028967.
60. Roorda A, Metha AB, Lennie P, Williams DR. Packing arrangement of the three cone classes in primate retina. *Vision Res*. 2001; 41(10–11):1291–306. PMID: 11322974.
61. Villegas EA, Alcon E, Artal P. Optical quality of the eye in subjects with normal and excellent visual acuity. *Invest Ophthalmol Vis Sci*. 2008; 49(10):4688–96. <https://doi.org/10.1167/iovs.08-2316> PMID: 18552387
62. Thibos LN. Retinal image quality for virtual eyes generated by a statistical model of ocular wavefront aberrations. *Ophthalmic Physiol Opt*. 2009; 29(3):288–91. <https://doi.org/10.1111/j.1475-1313.2009.00662.x> PMID: 19422560.
63. Thibos LN, Hong X, Bradley A, Cheng X. Statistical variation of aberration structure and image quality in a normal population of healthy eyes. *J Opt Soc Am A*. 2002; 19(12):2329–48. <https://doi.org/10.1364/JOSAA.19.002329>
64. Curcio CA, Sloan KR Jr., Packer O, Hendrickson AE, Kalina RE. Distribution of cones in human and monkey retina: individual variability and radial asymmetry. *Science*. 1987; 236(4801):579–82. <https://doi.org/10.1126/science.3576186> PMID: 3576186.
65. Jiang H, Cottaris N, Golden J, Brainard D, Farrell J, Wandell B, editors. *Simulating retinal encoding: factors influencing vernier acuity*. *Proceedings of Electronic Imaging*; 2017; Burlingame, CA.
66. Ding X, Radonjić A, Cottaris NP, Jiang H, Wandell BA, Brainard D. Computational-Observer analysis of illumination discrimination. *BioRxiv*. 2018.
67. Brainard DH, Cottaris NP, Radonjić A. The perception of colour and material in naturalistic tasks. *Interface Focus*. 2018; 8:20180012. <http://dx.doi.org/10.1098/rsfs.2018.0012>. PMID: 29951192
68. Wang X, Pedersen M, Thomas JB. The influence of chromatic aberration on mosaicking. 5th European Workshop on Visual Information Processing (EUVIP); Paris, France 2014.
69. Golden JR, Erickson-Davis C, Cottaris NP, Parthasarathy N, Rieke F, Brainard DH, Wandell BA, Chichilnisky EJ. Simulation of visual perception and learning with a retinal prosthesis. *J Neural Eng*. 2019; 16(2):025003. <https://doi.org/10.1088/1741-2552/aaf270> PMID: 30523985
70. Tuten WS, Cooper RF, Tiruveedhula P, Dubra A, Roorda A, Cottaris NP, Brainard DH, Morgan JIW. Spatial summation in the human fovea: the effect of optical aberrations and fixational eye movements. *J Vis*. 2018; 18(8):6. <https://doi.org/10.1167/18.8.6> PMID: 30105385
71. Curcio CA, Allen KA. Topography of ganglion cells in human retina. *J Comp Neurol*. 1990; 300(1):5–25. <https://doi.org/10.1002/cne.903000103> PMID: 2229487.
72. Perry VH, Cowey A. The ganglion cell and cone distributions in the monkey's retina: implications for central magnification factors. *Vision Res*. 1985; 25(12):1795–810. [https://doi.org/10.1016/0042-6989\(85\)90004-5](https://doi.org/10.1016/0042-6989(85)90004-5) PMID: 3832605.
73. Watson AB. A formula for human retinal ganglion cell receptive field density as a function of visual field location. *J Vis*. 2014; 14(7). <https://doi.org/10.1167/14.7.15> PMID: 24982468.

74. Mullen KT, Kingdom FA. Losses in peripheral colour sensitivity predicted from "hit and miss" post-receptoral cone connections. *Vision Res.* 1996; 36(13):1995–2000. PMID: [8759439](#).
75. Dacey DM. The mosaic of midget ganglion cells in the human retina. *J Neurosci.* 1993; 13(12):5334–55. <https://doi.org/10.1523/JNEUROSCI.13-12-05334.1993> PMID: [8254378](#).
76. Dacey DM, Petersen MR. Dendritic field size and morphology of midget and parasol ganglion cells of the human retina. *Proc Natl Acad Sci U S A.* 1992; 89(20):9666–70. <https://doi.org/10.1073/pnas.89.20.9666> PMID: [1409680](#)
77. Croner LJ, Kaplan E. Receptive fields of P and M ganglion cells across the primate retina. *Vision Res.* 1995; 35(1):7–24. [https://doi.org/10.1016/0042-6989\(94\)E0066-T](https://doi.org/10.1016/0042-6989(94)E0066-T) PMID: [7839612](#).
78. Liu T, Heeger DJ, Carrasco M. Neural correlates of the visual vertical meridian asymmetry. *J Vis.* 2006; 6(11):1294–306. PMID: [17209736](#)
79. Benson NC, Butt OH, Datta R, Radoeva PD, Brainard DH, Aguirre GK. The retinotopic organization of striate cortex is well predicted by surface topology. *Current biology: CB.* 2012; 22(21):2081–5. <https://doi.org/10.1016/j.cub.2012.09.014> PMID: [23041195](#)
80. Silva MF, Brascamp JW, Ferreira S, Castelo-Branco M, Dumoulin SO, Harvey BM. Radial asymmetries in population receptive field size and cortical magnification factor in early visual cortex. *NeuroImage.* 2018; 167:41–52. <https://doi.org/10.1016/j.neuroimage.2017.11.021> PMID: [29155078](#).
81. De Valois RL, Albrecht DG, Thorell LG. Spatial frequency selectivity of cells in macaque visual cortex. *Vision Res.* 1982; 22(5):545–59. [https://doi.org/10.1016/0042-6989\(82\)90113-4](https://doi.org/10.1016/0042-6989(82)90113-4) PMID: [7112954](#).
82. Legge GE. Binocular contrast summation—I. Detection and discrimination. *Vision Res.* 1984; 24(4):373–83. PMID: [6740958](#).
83. Barlow HB, Blakemore C, Pettigrew JD. The neural mechanism of binocular depth discrimination. *J Physiol.* 1967; 193(2):327–42. <https://doi.org/10.1113/jphysiol.1967.sp008360> PMID: [6065881](#)
84. Hubel DH, Wiesel TN. Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *J Physiol.* 1962; 160:106–54. <https://doi.org/10.1113/jphysiol.1962.sp006837> PMID: [14449617](#)
85. Poggio GF, Fischer B. Binocular interaction and depth sensitivity in striate and prestriate cortex of behaving rhesus monkey. *Journal of neurophysiology.* 1977; 40(6):1392–405. <https://doi.org/10.1152/jn.1977.40.6.1392> PMID: [411898](#).
86. Brascamp JW. Neurophysiology: Charting the Confluence of the Two Eyes' Information Streams. *Current biology: CB.* 2019; 29(4):R134–R6. PMID: [30779904](#).
87. Dougherty K, Cox MA, Westerberg JA, Maier A. Binocular Modulation of Monocular V1 Neurons. *Current biology: CB.* 2019; 29(3):381–91 e4. <https://doi.org/10.1016/j.cub.2018.12.004> PMID: [30661798](#)
88. Robson AG, Moreland JD, Pauleikhoff D, Morrissey T, Holder GE, Fitzke FW, et al. Macular pigment density and distribution: comparison of fundus autofluorescence with minimum motion photometry. *Vision Res.* 2003; 43(16):1765–75. [https://doi.org/10.1016/S0042-6989\(03\)00280-3](https://doi.org/10.1016/S0042-6989(03)00280-3) PMID: [12818346](#).
89. Green DM, Swets JA. Signal detection theory and psychophysics. New York: Wiley; 1966. 455 p.
90. Geisler WS. Contributions of ideal observer theory to vision research. *Vision Res.* 2011; 51(7):771–81. <https://doi.org/10.1016/j.visres.2010.09.027> PMID: [20920517](#)
91. Najemnik J, Geisler WS. Optimal eye movement strategies in visual search. *Nature.* 2005; 434(7031):387–91. <https://doi.org/10.1038/nature03390> PMID: [15772663](#).
92. Banks MS, Geisler WS, Bennett PJ. The physical limits of grating visibility. *Vision Res.* 1987; 27(11):1915–24. [https://doi.org/10.1016/0042-6989\(87\)90057-5](https://doi.org/10.1016/0042-6989(87)90057-5) PMID: [3447346](#).
93. Blakemore C, Campbell FW. On the existence of neurones in the human visual system selectively sensitive to the orientation and size of retinal images. *J Physiol.* 1969; 203(1):237–60. <https://doi.org/10.1113/jphysiol.1969.sp008862> PMID: [5821879](#)
94. Carandini M, Demb JB, Mante V, Tolhurst DJ, Dan Y, Olshausen BA, et al. Do we know what the early visual system does? *J Neurosci.* 2005; 25(46):10577–97. <https://doi.org/10.1523/JNEUROSCI.3726-05.2005> PMID: [16291931](#).
95. Rust NC, Schwartz O, Movshon JA, Simoncelli EP. Spatiotemporal elements of macaque v1 receptive fields. *Neuron.* 2005; 46(6):945–56. <https://doi.org/10.1016/j.neuron.2005.05.021> PMID: [15953422](#).
96. Zernike F, Stratton FJM. Diffraction theory of the knife-edge test and its improved form, the phase-contrast method. *Monthly Notices of the Royal Astronomical Society.* 1934; 94(5):377–84. <https://doi.org/10.1093/mnras/94.5.377>
97. Stockman A, Sharpe LT. The spectral sensitivities of the middle- and long-wavelength-sensitive cones derived from measurements in observers of known genotype. *Vision Res.* 2000; 40(13):1711–37. [https://doi.org/10.1016/S0042-6989\(00\)00021-3](https://doi.org/10.1016/S0042-6989(00)00021-3) PMID: [10814758](#).