Research Article Vol. 42, No. 5 / May 2025 / Journal of the Optical Society of America A B495

Journal of the Optical Society of America A

OPTICS, IMAGE SCIENCE, AND VISION

Emergence of strategic cone weighting from efficient coding of spatiochromatic natural images

ALEXANDER BELSTEN[1,2,*] AND BRUNO A. OLSHAUSEN[1,2,3]

[1]Redwood Center for Theoretical Neuroscience, University of California, Berkeley, Berkeley, California 94720, USA
[2]Herbert Wertheim School of Optometry & Vision Science, University of California, Berkeley, Berkeley, California 94720, USA
[3]Helen Wills Neuroscience Institute, University of California, Berkeley, Berkeley, California 94720, USA
*belsten@berkeley.edu

Received 16 October 2024; revised 4 March 2025; accepted 4 March 2025; posted 5 March 2025; published 2 April 2025

We develop an efficient coding model to address how a population of retinal ganglion cells (RGCs) can optimally combine signals from the retinal cone mosaic to maximize information transfer through the optic nerve. The model takes into account the redundancies inherent in color natural images and predicts how they should be reduced in order to make the best use of channel capacity in the optic nerve, given metabolic constraints, wiring constraints, and input and channel noise. RGCs are modeled as a set of linear–nonlinear neurons whose instantaneous firing rate is computed via a weighted sum of cone responses from a simulated L- and M-cone mosaic, followed by a rectifying nonlinearity. When adapted to a set of calibrated color natural images so as to maximize mutual information between the retinal image and RGC outputs, the learned weights exhibit a circularly symmetric, center–surround structure, and the population of RGCs tile visual space via ON- and OFF-mosaics, in line with previous studies that use only luminance variations in natural scenes. Over a range of cone-to-neuron ratios, the model RGCs strategically weight cones of a particular spectral type to construct a stronger form of L–M cone-opponency than would be obtained with purely random sampling, implying that such a specific arrangement increases information transfer through the optic nerve. Additionally, we find that the degree of cone-type-specific adaptation varies with the amount of noise in the cone activations, with less noise leading to more specific adaptation. The results of this study point to the benefits of strategic cone weighting for maximizing information transfer for spatiochromatic natural scenes.

https://doi.org/10.1364/JOSAA.545141

1. INTRODUCTION

Color vision originates from cone cells with different spectral sensitivities that are combined in an opponent manner. In trichromatic humans and primates, long- (L) and medium- (M) wavelength-sensitive cones are compared via midget retinal ganglion cells (RGCs), which form the so-called L–M cone-opponent pathway [1]. The spatial organization of cones into a mosaic, along with the spatially antagonistic center–surround receptive fields of midget RGCs, facilitates the generation of L–M cone-opponency [2,3]. However, the precise connectivity underlying these receptive fields remains unresolved, and two primary possibilities have been proposed to explain how midget RGCs construct cone-opponent signals [4].

One possibility is that midget RGCs construct cone-opponent signals by forming circularly symmetric center–surround receptive fields that are indiscriminate to cone type. In this scenario, cone-opponent signals emerge by chance from the random arrangement of cone types in the mosaic [4]. Alternatively, midget RGCs may exhibit cone-type selective receptive fields, where inputs to the center and/or surround are biased toward specific cone types, thus amplifying cone-opponent signals. While experimental evidence supports the presence of cone-type selective biasing in some contexts [5–7], other studies suggest that non-selective inputs can adequately explain key features of cone-opponent responses [4,8]. Thus, it remains an open question as to how or why RGCs construct cone-opponent signals.

Efficient coding principles [9,10] provide a powerful framework for understanding the function of sensory systems, including the retina. This framework posits that neural systems are adapted to the statistics of sensory input so as to maximize the information conveyed in a neural representation subject to energetic costs and resource constraints. In the retina, efficient coding has led to a deeper understanding of RGC center–surround receptive fields [11–14], preference for increments or decrements of light (i.e., ON- and OFF-cell types) [15], mosaic alignment [16,17], and temporal processing [18–20]. Of particular relevance to this work is the efficient coding of chromatic

structure in natural scenes. In previous work, Ruderman *et al*. [21] found that decorrelating simulated cone activations to natural images using an orthonormal basis reveals two chromatic axes that have close alignment with the physiological cone-opponent representation [3] (see also Refs. [22–24]). In addition, Doi *et al*. [13] derived a whitening transform for decorrelating simulated, spatially non-overlapping cone activations and found circularly symmetric center–surround receptive fields, but also found no evidence for cone-type specific wiring that strategically constructs a L–M opponent signal (see also Ref. [25]). While the approach of computing a decorrelating basis (via zero-phase whitening, for example) for stimulus statistics provides an insightful first-order approximation to retinal function, it has notable limitations. Specifically, it relies solely on pairwise (second-order) statistics and thus cannot capture higher-order structures in natural images [26,27]. Interestingly, Doi *et al*. [13] applied independent component analysis to the whitened responses, which is sensitive to higher-order structure. The resulting receptive fields exhibit a bias toward specific cone types but now with an oriented structure as found in the cortex, as opposed to the circularly symmetric center–surround type as found in the retina. Additionally, both ICA and the whitening model assume no noise in the input stimulus or neural representation, and their analysis was restricted to simulations with 1:1 input-to-output ratios, preventing it from including features such as the bottleneck of the optic nerve, where there is a convergence of cones onto retinal ganglion cells.

Here, we adopt a more general approach by adapting a model population of noisy linear–nonlinear neurons [15] to efficiently encode simulated non-overlapping L- and M-cone activations to spatiochromatic natural scenes. Specifically, the linear weights and nonlinearities of each model neuron are jointly adapted to maximize mutual information between cone inputs and their corresponding noisy output responses—subject to a constraint on the mean response rate (Fig. 1). By analyzing the neurons' optimal linear weights, we quantitatively examine the role of adaptation to the spectral topography of the cone mosaic in optimizing information transfer.

In line with prior work [15,17], we find that the optimized model weights exhibit a circularly symmetric center–surround structure and form ON- and OFF-mosaics that tile visual space. Notably, we find that the model adapts these weights to the cone mosaic in both the center and surround regions to produce a strong cone-opponent signal. The model achieves this by strategically weighting cone inputs rather than relying on spatial clumping within the cone mosaic. Moreover, by varying the number of neurons in the simulation, we demonstrate that adaptation to the cone mosaic is more pronounced in models simulating retinal regions closer to the fovea. We also find that the degree of adaptation depends on the level of noise injected into the cone activations, with higher noise reducing the degree of adaptation. Together, these findings indicate that cone-type selective receptive fields provide a computationally favorable representation of chromaticity in natural scenes.

2. MODEL

Our model extends the joint source-channel coding framework described in Ref. [15] by incorporating a cone mosaic

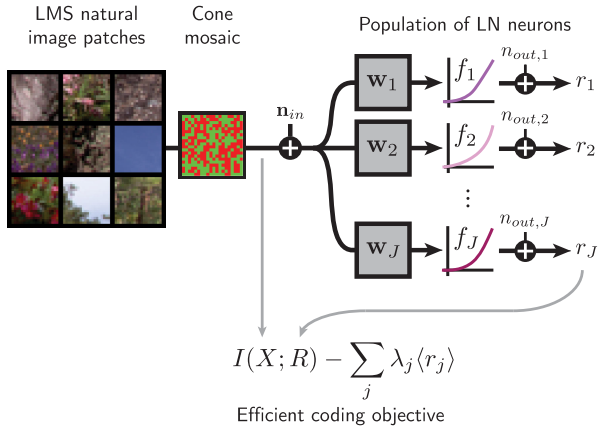

Fig. 1. Efficient coding model of the retina. A population of linear–nonlinear (LN) neurons is optimized to maximize an efficient coding objective over a dataset of simulated cone activations to natural image patches. Input LMS image patches are sampled by a simulated non-overlapping cone mosaic, to which noise is added. Each neuron's response is computed by applying a weighted sum to these noisy activations, followed by a rectifying nonlinearity and additional output noise. The linear weights and nonlinearities are adapted to maximize information transfer, subject to a constraint on the mean firing rate.

with distinct spectral sensitivities (Fig. 1). The model consists of a population of J linear–nonlinear neurons. The linear component u_j for neuron j is computed via

$$u_j = \mathbf{w}_j^T(\mathbf{x} + \mathbf{n}_{\text{in}}), \tag{1}$$

where $\mathbf{w}_j$ is the unit-norm weight vector of neuron j, $\mathbf{x} \in \mathbb{R}^D$ is a vector of D cone activations, and $\mathbf{n}_{\text{in}} \sim \mathcal{N}(0, \mathbf{C}_{\text{in}})$ is input noise mimicking randomness in cone activations (due to photon shot noise and neural noise). The response r_j of neuron j is computed by rectifying the linear component u_j with nonlinearity f_j and adding output noise:

$$r_j = f_j(u_j) + n_{j,\text{out}}, \tag{2}$$

where noise $\mathbf{n}_{\text{out}} \sim \mathcal{N}(0, \mathbf{C}_{\text{out}})$ simulates variability in neuron spike rates. The adjustable nonlinearity $f_j(u_j) = \theta_j \eta(u_j - \tau_j)$ is composed of the softplus nonlinearity $\eta(y) = \log(1 + e^{\beta y})/\beta$, a learnable gain θ_j, and a shift τ_j (we set $\beta = 2.5$). Input and output noise are defined as i.i.d., thus their covariances can be written as $\mathbf{C}_{\text{in}} = \sigma_{\text{in}}^2 \mathbf{I}_D$ and $\mathbf{C}_{\text{out}} = \sigma_{\text{out}}^2 \mathbf{I}_J$, where $\mathbf{I}_N$ is the identity matrix of dimension N.

The model parameters $\{\mathbf{w}_j, \theta_j, \tau_j\}_{j=1}^J$ are fit to maximize the mutual information $I(X; R)$ between a dataset of cone activations X and their corresponding noisy output responses R, subject to a constraint on the mean firing rate, resulting in the following objective:

$$\max_{\{\mathbf{w}_j, \theta_j, \tau_j\}_{j=1}^J} I(X; R) - \sum_{j=1}^{J} \lambda_j \langle r_j \rangle. \tag{3}$$

Linearization of f_j via first-order Taylor expansion is performed to make the calculation of mutual information tractable (see Section 3.B for details). Lagrange multipliers λ_j are adjusted during optimization so as to constrain the mean activation of each neuron to be equal to 1 in expectation

(i.e., enforce that $\langle r_j \rangle = 1$). Model parameters are optimized via stochastic gradient descent, and after each gradient update, the weights are normalized such that $||\mathbf{w}_j||_2 = 1 \; \forall j$.

3. SIMULATION METHODS

A. Dataset of Model Cone Activations

The dataset of simulated L- and M-cone activations is extracted from a dataset of calibrated natural images [13]. A total of 102,300 randomly extracted 20×20 pixel patches are obtained from the images, which are then sampled by a simulated cone mosaic of equivalent size, consisting of non-overlapping L- and M-cones (resulting in $D = 400$ cones). We do not model S-cones in order to compare with Refs. [4,7], although it is of future interest. Cone types at each pixel location are assigned randomly using a fair coin flip. This results in an equal distribution of 200 L-cones and 200 M-cones, which is within the large range of observed L-to-M ratios and the spatial randomness of cone-types reported empirically [28]. The resulting mosaic can be seen in Fig. 4(a). No other preprocessing methods were performed (i.e., whitening) other than shifting and scaling the entire dataset to have a zero mean and unit variance.

B. Maximizing Mutual Information

Mutual information $I(X;\, R)$ between the input distribution X and the output distribution R can be written as $I(X;\, R) = H(R) - H(R|X)$. The first term $H(R)$ is the entropy over the responses, and the second term $H(R|X)$ is the entropy of the response distribution given a fixed input. Following the conventions of Refs. [15–17,19], we locally approximate the nonlinearity with a first-order Taylor expansion, such that the responses $\mathbf{r}$ can be approximated as

$$\mathbf{r} \approx \mathbf{G}\mathbf{W}^T(\mathbf{x} + \mathbf{n}_{\text{in}}) + \mathbf{n}_r + \mathbf{f}_0, \tag{4}$$

where $\mathbf{W}$ is a matrix containing the weights, $\mathbf{G} = \text{diag}(f'(\mathbf{W}^T(\mathbf{x} + \mathbf{n}_{\text{in}})))$ such that the element-wise nonlinearity $f'(\mathbf{y})$ is $\frac{d}{dy} f_j(y)$ at index j, and $\mathbf{f}_0$ is a constant. Similarly, we assume a Gaussian prior over X with covariance $\mathbf{C}_x$, which is estimated from the dataset. The local linear approximation and Gaussian prior enable local analytic Gaussian approximations to the output response distribution and conditional distribution (characterized by covariances $\mathbf{C}_r$ and $\mathbf{C}_{r|x}$, respectively) found in the mutual information objective:

$$I(X;\, R) = H(R) - H(R|X) \tag{5}$$

$$\approx \frac{1}{2}\log\det(\mathbf{C}_r) - \frac{1}{2}\log\det(\mathbf{C}_{r|x}) \tag{6}$$

$$= \frac{1}{2}\log\det\left(\mathbf{G}\mathbf{W}^T(\mathbf{C}_x + \mathbf{C}_{\text{in}})\mathbf{W}\mathbf{G} + \mathbf{C}_{\text{out}}\right) - \frac{1}{2}\log\det\left(\mathbf{G}\mathbf{W}^T\mathbf{C}_{\text{in}}\mathbf{W}\mathbf{G} + \mathbf{C}_{\text{out}}\right). \tag{7}$$

We optimize the approximate mutual information in Eq. (7) above in Eq. (3). Note that this form differs from the one originating in Ref. [15] and is more similar to that of Linsker [29,30]; however, they are mathematically equivalent due to the dual form of mutual information $I(X;\, R) = H(R) - H(R|X) = H(X) - H(X|R)$. In practice, the approximation made here is more efficient to compute, as multiple matrix inversions are not necessary.

4. EXPERIMENTS

A. Model Neurons Are Adapted to the Spectral Topography of the Cone Mosaic

In this section, we analyze optimal models with various numbers of neurons and investigate the degree to which they adapt to the cone mosaic to produce a cone-opponent representation. We do so under a noise regime ($\sigma_{\text{in}} = 0.2$ and $\sigma_{\text{out}} = 2.0$), which has previously been shown to result in center–surround weights [15,17] (in Section 4.C, we explore how noise effects the models). In accordance with these studies, the optimal weights bifurcate into ON- and OFF-types with circularly symmetric center–surround profiles [Fig. 2(a)]. These models spatially coordinate the two types to form mosaics that tile space [Fig. 2(b)]. Moreover, models with fewer neurons exhibit smaller center–surround profiles compared to those with more neurons and transfer more information [Fig. 3(a)].

To investigate the degree of adaptation to the non-overlapping spectral topography of the simulated L- and M-cone mosaic, we compared the information transfer using the original mosaic to that of other random mosaics. Specifically, we computed the average information transfer (measured in bits) for each input image patch in the dataset using the original cone mosaic to which the model is adapted. We then permuted the cone identities in the mosaic so as to scramble the spectral topography. Cone activations to all image patches in the dataset were recalculated for the permuted mosaics, and information transfer was re-evaluated for each one.

We observe a slight reduction in information transfer due to mosaic permutation [Fig. 3(a)], supporting the notion that the models strategically adapt to the topography of the cone mosaic. To assess the significance of this reduction, the difference between information transfer in the original and permuted mosaics was plotted [Fig. 3(b)]. While small, this reduction is statistically robust. Moreover, as the number of neurons in the simulation increases, a disproportionate reduction in mutual information is seen, indicating that smaller center–surround profiles exhibit more adaptation to the cone topography of the mosaic [Fig. 3(b) and see Fig. S1 for percent reduction in mutual information as opposed to bits].

To identify where the center–surround weights' cone-mosaic-specific specialization occurs, we applied the analysis of Field *et al*. [7]. While the full details of this method are provided in their paper, a brief description is included here. The following purity index [Fig. 3(c)] quantifies the degree of strategic weighting in the model's learned weights $\{\mathbf{w}_j\}_{j=1}^J$:

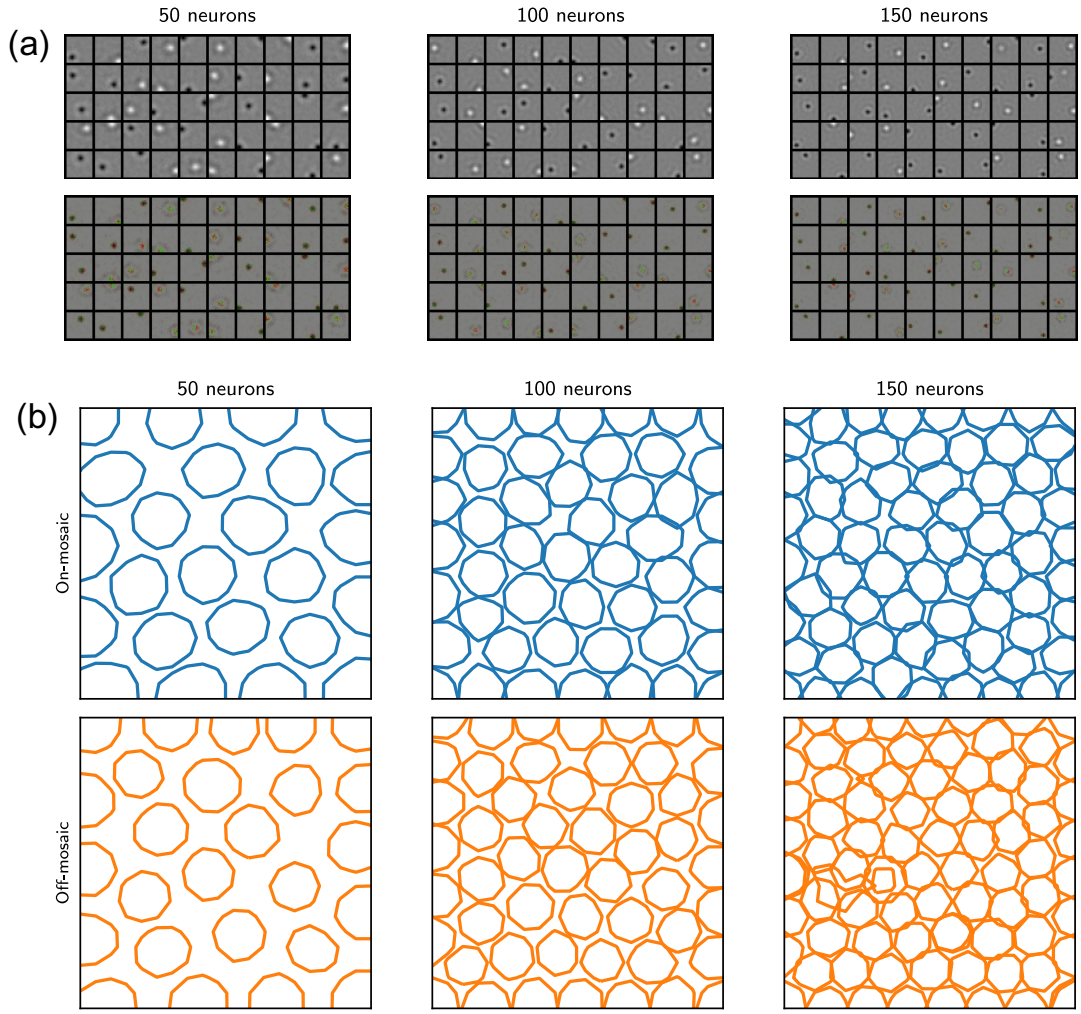

Fig. 2. Learned weights bifurcate into ON- and OFF-types and tile space. (a) Top row: A subset of optimal weights for models with varying numbers of neurons. Bottom row: visualizing the same weights false-colored by the cone type they sample from. The luminance value is set to the weight value, the saturation is set to the weight magnitude, and the hue is set to red and green for L- and M-cones, respectively. (b) Contours of ON and OFF weights at their half-maximum magnitude value.

$$c_L = \sum_{i \in \text{L-cones in center/surround}} |\mathbf{w}_i|, \quad c_M = \sum_{i \in \text{M-cones in center/surround}} |\mathbf{w}_i|,$$

$$\text{Purity index} = \frac{c_L - c_M}{c_L + c_M}. \tag{8}$$

The purity index ranges from −1 to 1, where −1 indicates exclusive M-cone connections, 0 represents equal contributions from L- and M-cones, and 1 signifies exclusive L-cone connections. This metric is calculated separately for the center and surround regions of the weights, providing each model neuron with a center and surround purity index. These regions are defined as follows: if the cell is ON-type (OFF-type), then the center is defined to be all positive (negative) weights in the weight vector, and the surround is all negative (positive) weights in the weight vector. Intuitively, models with weights that strategically organize to favor one cone type will produce purity indices clustered near the extremes of the range. The standard deviation calculated over the distribution of purity indices for all learned weights in a particular model is used to characterize this shift toward the extremes. By permuting cone identities within the mosaic and recalculating the standard deviations, the degree of biased sampling in the center and surround regions can be determined.

The results [Fig. 3(d)] reveal that biased sampling occurs in both the center and surround regions, as the standard deviations of the purity indices are significantly higher for the original mosaic compared to the permuted mosaics. Additionally, increasing the number of neurons in the simulation enhances the standard deviation of the original mosaic, again suggesting a stronger adaptation to the cone mosaic. An exception occurs in the $J = 200$ model, of which a subset of the weights departs from center–surround motifs (Fig. S2); however, this model still exhibits adaptation to the mosaic [Figs. 3(a) and 3(b)].

This analysis demonstrates that the model's weights adapt to the spectral topography of the cone mosaic to establish connections biased toward a particular cone type (L or M). However, the analysis does not elucidate the precise relationship

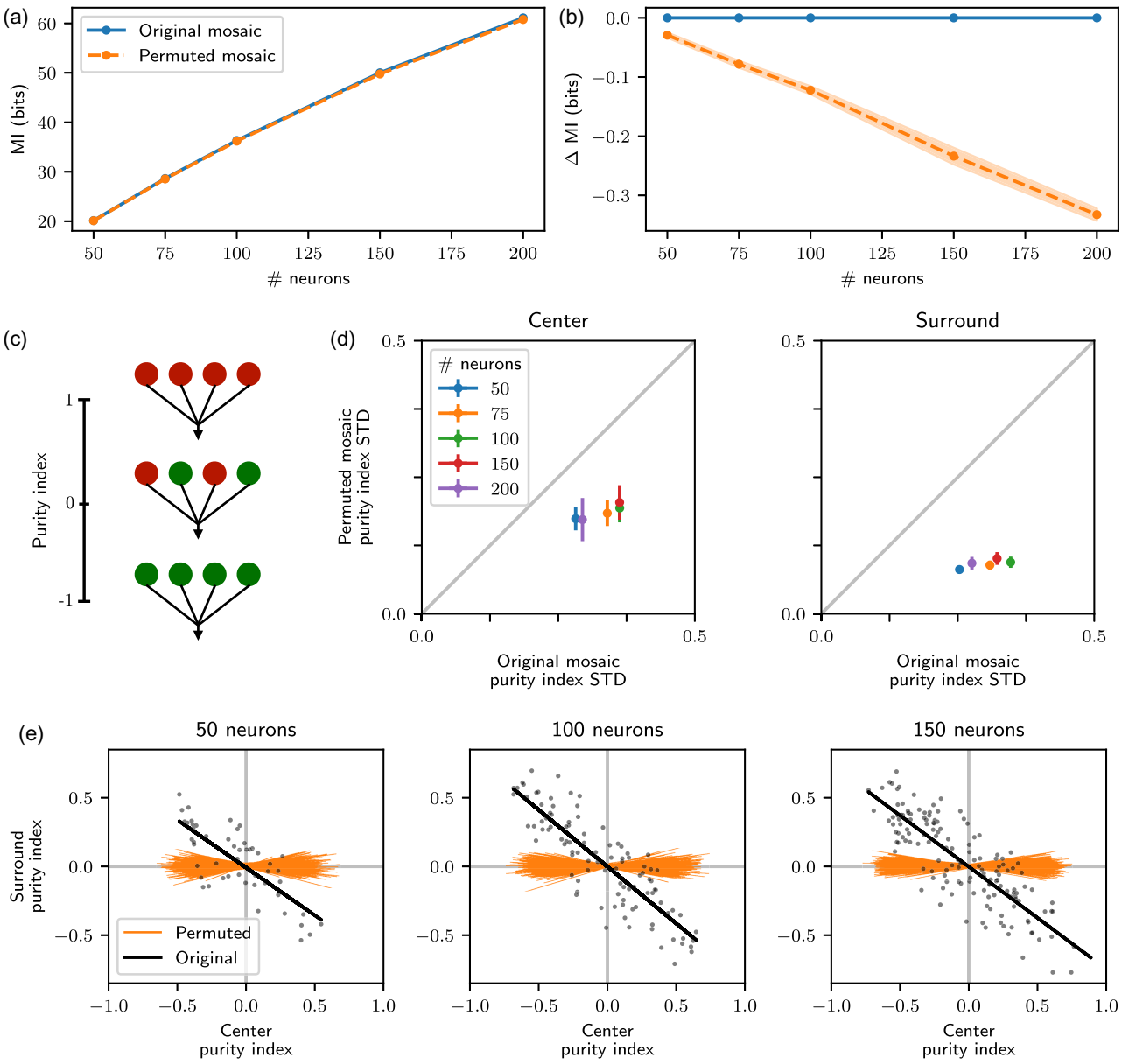

Fig. 3. Optimal weights are adapted to the cone mosaic to form a strong cone-opponent signal. (a) Mutual information for the original mosaic and the mean mutual information after permuting the cone types in the mosaic as a function of the number of neurons in the simulation. (b) Comparing the permuted mosaic mutual information to that of the original. The shaded region indicates one standard deviation. (c) Schematic of the purity index used in the analysis. The index ranges from −1 to 1 and is computed over the center or surround region in isolation. (d) Comparing the observed purity index standard deviation to the permuted purity index standard deviation for the linear weight center (left) and surround (right). Error bars indicate one standard deviation computed over 1000 permutations. (e) The black scatter plot shows the purity indices for the center and surround regions computed over the adapted mosaic. The black line of best fit is plotted over the observed range of center purity indices. Orange lines indicated the same but computed over a permuted cone mosaic (repeated 1000 times resulting in 1000 orange lines). Scatter plots are left out for the permuted mosaics. The analysis is shown for model simulations with varying numbers of neurons.

between the center and surround region's adaptation (i.e., if they strategically organize to produce a cone-opponent signal). By plotting the center purity index against the surround purity index, we can determine that the model weights strategically construct L–M cone-opponent signals via the bias observed in the center–surround weights [Fig. 3(e)]. Specifically, stronger L-cone inputs in the center are accompanied by stronger M-cone inputs in the surround, and vice versa. Interestingly, not all of the model RGC weights exhibit strong L–M opponency, as many points lie near the origin in Fig. 3(e). However, when comparing the overall trend of the adapted mosaic to that of the permuted mosaics [by comparing the black line to the orange lines in Fig. 3(e)], it is clear that the model strategically constructs strongly cone-opponent neurons.

B. Strategic Weighting Underlies Cone-Opponency

The model could construct a cone-opponent signal from the cone mosaic through two possible mechanisms: exploitation

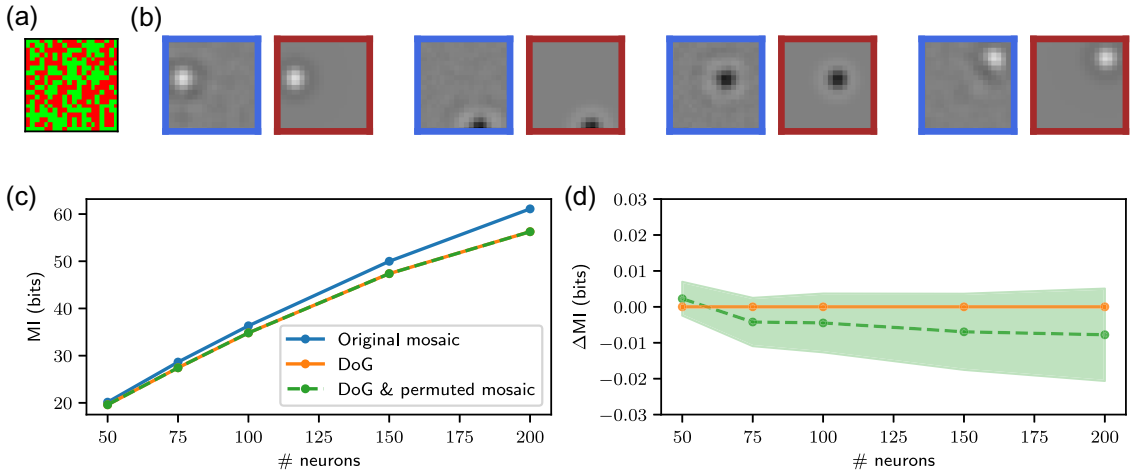

Fig. 4. Optimal weights adapt to the mosaic by leveraging strategic weighting from the cone mosaic over cone clumping. (a) Cone mosaic used to sample the spatiochromatic image patches. (b) Four adapted weights for the $J = 100$ model shown in the blue border and their corresponding DoG fits in the red border. (c) Mutual information for the model adapted to the original mosaic, the model with weights fit by DoG, and the mean mutual information for the DoG model after permuting the cone types in the original mosaic. (d) Comparison of mutual information for the DoG model and the DoG model under permutation. The shaded region indicated one standard deviation.

of cone clumping or strategic weighting. Despite the random assignment of cone identities in the mosaic, the resulting pattern shows spectral-type clumping [Fig. 4(a)]. One possibility is that the model leverages this clumping while maintaining a radially symmetric center–surround structure to produce the opponent signal [4]. Alternatively, the model could deviate from a purely circularly symmetric center–surround configuration by strategically adjusting weights based on spectral type, a process referred to here as strategic weighting. To assess the extent to which each mechanism contributes, the optimized weights were fitted with a parametric difference of Gaussians (DoG) model. The DoG model is a common parametrization of center–surround profiles of RGCs [4,17,19]. It describes the spatial weight vector $\mathbf{m}$ over spatial position $\mathbf{z} \in \mathbb{R}^2$ as

$$\mathbf{m}(\mathbf{z}) = a\left(e^{-b(\mathbf{z}-\boldsymbol{\mu})^2} - c\,e^{-d(\mathbf{z}-\boldsymbol{\mu})^2}\right), \tag{9}$$

where adapted parameters are defined as follows: $\boldsymbol{\mu} \in \mathbb{R}^2$ defines the spatial mean of both Gaussians, b, $d \in \mathbb{R}^+$ correspond to the positive-only scales of the center and surround Gaussians, respectively, and $c \in \mathbb{R}^+$ sets the relative weighting between the two. The parameter $a \in \{1, -1\}$ is set to correspond to the type of the learned linear weight (ON-/OFF-type, respectively). These parameters were optimized for each learned weight vector $\{\mathbf{w}_j\}_{j=1}^{J}$ to minimize the squared error between the DoG model weights, $\mathbf{m}$, and the learned weights, $\mathbf{w}_j$. The resulting fits produce spatial profiles that have a strong resemblance to the original weights [Fig. 4(b) and Fig. S4].

The reduction of information transfer was then quantified for two scenarios: first, the loss due to the DoG fit alone, and second, the additional loss incurred by permuting the cone mosaic with the DoG fit in place. A significant reduction in mutual information is observed when comparing the DoG model to the adapted model [Fig. 4(c)]. Furthermore, when the mosaic is permuted with the DoG weights, the loss in mutual information is comparatively small [Fig. 4(d)]. This disparity suggests that the model primarily relies on strategic weighting rather than exploiting cone clumping to generate the cone-opponent signal. Additionally, we found that the DoG-fit weights do not exhibit strong cone-type-specific sampling (Fig. S4). To further validate this result, models with parametric DoG weight vectors were directly optimized for mutual information (instead of adapting DoG fits to pre-optimized weights). Consistent with the findings of Fig. 4(d), these models also showed minimal mutual information loss under cone permutation (Fig. S5).

C. Mosaic Adaptation Depends on Input Noise

The sections above demonstrate that the model's weight vectors adapt to the cone mosaic's topography by strategically adjusting the weight for each input to produce a cone-opponent signal within a specific noise regime. This section examines how the input σ_{in} and output σ_{out} noise parameters influence the level of adaptation to the mosaic. To do so, we optimize $J = 100$ neuron models with varying conditions of input and output noise and investigate the degree to which they adapt to the mosaic.

Consistent circularly symmetric center–surround weights were observed in noise conditions where $0 \leq \sigma_{\text{in}} \leq 0.3$ and $1.5 \leq \sigma_{\text{out}} \leq 3.0$ (Fig. S6). Therefore, the analysis focused on models trained within these noise ranges. As expected, models optimized in low-noise regimes exhibited higher information transfer compared to those trained under high-noise conditions [Fig. 5(a)]. Interestingly, when the cone mosaic was permuted and the mutual information was reevaluated, models trained in low-input-noise regimes showed a greater reduction in information transfer due to mosaic permutation [Figs. 5(b) and 5(c)].

This result aligns with intuition: in high-input-noise conditions, subtle differences between the highly correlated L- and M-cone activations [21,25] are more likely to be masked by noise. Consequently, the chromatic signal is smaller; thus, the model exhibits less cone-type-specific adaptation. In contrast, when input noise is low, the model can detect the chromatic

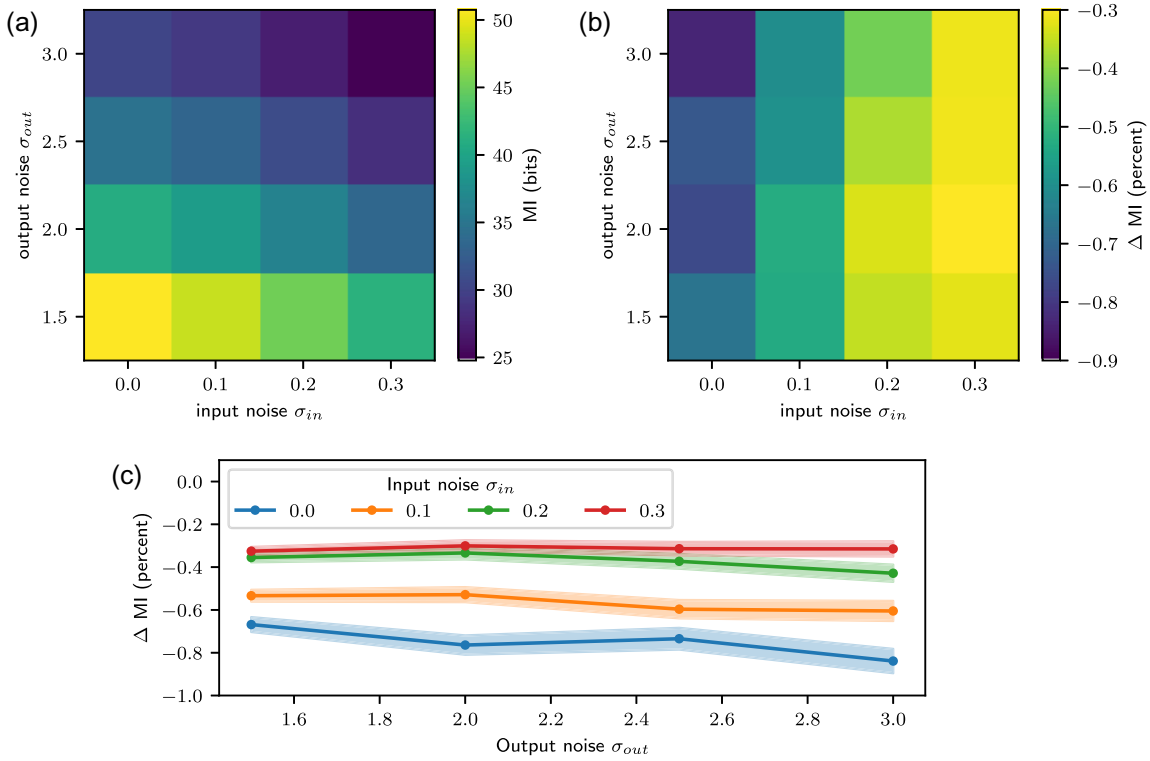

Fig. 5. Less adaptation to the mosaic occurs when input noise is larger. (a) Mutual information for $J = 100$ models adapted in different noise environments. (b) Mean loss in mutual information (as a percentage) due to permuting the mosaic. (c) Same data as (b) but shown with error bars that indicate one standard deviation over 1000 permutations.

signal more effectively and adjust its weight vectors to optimize encoding. Interestingly, the output noise did not seem to have a strong effect on cone-type-specific adaptation to the mosaic; however, it did change the spatial scale of the circularly symmetric center–surround receptive fields (Fig. S6).

5. DISCUSSION

Although it was originally thought that chromatic visual information was encoded in the retina by excitatory connections made exclusively to one type of cone and inhibitory connections made to another [31], more recent investigations have revealed a more complex and mixed representation. Furthermore, some have proposed that completely indiscriminate center–surround receptive fields suffice to encode chromatic information outside the fovea, where RGCs pool inputs from multiple cones [4]. While the results of this study do not answer the empirical question of whether or not RGCs strategically adapt to the spectral topography of the mosaic, it does provide evidence supporting its utility through simulation.

We find that by maximizing information transfer between cone activations and the responses of a population of linear–nonlinear neurons, the neurons' spatial weights adapt to the topography of the L- and M-cone mosaic. Specifically, the model strategically weights inputs toward a particular cone type in both the center and surround regions of their weight vectors, producing a strong L–M cone-opponent signal. Moreover, this adaptation enhances information transfer, particularly for models with more neurons (simulating areas of the retina at lower eccentricities). Additionally, we demonstrate that adaptation to the mosaic can be influenced by additive noise in the input cone activations, with less noise leading to more adaptation to the mosaic (Fig. 5). These results corroborate the benefits of biased cone sampling, which have been observed empirically [6,7]. Moreover, our finding of cone-type specificity is in line with prior work on information maximization applied to simulated non-overlapping cone activations, which revealed cone-type specificity [13,25]. This consistency reinforces the idea that learning cone-type-specific wiring is driven by higher-order statistics. A key distinction in our work is the explicit assumption of noise in the model, which leads to the emergence of center–surround receptive fields [15].

Although the information loss from cone permutation is statistically robust [Fig. 3(b)], its magnitude is relatively small [bear in mind, however, that the mutual information objective being optimized is the approximation seen in Eq. (7)]. The extent to which this small effect is truly beneficial remains an open question. This minor gain could compound significantly when considering the entirety of the retina (as opposed to the 400 cones that we simulated here) and downstream cortical mechanisms that amplify signals via pooling. It is important to acknowledge that the measure of mutual information used in this study is a proxy for the ultimate objective of any organism—survival. Our premise is that these two objectives are strongly related to one another, and it could well be the case that slight changes in mutual information lead to large changes in utility to the organism, or vice versa.

An important consideration is the mechanism by which midget RGCs could adapt to the topography of the cone mosaic.

The center regions of these cells are constructed through bipolar cell private line connectivity, which is maintained to high eccentricities [32,33]. The strength of the bipolar cell/RGC synapse could potentially be adjusted based on the activity of the presynaptic cone type [25]; however, the precise learning mechanism facilitating this weighting remains unknown. Conversely, the surround of the midget RGC is produced via the horizontal cell network's diffuse inhibition [8,34,35]. How this seemingly indiscriminate network facilitates strategic weighting is unclear, although functional biases toward one cone type have been observed in the surrounds of parvocellular neurons [5,6]. A possible explanation for this observed bias could lie outside the retinal circuitry, such as within the optics of the eye [36].

Future theoretical work includes incorporating physiologically accurate optics [14,37], modeling mosaics with varying L-to-M cone ratios, incorporating short- (S) wavelength-sensitive cones to the mosaic to model the S-(L+M) cone-opponent pathway [1,3], and including temporal information [19]. Investigating how the model allocates neurons to code the joint domain of space, time, and color will be particularly interesting.

Funding. Air Force Office of Scientific Research (FA9550-21-1-0230); Division of Information and Intelligent Systems (231-3149).

Disclosures. The authors declare no conflicts of interest.

Data availability. The calibrated natural image database underlying the results presented in this paper is available in Ref. [13]. Code is available upon request.

Supplemental document. See Supplement 1 for supporting content and figures.

REFERENCES

1. J. D. Mollon, "'tho'she kneel'd in that place where they grew…' the uses and origins of primate colour vision," J. Exp. Biol. **146**, 21–38 (1989).
2. F. De Monasterio and P. Gouras, "Functional properties of ganglion cells of the rhesus monkey retina," J. Physiol. **251**, 167–195 (1975).
3. A. M. Derrington, J. Krauskopf, and P. Lennie, "Chromatic mechanisms in lateral geniculate nucleus of macaque," J. Physiol. **357**, 241–265 (1984).
4. L. E. Wool, J. D. Crook, J. B. Troy, *et al.*, "Nonselective wiring accounts for red-green opponency in midget ganglion cells of the primate retina," J. Neurosci. **38**, 1520–1540 (2018).
5. B. B. Lee, J. Kremers, and T. Yeh, "Receptive fields of primate retinal ganglion cells studied with a novel technique," Vis. Neurosci. **15**, 161–175 (1998).
6. P. Buzás, E. M. Blessing, B. A. Szmajda, *et al.*, "Specificity of m and l cone inputs to receptive fields in the parvocellular pathway: random wiring with functional bias," J. Neurosci. **26**, 11148–11161 (2006).
7. G. D. Field, J. L. Gauthier, A. Sher, *et al.*, "Functional connectivity in the retina at the resolution of photoreceptors," Nature **467**, 673–677 (2010).
8. J. D. Crook, M. B. Manookin, O. S. Packer, *et al.*, "Horizontal cell feedback without cone type-selective inhibition mediates "red–green" color opponency in midget ganglion cells of the primate retina," J. Neurosci. **31**, 1762–1772 (2011).
9. F. Attneave, "Some informational aspects of visual perception," Psychol. Rev. **61**, 183 (1954).
10. H. B. Barlow, "Possible principles underlying the transformation of sensory messages," in *Sensory Communication* (1961), Vol. **1**, pp. 217–233.
11. J. J. Atick and A. N. Redlich, "Towards a theory of early visual processing," Neural Comput. **2**, 308–320 (1990).
12. J. J. Atick and A. N. Redlich, "What does the retina know about natural scenes?" Neural Comput. **4**, 196–210 (1992).
13. E. Doi, T. Inui, T.-W. Lee, *et al.*, "Spatiochromatic receptive field properties derived from information-theoretic analyses of cone mosaic responses to natural scenes," Neural Comput. **15**, 397–417 (2003).
14. E. Doi and M. S. Lewicki, "A theory of retinal population coding," in *Advances in Neural Information Processing Systems* (2007), Vol. **19**, p. 353.
15. Y. Karklin and E. Simoncelli, "Efficient coding of natural images with a population of noisy linear-nonlinear neurons," in *Advances in Neural Information Processing Systems* (2011), Vol. **24**.
16. S. Roy, N. Y. Jun, E. L. Davis, *et al.*, "Inter-mosaic coordination of retinal receptive fields," Nature **592**, 409–413 (2021).
17. N. Y. Jun, G. D. Field, and J. Pearson, "Scene statistics and noise determine the relative arrangement of receptive field mosaics," Proc. Natl. Acad. Sci. USA **118**, e2105115118 (2021).
18. S. Ocko, J. Lindsey, S. Ganguli, *et al.*, "The emergence of multiple retinal cell types through efficient coding of natural movies," in *Advances in Neural Information Processing Systems* (2018), Vol. **31**.
19. N. Y. Jun, G. Field, and J. Pearson, "Efficient coding, channel capacity, and the emergence of retinal mosaics," in *Advances in Neural Information Processing Systems* (2022), Vol. **35**, pp. 32311–32324.
20. Y. Singer, L. Taylor, B. D. Willmore, *et al.*, "Hierarchical temporal prediction captures motion processing along the visual pathway," Elife **12**, e52599 (2023).
21. D. L. Ruderman, T. W. Cronin, and C.-C. Chiao, "Statistics of cone responses to natural images: implications for visual coding," J. Opt. Soc. Am. A **15**, 2036–2045 (1998).
22. G. Buchsbaum and A. Gottschalk, "Trichromacy, opponent colours coding and optimum colour information transmission in the retina," Proc. R. Soc. Lond. B **220**, 89–113 (1983).
23. J. J. Atick, Z. Li, and A. N. Redlich, "Understanding retinal color coding from first principles," Neural Comput. **4**, 559–572 (1992).
24. T.-W. Lee, T. Wachtler, and T. J. Sejnowski, "Color opponency is an efficient representation of spectral properties in natural scenes," Vis. Res. **42**, 2095–2103 (2002).
25. T. Wachtler, E. Doi, T.-W. Lee, *et al.*, "Cone selectivity derived from the responses of the retinal cone mosaic to natural scenes," J. Vis. **7**(8), 6 (2007).
26. D. J. Field, "Relations between the statistics of natural images and the response properties of cortical cells," J. Opt. Soc. Am. A **4**, 2379–2394 (1987).
27. E. P. Simoncelli and B. A. Olshausen, "Natural image statistics and neural representation," Annu. Rev. Neurosci. **24**, 1193–1216 (2001).
28. H. Hofer, J. Carroll, J. Neitz, *et al.*, "Organization of the human trichromatic cone mosaic," J. Neurosci. **25**, 9669–9679 (2005).
29. R. Linsker, "Deriving receptive fields using an optimal encoding criterion," in *Advances in Neural Information Processing Systems* (1992), Vol. **5**.
30. R. Linsker, "Local synaptic learning rules suffice to maximize mutual information in a linear network," Neural Comput. **4**, 691–702 (1992).
31. T. N. Wiesel and D. H. Hubel, "Spatial and chromatic interactions in the lateral geniculate body of the rhesus monkey," J. Neurophysiol. **29**, 1115–1156 (1966).
32. A. H. Milam, D. M. Dacey, and A. M. Dizhoor, "Recoverin immunoreactivity in mammalian cone bipolar cells," Vis. Neurosci. **10**, 1–12 (1993).
33. U. Grünert, P. R. Martin, and H. Wässle, "Immunocytochemical analysis of bipolar cells in the macaque monkey retina," J. Comp. Neurol. **348**, 607–627 (1994).
34. D. M. Dacey, L. C. Diller, J. Verweij, *et al.*, "Physiology of L-and M-cone inputs to H1 horizontal cells in the primate retina," J. Opt. Soc. Am. A **17**, 589–596 (2000).
35. L. E. Wool, O. S. Packer, Q. Zaidi, *et al.*, "Connectomic identification and three-dimensional color tuning of s-off midget ganglion cells in the primate retina," J. Neurosci. **39**, 7893–7909 (2019).
36. N. Cottaris, B. Wandell, and D. Brainard, "Midget retinal ganglion cell surrounds in macaque: cone-selective or not?" J. Vis. **24**(10), 1090 (2024).
37. N. P. Cottaris, H. Jiang, X. Ding, *et al.*, "A computational-observer model of spatial contrast sensitivity: effects of wave-front-based optics, cone-mosaic structure, and inference engine," J. Vis. **19**(4), 8 (2019).