Inseparable color and form in mouse visual cortex described by random pooling from rods and cones

Rhim I.[1,3], Nauhaus I.[1,2,3]*

Affiliations:
UT Austin Department of Psychology[1], Department of Neuroscience[2], Center for Perceptual Systems[3]

***Correspondence:**
Ian Nauhaus (nauhaus@utexas.edu)
Department of Psychology
Center for Perceptual Systems
University of Texas at Austin
108 E. Dean Keeton
Austin TX, 78712

Abstract

An image projected onto the retina is composed of local contrasts in color and brightness, both of which can aid in any visual perception task. Recent investigations of the mouse's ventral retina demonstrate that rod and cone responses are combined to detect changes between UV and green light, thus providing a new model for color vision. An important question is how the spatial representations of both color and brightness contrast are transformed by downstream circuits. Its known that SF tuning of brightness contrast is sharpened at the level of mouse primary visual cortex, yet color contrast is untested. Here, we presented sinewave gratings that drive one of four axes of rod and cone contrast space, including brightness contrast (rod+cone) and color contrast (rod-cone). We find that V1 neurons are tuned to higher spatial frequencies of brightness contrast than color contrast, and are most responsive to color at the lowest spatial frequencies. These results are consistent with a model of "single-opponency" between rods and cones, but do not match its classic description. The data can instead be described by a simple model of convergent ON and OFF inputs to V1, which randomly pool discrete quantities of each photoreceptor class. Unlike classic depictions of single-opponency, this model requires minimal constraints on the circuit, accounts for our observed bandpass spatial frequency tuning of rod and cone isolating contrast, and is consistent with recent studies showing unselective pooling from photoreceptors in the retina.

Introduction

Photoreceptors in the mammalian retina are sensitive to the local amplitude and wavelength of light, and are thus able to encode local changes in brightness and color in an image. Each cone in the mouse retina expresses short (S) and/or medium (M) wavelength opsins, with peak sensitivity to 360 nm and 555 nm, respectively. The spectral sensitivity of rods is very similar to the cone M-opsin, so discrimination between different opsin contrasts (i.e. colors) requires stimulation of S-opsin. That is, spatial color gradations in an image, from UV-to-green wavelengths, can be encoded by comparing responses between the "UV opsin" (S-opsin) and either of the other two "green opsins" (rhodopsin or M-opsin). Recent studies show that UV-to-green contrast is encoded in the ventral retina via opponency between cone and rod signals [1–3]. Furthermore, rod-cone opponency is likely to account for behavioral identification of UV-green contrast [4], which is enriched in the upper visual field at twilight [5]. Motivated by these recent studies, we designed experiments to probe the interplay of color and spatial contrast in the upper visual field representation of mouse V1.

Selectivity for spatial contrast, or form, is considerably sharpened at the level of primary visual cortex (V1) [6–8]. Perhaps most notably, the V1 population exhibits a diverse population of tuning for the spatial frequency (SF) and orientation of sinewave gratings. An important and complex question is whether the exquisite tuning for spatial form in V1 is limited to brightness contrast, or if V1 also encodes form via color contrast [9]. Several studies in primate V1 have addressed this question by measuring spatio-chromatic RFs [10–13], yielding variable results. A handle on this seemingly complex problem starts with our pre-existing knowledge of V1's responses to grayscale sinewave gratings, and generalizes to a multi-dimensional color space where each grating modulates a specified combination of opsins [12]. When normalized for "total opsin contrast" along each color dimension, the method becomes even more valuable at interpreting the underlying spatial RF of each opsin [9,10]. However, this framework had not been

used to study mouse V1, so little was known about its spatial representation of color contrast, which could lead to new insights into the underlying mechanism with genetic tools.

Here, we investigated the joint representation of color and form in mouse V1, mediated by rod and cone S-opsin inputs. With two-photon imaging, we measured SF tuning along four axes of S-opsin vs. rod ("R-opsin") contrast space: S, R, R-S, and R+S. We find that V1's SF tuning is more low-pass in response to color (R-S) contrast than brightness (R+S) contrast. In addition, V1 is more strongly activated by color contrast at low SFs. These observations of SF tuning imply that the average V1 RF in our data is "single-opponent" for color, meaning that the polarity of color sensitivity (i.e. S-R vs. R-S) is the same across the RF [9]. However, classic models of single-opponency require specific connectivity between photoreceptors and the ON-OFF channels that ultimately provide input to V1. This classic model was not consistent with our data. Instead, we account for the data with a random wiring model, whereby ON and OFF subfields indiscriminately sample from each photoreceptor class. Random wiring models have been shown to account color tuning of midget cells in the primate retina [14], and may also describe much of the rod-cone opponency in mouse retinal ganglion cells [1]. At the level of V1, we show that this allows for a simple addendum to classic achromatic RF models that is consistent with many spatio-chromatic tuning properties in the mouse and primate [6].

Results

Models of rod-cone opponency

To help guide the interpretation of results, we first describe simulated spatio-chromatic RFs in V1. For each example, we start with a 1-dimensional S-opsin and rhodopsin ("R-opsin") spatial RF (Fig. 1b). From here, one can use the Fourier transform to estimate SF tuning along any axis of R- and S-opsin contrast space (Fig. 1c,d). In the first example ('Ex 1'), the S and R RFs are unimodal (e.g. Gaussian), which yields low-pass SF tuning along the S and R axes of color space. However, since the Gaussians in the spatial domain have a different width and polarity, they yield bandpass tuning along S+R, and lowpass tuning along S-R. As shown in the next example ('Ex 2'), similar SF tuning is created if the S and R RFs are the same, but shifted relative to one another. Ex 1 and 2 are more common depictions of "single-opponency" in V1, whereby ON and OFF subfields must systematically pool from either opsin. However, this is not the only way to achieve single-opponency. More generally, single-opponency is achieved when the RF of one opsin is more positive than that of another, at all positions. The third example ('Ex 3') shows a model of single opponency with bimodal S and R RFs; i.e. they each contain an ON and OFF subfield. To achieve single opponency in this case, the ON and OFF subfields are not balanced in amplitude, such that $R_{ON} < R_{OFF}$ and $S_{ON} > S_{OFF}$. Finally, we show an example of a so-called "luminance cell" ('Ex 4'). Similar to the single-opponent neuron in Ex 3, the S and R spatial RFs in Ex 4 have the same polarity and phase. However, for a luminance cell, they are scaled versions of each other, which creates separability in color and form – i.e. the SF tuning curves have the same bandpass shape along any axis of opsin contrast space, though with varying amplitude.

Our data shows that the average mouse V1 neuron is generally consistent with Ex 3, yet many also align with Ex 4. Ex1 and Ex2 are not consistent with the data. Ultimately, we show that Ex 3 and Ex 4 are typical instantiations of a random pooling model. Each instantiation is built from ON and OFF channels that independently draw from rods and cones.

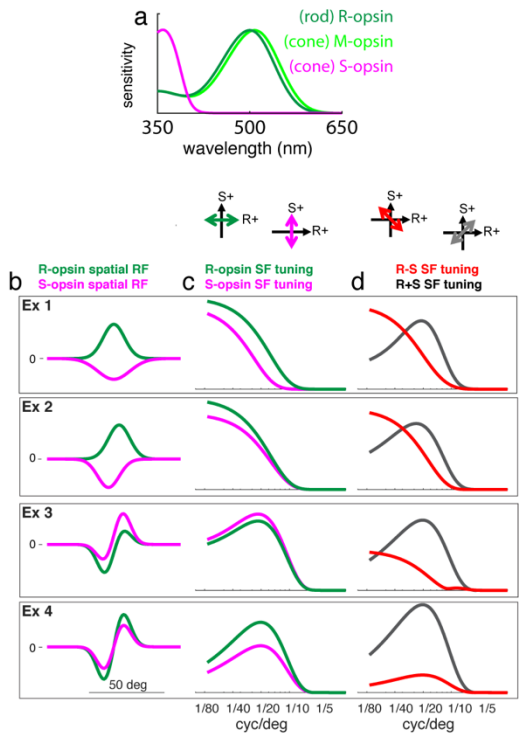

Figure 1: Examples of rod vs. S opponency. (a) Approximate spectral sensitivity functions of the three main opsins in the mouse retina. Data in this study is from the upper visual field representation of V1, which is mostly limited to input from rhodopsin ("R-opsin") and cone S-opsin, as are the simulated neurons here. Each row below shows an example spatio-chromatic RF and the predicted SF tuning along different axes of opsin contrast. (b) Each neuron (row) has two linear RFs, one for R-opsin (green) and one for S-opsin (violet). (c) Taking the Fourier transform of each opsin's RF in 'b' approximates the sensitivity of the neuron to sinewave gratings with contrast that isolates either the R- or S-opsin axis of color space. (d) To obtain the SF tuning to S+M (black) and S-M (red) contrast, we add and subtract the RFs in 'a', and then take the Fourier transform. The first three examples ('Ex1'-'Ex3') are of "single-opponency", which leads to the inseparable SF tuning. The last example shows a brightness-preferring neuron that is separable in color and SF tuning.

Paradigm for measuring rod-cone contributions to color and form tuning in V1

To build on recent studies showing rod-cone opponency in the ventral retina [1,2], we biased two-photon imaging fields-of-view to the upper visual field representation of V1 (Fig. 2a,b), while the retina was in a mesopic state of light adaptation [15]. The lack of M-opsin expression by cones in the ventral retina allowed for rod isolation, and more generally, calibrated stimuli within the R- and S-opsin contrast plane. Neuronal responses were measured with two-photon imaging of GCaMP6f in WT mice (AAV1.syn.GCaMP6f) [16].

We measured SF tuning along four different axes in the S- and R-opsin contrast plane: R, S, S-R and S+R (Fig. 2d). Total opsin contrast was the same for each axis. Figure 2e shows the responses of six example neurons, where each trace was elicited by a different opsin contrast at

the optimal SF. The other stimulus used in this study was a monochrome drifting bar (not shown) to verify the upper visual field representation of each neuron [17,18] .

A requirement for neural processing of color is simultaneous input from two types of photoreceptors. As shown in Figure 3a,b there is considerable variability in the balance of rod and cone input, as assessed by "%R". %R is the response to R contrast, divided by the summed response to S and R contrast. The variability may be due to a number of factors, such as slight changes in the state of retinal adaptation across experiments. For instance, some experiments had relatively high bias toward S-opsin input, implying that the rods were more saturated. To proceed with an assessment of color selectivity – viz. comparing responses to S-R and S+R contrast – we selected a subpopulation of cells that were balanced in their response to the S- and R- isolating contrast: $1/3 < \%R < 2/3$. Finally, the visual field representation of each neuron was verified by mapping the vertical retinotopy. A small subset of neurons with RFs below the midline were excluded (Fig. 3a,c).

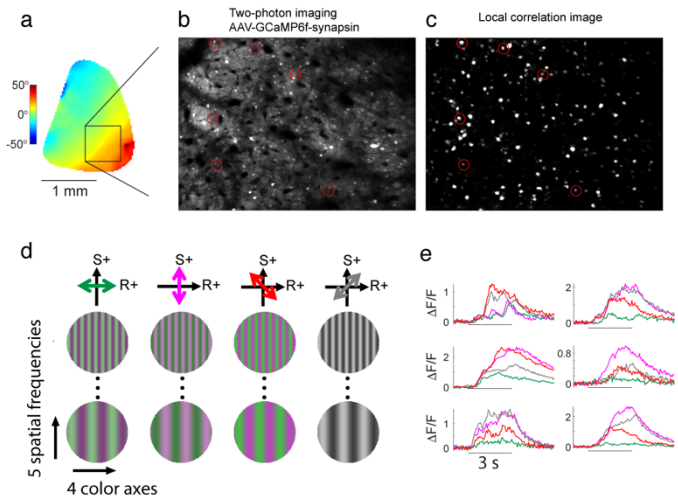

Figure 2: Two-photon imaging of responses to drifting gratings in the S and R-opsin color plane. (a) A map of vertical retinotopy in V1, using intrinsic signal imaging. Red-to-blue is upper-to-lower fields. Subsequent two-photon imaging was targeted to the upper visual field representation. (b) Raw image from a two-photon imaging session, 900 μm wide. (c) Local correlation image used to select cell bodies. Bright pixels are highly correlated with the immediate neighborhood. (d) Illustration of the stimulus set. There were five SFs, spaced 1 octave apart, ranging from 0.0125 c/° to 0.2 c/°. In a small subset of experiments, a sixth SF was included at 0.4 c/°. Above each column of gratings is a diagram of the corresponding color axis. e) Each panel shows the time courses from one of six circled neurons in 'b' and 'c'. Each color is the response to a different color axis, at the optimal drift direction and spatial frequency, averaged over 5 repeats. The color of the trace indicates the color axis, shown in 'd', top.

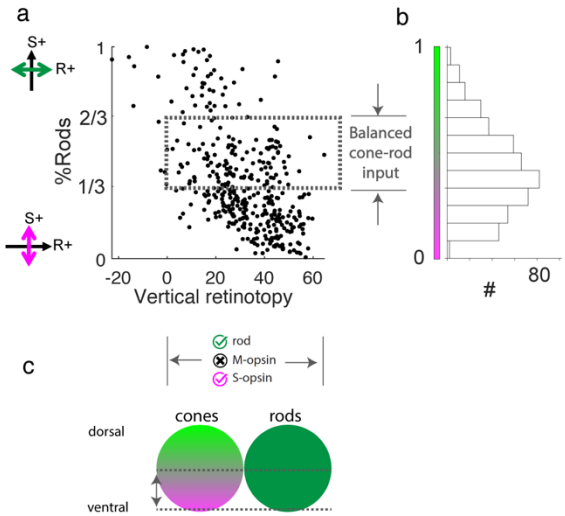

Figure 3: Data selection to allow for testing of rod-cone color opponency in V1 of WT mice (a) Each data point is a neuron that passed the criteria for SF tuning and retinotopy, described in the Methods (n = 388). The y-axis is the relative response to the S- and R-opsin isolating contrast. A value of 1 and 0 indicates that it only responded to R- and S-contrast, respectively. The x-axis is the RF's position in vertical retinotopy, with positive values being above the midline. The overlaid red rectangle outlines the population used for subsequent analyses, with the goal of examining R vs. S color opponency in V1. (b) Histogram of S vs. R selectivity, which is the marginal of the scatter plot in 'a'. The thresholds along this axis limits the analysis to neurons with relatively balanced response to the rod and cone gratings. (c) Illustration of the spatial distribution of opsins in the mouse retina. In the cone mosaic on the left, green and violet indicate concentration of M- and S-opsin, respectively. The green disk on the right represents the uniform rhodopsin ("R-opsin") distribution. Limiting analyses to the upper visual field (ventral retina) of the WT mouse excluded M-opsin, thus allowing for the study of rod-cone color opponency with gratings that contain contrast in the R- and S-opsin color plane.

A preference for color contrast at lower SFs, and brightness contrast at higher SFs

We asked whether mouse V1 neurons respond more strongly to color contrast (S-M) or brightness contrast (S+M), at each SF in the stimulus set. The population's mean response (Fig. 4a) at each SF shows a preference for color at the lowest SF, and a preference for brightness at higher SFs. To further quantify this relationship, we calculated population statistics of "preference for color" at each SF. Each neuron's preference for color is defined as the responsiveness to S-R contrast, divided by the sum of the responses to S-R and S+R contrast. That is, 0 and 1 indicates exclusive response to S+R and S-R, respectively. The distribution at each SF (Fig. 4c,d) is consistent with the plot in Figure 4a. The mean preference for color at the lowest SF is significantly greater than 0.5 (t-test ; $p < 10^{-12}$). At higher SFs, the preference for color is significantly less than 0.5 (see Fig. 4d for p-values).

Next, to summarize each neuron's preference for color with a single value, the peak of each SF tuning curve was used. For a given neuron, this can entail comparing responses between S-R and S+R contrast at different SFs. The distribution appears unimodal, with a mean of 0.48 and SD of 0.12 (Fig. 4e). We nonetheless categorized neurons as "brightness neurons" (n = 22), "color-brightness neurons" (n = 116), and "color neurons" (n = 5). This method of

categorization is used in the next section to show that preference for color is not coupled to inseparability of color and form in mouse V1.

Overall, Figure 4 shows that the mean V1 RF is consistent with a single-opponent spatial RF (e.g. Fig. 1, 'Ex 1-3'). The lowest SF in this stimulus, 1/80 cyc/deg, has a spatial period that is much wider than a mouse V1 RF [8], whereas the other SFs do not. In a model of single-opponency, this is the key criterion that leads to a preference for color. Additionally, the histogram in Figure 4e suggests that neurons in V1 are not "double-opponent" for rod and cone input. Double-opponent neurons would have strong overall preference for the color contrast [9], but no such cells exist in our data.

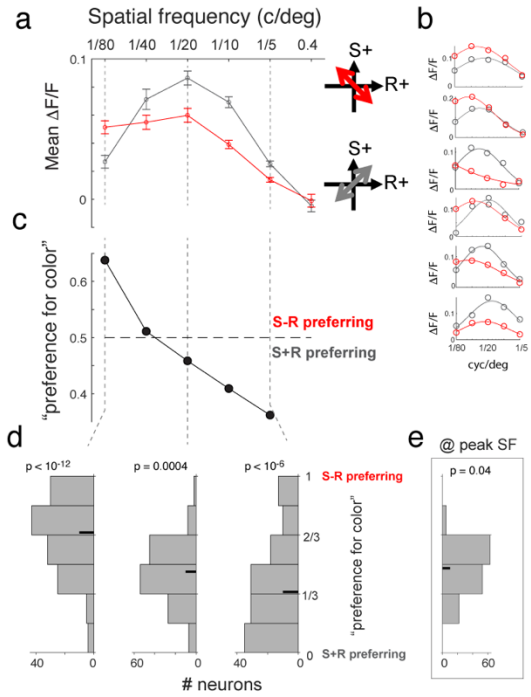

Figure 4: Preference for opponency at each SF. (a) Mean and standard error SF tuning curve to the S-R and S+R color axes (143 cells). (b) Four examples of SF tuning and Gaussian fit, for the S-R and S+R color axes. (c) Mean preference for color (S-R) over sum of brightness (S+R) and color (S-R), at each SF. (d) The distribution of preference for color, at three SFs spaced two octaves apart: 1/80, 1/20, and 1/5 cyc/deg. The black dash on the y-axis is at the median. Above each histogram is the p-value of a t-test to indicate whether the mean is significantly different than 0.5. (e) This distribution uses the peak of each SF tuning curve (S-M and S+M) to obtain a single measure of preference for color for each cell.

Neurons respond to higher SFs of brightness contrast than color contrast

Next, using the same data as in Figure 4, we compared the shape of SF tuning between brightness (S+R) and color (S-R) contrast. Whereas Figure 4 compares the magnitude of response at each SF, Figure 5 compares SF tuning parameters that are independent of response magnitude. The results in Figure 5 are nonetheless consistent, in that SF tuning for color contrast is shifted to lower values.

For each neuron, we used 4 metrics to quantify SF tuning: peak location ($peak_{S+R}$ & $peak_{S-R}$), center-of-mass (CoM_{S+R} & CoM_{S-R}), high-pass cut-off ($HPCO_{S+R}$ & $HPCO_{S-R}$), and bandpass factor (BPF_{S+R} & BPF_{S-R}). Each row in Figure 5 corresponds to a different metric, and each data point shows a neuronal measurement for both types of contrast, S+R and S-R. To determine whether a given SF tuning metric tended to be higher for color or brightness contrast, we calculated the mean-difference between pairs. For parameters with units in cyc/deg (i.e. all but 'BPF'), the log of the tuning parameters was first taken. For example, the mean of $\log(peak_{S+R}) - \log(peak_{S-R})$ was 0.43 ($p < 10^{-19}$; paired t-test), which corresponds to a 53% increase from $peak_{S-R}$ to $peak_{S+R}$. That is, 100*(exp(0.43)-1) = 53%. These summary statistics

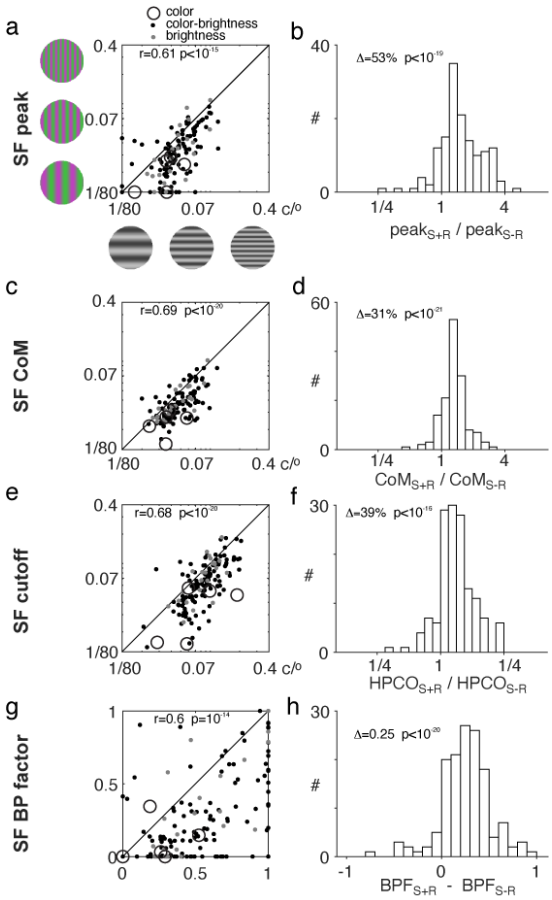

Figure 5: SF tuning of brightness contrast vs. color contrast. Each row quantifies the difference in SF tuning using a different parameter. (a) X and Y-axes show each neuron's SF tuning for brightness and color, respectively. The three symbols indicate the overall preference for color vs. brightness, as defined previously in Fig. 4e. Unity line is the diagonal. The inset statistics give the Pearson correlation coefficient and p-value. (b) The histogram is the marginal of the scatter plot in 'a', projected along the unity line. Inset gives the mean % change (Δ) of the peak, along with the t-test of significance that the mean of $\log(Peak_{S+M}/Peak_{S-M})$ is greater than zero. (c,d) Same as in 'a,b', but the SF parameter is the center-of-mass ("CoM") of the raw tuning curve. (e,f) Same as in 'a,b', but the SF parameter is the high-pass cut-off ("HPCO") from the Gaussian fit. (g,h) The SF parameter is the bandpass factor ("BPF") of the raw tuning curve. Unlike the other SF parameters in this figure, BPF is plotted on a linear axis. The inset of the histogram simply gives the mean and p-value for significance > 0.

are provided as insets in each histogram of Figure 5, all of which show that SF tuning for color contrast is significantly skewed toward lower SFs.

Lastly, Figure 5 shows a lack of modularity in color vs. form representation across the population. A classic question is whether visual neurons that are more sharply tuned for form (e.g. bandpass SF tuning) tend to be monochromatic (e.g. S+R preferring) [13]. This does not appear to be the case in mouse V1. For instance, the black data points do not cluster near the origin more than the gray data points.

SF tuning for R- and S-opsin place additional constraints on the single-opponent RF model. Results in Figure 4 and 5 provide evidence of single-opponency between rods and cones in V1. However, as shown in Figure 1, defining a neuron as "single-opponent" does not place very strong constraints on the shape of SF tuning of opsin-isolating (viz. S and R) contrast. For example, they could be low-pass (Fig. 1, 'Ex1' & 'Ex2') or bandpass (Fig. 1, 'Ex3').

The mean S and R SF tuning curves are roughly scaled versions of each other – i.e. they are separable (Fig. 6a). Furthermore, there is a clear peak shifted away from 0 cyc/deg – i.e. they are bandpass. These results are recapitulated in Figures 6b-e, where the shape of SF tuning is correlated along the unity line (separable), and the peak SF is consistently away from the origin (bandpass). It should also be noted that Fig. 6 replicates a result from our prior study, but is placed in a different context [15]. We found it useful to repeat the analysis on the same sample of neurons in Figures 4 and 5, which made it through the inclusion criteria described in Figure 3.

Another observation that helps constrain our modeling is that the SF tuning with opsin-isolating (S or R) contrast sits "between" that of S-R and S+R contrast. In other words, they are responsive to lower SFs than S+R, and higher SFs than S-R. This can be gleaned by visual

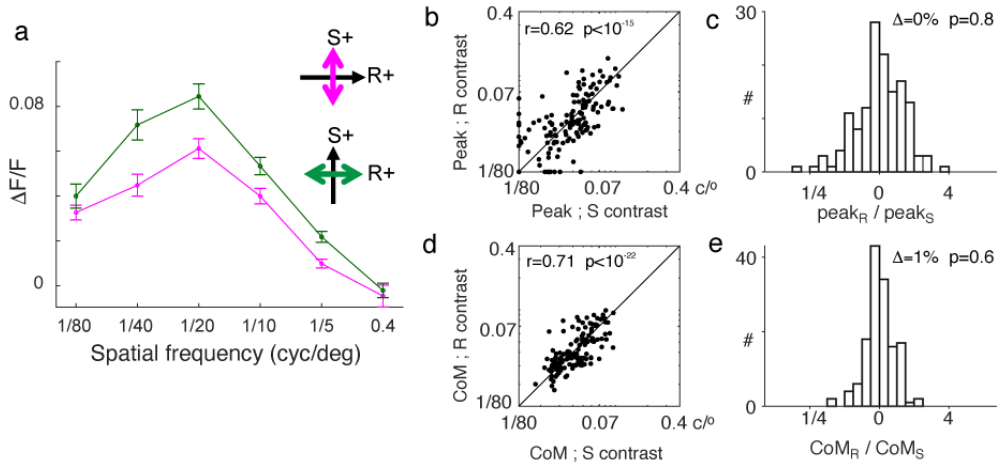

Figure 6: Summary of SF tuning for S- and R-contrast. (a) Mean and standard error SF tuning curve to the S and R color axes (143 cells). (b) X and Y-axes show each neuron's SF tuning for S and R contrast, respectively. Unity line is in black. The inset statistics give the Pearson correlation coefficient and p-value. (c) The histogram is the marginal of the scatter plot in 'b', projected along the unity line. Inset gives the mean % change (Δ) of the peak, along with the p-value (t-test) that the mean of $\log(peak_{S+M}/peak_{S-M})$ is greater than zero. (d,e) Same as in 'b,c', but the SF parameter is the center-of-mass ("CoM") of the raw tuning curve.

comparison of Figures 4a and 6a, and was quantified by comparing the distributions of peak SF. On average, the peak SF of S contrast is 29% higher than that of S-R contrast ($p < 10^{-11}$), and

18% lower than that of S+M contrast ($p < 10^{-5}$). Similarly, the peak SF of R contrast is 30% higher than S-R contrast ($p < 10^{-12}$), and 17% lower than that of S+R contrast ($p < 10^{-4}$). The p-values are based on the same paired t-test described for relating S-R vs. S+R tuning in Figure 5.

Modeling the observed inseparability of color and form

Next, we constructed a model that explains our observations of SF tuning vs. opsin contrast. We first summarize four key constraints given by the data: (1) Neurons tend to prefer brightness over color contrast at all but the lowest SFs. This constraint aligns with a model of single-opponency between R and S input, but still leaves open several possible scenarios. (2) Both S-opsin and R-opsin SF tuning curves have a peak that is shifted away from the origin (i.e. bandpass). This second constraint rules out unimodal S- and R-opsin spatial RFs, such as 'Ex1' and 'Ex2' in Figure 1, which is the more classic depiction of single opponency, even in V1. (3) There is no systematic difference between the shape of S and R SF tuning (Fig. 6), so S and R spatial RFs should be generated with the same statistics in the model. (4) A subtler, but important observation, is that the SF tuning for S and R contrast are tuned to higher SFs than S-R tuning, and lower SFs than S+R tuning. This result is consistent with an S+M spatial RF having relatively balanced ON and OFF subfields, compared to that of the individual S and M spatial RFs.

In Figure 1, 'Ex3' is consistent with the aforementioned constraints. 'Ex4' is also consistent, but only at one side of a model continuum described below. To simulate a comparable population with minimal assumptions, we start with a classic achromatic V1 model. Each V1 neuron receives input from adjacent ON and OFF lateral geniculate cells, which creates a multimodal spatial RF (and thus bandpass SF tuning curve) for brightness contrast [6,8]. Now, the difference from the achromatic model is that the ON and OFF channels each pool random amounts of S- and R-opsin. This implies that the two opsins have the same sign within each subfield, as shown in the majority of retinal ganglion cells in the ventral retina [1]. The integration of S-opsin by the ON channel creates its sensitivity (i.e. ON subfield amplitude) to S contrast, and similarly for R-opsin and R contrast. The OFF subfield is generated independently of the ON subfield, but uses the same statistics to pool from R- and S-opsin. Additional "noise" was inserted into the model for both the spatial scale of the RFs, along with Poisson-distributed responses.

A population of 250 neurons were simulated to create figure panels that mimic those in Figures 4-6. Overall, this simple model does well at capturing key features of the data. Bandpass S- and R-opsin RFs (Fig. 7a,c) are used to build lowpass tuning for color and bandpass tuning for brightness (Figs. 7b-e). At the same time, bandpass selectivity for S and R contrast is greater than S-R contrast, and less than S+R contrast (Fig. 7c). Some neurons are separable for color and SF, such as the leftmost example in Figure 7a, which is one of the data points on the unity line in Figure 7d,e.

Unlike classic depictions of single-opponent V1 neurons [9], this model does not require that ON and OFF subfields systematically pool from each type of opsin, which is more consistent with circuits in the retina [1,14,19]. Finally, we note that the model results are quite robust to an overall bias in S or R sensitivity. For example, when the simulation is re-run with all of the R-opsin RFs scaled down by a factor of 2x, such as during brighter states of adaptation, the results still look quite similar to Figure 7. The difference is a slight increase in color-form separability. Ultimately, once rods are fully saturated, opponency is gone and the tuning becomes completely separable. In the Discussion, we elaborate on the issue of how graded rod saturation affects our data and the separability of color and form.

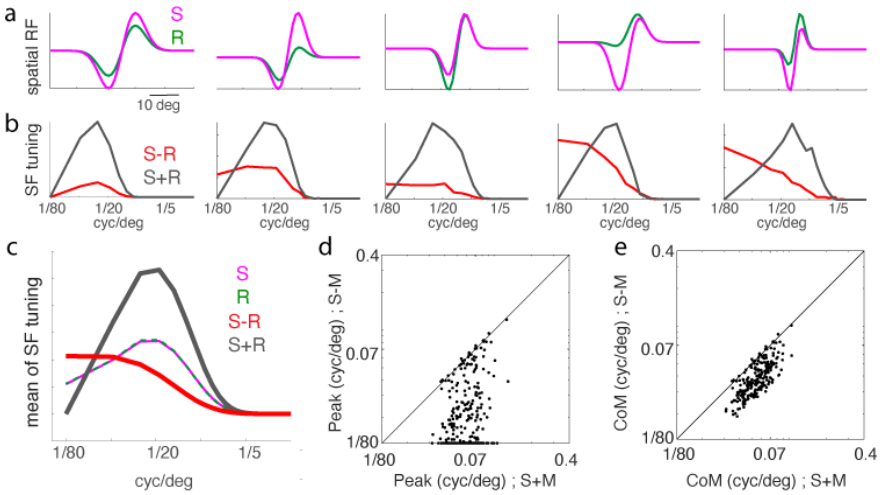

Figure 7: Modeling V1's inseparability of color and form with a random pooling model. (a) Each of the five panels shows the spatial S- an R-opsin RF of a random draw from the DoG model. They are organized from left-to-right in order of decreasing separability. The model's three independent random variables are ON amplitude-ratio (ON_S:ON_R), OFF amplitude-ratio (OFF_S:OFF_R), and spatial scale. (b) The SF tuning curves for brightness (gray) and color (red) contrast. The value at each SF is a sample from a Poisson distribution, where the mean/variance is the Fourier amplitude from the spatial RF in 'a'. (c) The mean SF tuning for each axis in our opsin contrast space, taken from 250 simulated V1 neurons. (d) Each data point gives the peak SF for brightness and color (Gaussian fit) from a neuron in the simulated population. (e) Each data point gives the SF tuning center-of-mass for brightness and color of a neuron in the simulated population.

A dependence of color-form inseparability on orientation preference predicts marginal color opponency from the cone-opsin gradient in the mesopic state

The mouse retina has a gradient of cone-opsin expression along its dorsoventral axis, with M- and S-opsin dominating the dorsal and ventral retina, respectively [20]. Recordings from retinal ganglion cells show that when center-surround RFs pool from this cone-opsin gradient, unselectively, color opponency arises [21]. What about V1 RFs, which have oriented and Gabor-like RFs? Like retinal ganglion cells, V1's color opponency from the cone-opsin gradient will be highest in the transition zone (i.e. near the horizontal midline). In addition, a V1 neuron's color opponency from this mechanism will depend on its orientation preference. Specifically, a Gabor-like RF tuned to a vertical orientation will not exhibit any opponency from co-expressing cones, regardless of its retinotopic position (Fig. 8b). However, when orientation preference is closer to

horizontal, we expect cone-opsin opponency, and thus inseparable color and form that resembles many of the RFs in our random wiring model for rod-cone opponency in Figure 7 (Fig. 8a).

As described above, opponency between M- and S-opsin from the retinotopic gradient has some clear predictions, which allowed us to check if it contributes to the color-form inseparability shown in Figures 4 and 5. The motivation behind this analysis is to gauge the impact of M-opsin on our results. Although the expression levels of M-opsin are small in the ventral retina, they may still contribute to the response to R contrast. R-opsin and M-opsin have very similar spectral sensitivity functions (Fig. 1a), so the R-opsin axis of our opsin-contrast space will drive M-opsin with similar contrast.

We first compared inseparability of color and form against vertical retinotopy. The prediction here is that M-S opponency, and thus inseparability, should be greater at lower retinotopic position which is near the midline. For the metric of inseparability, we used the change in peak SF tuning between brightness and color contrast, which was shown in Fig. 5b. There was not a significant correlation between retinotopy and inseparability ($p = 0.5$).

Next, we compared inseparability against orientation preference (Fig. 8c,d). Figure 8a shows the deviation of orientation preference against the shift in peak SF between color and brightness contrast. Interestingly, this yielded the expected trend from the cone-opsin gradient ($r = -0.25$; $p = 0.002$); neurons tuned closer to horizontal were more inseparable. However, this analysis also shows that the inseparability of color and form in our data cannot be accounted for by the cone opsin gradient - neurons tuned near vertical were still inseparable, and only slightly less so than the total population. Specifically, we took the subpopulation with a preferred orientation less than 30° from vertical ($n = 36$), and calculated the same statistics shown in the inset of Fig. 5b, which gave $\Delta = 41\%$, $p = 1.8\text{x}10^{-6}$. We performed the same analysis on the other SF tuning metrics shown in Figure 5, but orientation was only significantly correlated with change in peak SF. To summarize, the analyses reveal that there is indeed some likely opponency from the cone-opsin gradient in the mesopic regime, but it is minimal compared to the rod-cone opponency.

Opponency from co-expressing cones is quite limited in its ability to generate chromatic edge detection. The first reason is the orientation-dependence described above. In addition, this mechanism is incapable of generating "double-opponency", which is the hypothesized mechanism of spatial selectivity for color in the primate [22]. Double-opponency requires intermixing of different photoreceptors at each location of the retina.

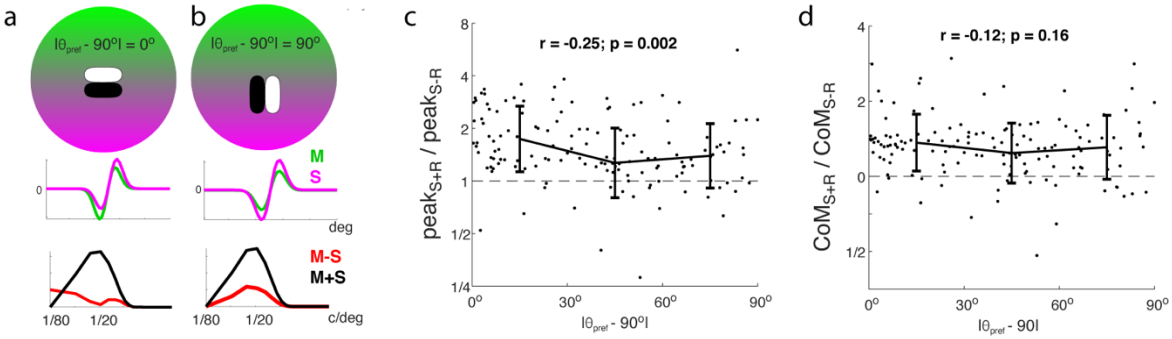

Figure 8: A dependence of inseparability on orientation preference predicts marginal opponency from the cone-opsin gradient in the mesopic state. (a) The background colored disk is a cartoon representation of the cone-opsin mosaic, where green and violet are M and S-opsin expression. A model of this gradient in retinal ganglion cell output was used (ref. Wang et al. 2011) for the simulation. Overlaid at the steepest part of the gradient is a cartoon V1 RF tuned for horizontal orientation ($\theta_{pref} = 90^o$). The RF is composed of an ON and OFF Gaussian under the same spatial constraints described in the model for Figure 7. The green and violet traces below are the 1D profiles of a simulated S and M-opsin RF, calculated by multiplying the spatial RF by the percentage of S and M-opsin expression at each location, given by the model in Wang et al. The traces on bottom are the corresponding SF tuning curves for S-M and S+M contrast. This RF is inseparable for color and form. (b) Same as 'a' but the RF prefers vertical orientation, and is positioned more ventral. This RF is separable for color and form. (c) Actual data showing color-form inseparability vs. orientation preference. y-axis is identical to Fig. 5b and contains the same data points. The x-axis is absolute deviation from horizontal orientation. Error bars show the median and SD in three bins, 0°-to-30°, 30°-to-60°, and 60°-to-90°. (d) Same as 'c', but the y-axis is from Fig. 5d.

Discussion

Building on recent work showing rod-cone opponency in the ventral portion of the mouse retina [1–3], we investigated the SF tuning of brightness and color contrast in the upper visual field representation of mouse V1. A dramatic sharpening of spatial tuning arises in V1, making it particularly relevant to the joint representation of color and form. We find that V1 neurons in the upper visual field are tuned to lower SFs of color contrast than brightness contrast, which is generally consistent with a model of single-opponency between rods and cones, as classically defined [9]. We take this an important step further to show that additional constraints imposed by our data can be described by an addendum to a classic feedforward model of V1, whereby an achromatic "Gabor-like" RF randomly pools from rods and cones. This makes it unique to classic depictions of single-opponency in the LGN or V1, which require more specific pooling of each opsin from the ON and OFF subfields (e.g. Fig. 1, 'Ex 1' & 'Ex 2'). Furthermore, a random pooling model is consistent with more recent recordings from subclasses of retinal ganglion cells in the mouse and primate [1,14,19].

Why UV-green color opponency in the mouse?

As the sun sets, irradiance at all UV-to-IR wavelengths are heavily attenuated, but longer wavelengths are attenuated the most. What remains until sunrise is a band of spectral irradiance

between roughly 300 and 550 nm [23]. The sensitivity of photoreceptors in the rodent retina cover this range quite well, thus suiting their nocturnal habits. But why would UV-visible color vision be specialized for the upper visual field? UV radiation is heavily scattered by the atmosphere and readily absorbed by most surfaces, which allows for "silhouette detection" in the upper visual field during daylight [5]. At twilight, rods are modulated by green reflection off plant matter, thus allowing improved discrimination from predators [5]. In contrast, light received in the lower half of the visual field is all reflected, and thus filtered of the ambient UV.

What about disadvantages of UV-visible sensitivity? One example is potential damage to the retina from UV. Naturally, transmittance of UV in the rodent eye is high. However, being nocturnal and having a relatively short life, any accumulated UV damage to the mouse retina is likely minimal compared to that in a larger diurnal animal [24]. Another potential problem is chromatic aberration, which will place the UV and visible contributions of an image at different focal planes. Chromatic aberration is particularly strong at these shorter wavelengths. However, there are a few things to consider. The first is that objects in the mouse's upper visual field, above ground, are effectively at "infinity". This keeps the aberration of a given object invariant to changes in its proximity, which limits potential confounds for downstream processing. Regardless of its invariance, chromatic aberration might not be a limiting factor to the overall poor resolution of the mouse's visual system. This is challenging to address definitively. For example, the maximum aberration-induced blur in the rat eye is about 60 microns, or ~1° of visual field [25], which is right around the upper limits of resolution at the level of V1 [26]. Overall, the pros seem to largely outweigh any cons of a UV-visible color system.

Comparisons with red-green opponency in the primate

Our methods borrow heavily from rigorous studies on the integration of color and form in primate V1 [10,12,27]. Here, we discuss a few of the clearer similarities and differences between the mouse and primate data.

We point to three similarities between mouse and primate V1 results on color and form tuning. First, one of the earliest color studies in primate V1 showed that color (i.e. red-green or "isoluminant") contrast yields tuning for lower SFs than brightness (i.e. "luminance") contrast [27] (e.g. Figs. 4 & 5). Next, a later study showed bandpass tuning for contrast that isolates each photoreceptor [10] (e.g. Fig. 6). Finally, similar to the mouse (Fig. 4e), neurons in primate V1 are rarely color-preferring, even when stimuli are equated for total cone contrast [10]. As a result, neurons with SF tuning that fit the classic depiction of "double-opponency" [9] have proved challenging to find in V1 of either species. The three aforementioned results in the primate are consistent with the random wiring model that we propose here. Furthermore, there is recent support that the midget pathway is built from unselective wiring of local biases in the cone mosaic. They showed that L-M cone opponency in midget cells, previously thought to be represented by a dichotomy of cell-types that receive either L or M inputs to the center RF component, should instead be modeled as a continuum of L and M inputs [14]. These observations are analogous to our rod-cone color model here.

Perhaps the most noteworthy difference from our mouse data is in a subpopulation of primate V1, called "color-luminance" cells [10]. These neurons are named for their similar response magnitude to color (L-M) and luminance (L+M) contrast. Unlike our mouse data, color-luminance cells in the primate are separable; i.e. there is not a systematic difference in SF tuning between color and luminance. Furthermore, these cells receive balanced levels of cone input with opponent phase, showing that they are not simply "imbalanced luminance cells" (e.g.

Fig. 1, Ex 4). At the same time, color-luminance cells dominate superficial layers, so their unique functionality may be due to additional feedforward and recurrent V1 processing.

Measuring single-opponency from the phase of response

We did not report the phase of neural responses, only the mean amplitude. However, a common strategy for measuring color opponency is to compare the response phase at the temporal frequency of the gratings, between gratings of different opsin contrast. Given the population's preference for color (S-R) at 1/80 cyc/deg (Fig. 4), we predict that the response to this SF has a phase differential of 180° between S-opsin and R-opsin gratings. Unfortunately, a combination of factors made a rigorous analysis of phase opponency overly challenging. For one, most F1-amplitudes at the stimulus frequency were weak, possibly due to the cells being more "complex", along with the relatively slow time course of the calcium signal. Furthermore, the PMT in our two-photon microscope was detecting subtle leakage of the green LED in the visual stimulus, which has a 180° phase shift between the R- and S-opsin gratings. Ultimately, this created an artifactual phase opponency that directly aligns with the aforementioned prediction. Importantly, the amplitude of this artifactual modulation does not affect the mean used in the rest of the analyses.

Graded discriminability between color and brightness contrast in the mesopic regime

The degree to which a mouse V1 neuron is single-opponent – and thus inseparable for SF and color – is in flux with the balance of rod and cone input induced by a given level of light adaptation. Below, we explain how a graded reduction in the balance of rod and cone inputs leads to reduced discriminability between color and brightness contrast by the V1 population. Figure 1 is used as a guide, and new statistics from our data are provided.

As light levels fall to the lower end of the mesopic range, R-opsin sensitivity (Fig. 1b,c, green amplitude) will dominate over S-opsin sensitivity (Fig. 1b,c, violet amplitude). With this imbalance, the SF tuning for S-R and S+R color axes (Fig. 1d) will begin to match the R-opsin SF curve. The results are analogous for an increase in background intensity, where the S-opsin RF dominates. In other words, the prediction is that the RF becomes more separable in color and form with greater opsin imbalance. Consistent with this, we found more separable RFs when we limited analyses to neurons with imbalanced ($1/3 > \%R > 2/3$) instead of balanced ($1/3 < \%R < 2/3$) S and R input. The balanced case was used for Figures 4-6. The more imbalanced population was still inseparable, but less so. Specifically, the change in peak SF was 29% (vs. 53% in Figure 5b), the change in center-of-mass was 18% (vs. 31% in Figure 5f), and the change in high-pass cut-off SF was 23% (39% in Figure 5f). Also, the mean S-R and S+R SF tuning curves for the imbalanced case were qualitatively much more separable than the balanced case shown in Figure 4a.

Under complete photopic or scotopic adaptation, a single opsin encodes the upper visual field, which cannot discriminate S-R and S+R contrast because separable color and spatial tuning of each neuron leads to the same pattern of population activity. Conversely, in the mesopic regime, 2 images with the same form but different color will generate a different pattern of population activity because the RFs are inseparable in color and form. Such a population code in V1 may be the cortical basis of a behavioral discrimination between color and brightness contrast [4].

In search of double-opponency and bandpass tuning for chromatic edges

Double-opponency is a hypothesized mechanism by which chromatic edge detection is encoded in the primate visual system [9,22]. The term double-opponent refers to the spatial RF model from which bandpass SF tuning for color contrast may arise. A double-opponent spatial RF is composed of adjacent single-opponent subfields that are of opposite sign – e.g. an S-R subfield that is adjacent to an R-S subfields. Therefore, double-opponent RFs in the mouse could be built in downstream areas by combining V1's single-opponent outputs [28,29], including the type consistent with our data and the random wiring model in Figure 7. We believe this warrants future theoretical and empirical investigation on the transformations of color and form beyond mouse V1, along with its behavioral correlates.

Methods

Animal preparation for surgery and imaging

All experiments were approved by the University of Texas at Austin's Institutional Animal Care and Use Committee, which maintains AAALAC accreditation. All methods were carried out in accordance with relevant guidelines and regulations. 8 male and 5 female wild-type mice (C57BL/6; aged 2-5 months) were used.

For all surgical procedures, mice were anesthetized with isoflurane (3% induction, 1-1.5% surgery), and given a pre-operative subcutaneous injection of analgesia (5mg/kg Carprofen) and anti-inflammatory agent (Dexamethasone, 1%). Mice were kept warm with a water-regulated pump pad. Each mouse underwent two surgical procedures. The first was the mounting of a metal frame over the visual cortex using dental acrylic, which allowed for mechanical stability during imaging. The second was a craniotomy (4-4.5mm in diameter) over V1 in the right hemisphere, virus injections, and a window implant (3-4mm glass window). Surgical procedures were always performed on a separate day from functional imaging. On a day between the frame implant and virus injections, we identified the outline of the V1 border by measuring the retinotopic map with intrinsic signal imaging (ISI) through the intact bone[30,31].

The virus (pAAV.Syn.GCaMP6f.WPRE.SV40, Addgene viral prep # 100837-AAV1)[32] was delivered using a Picospritzer III (Parker) or Nanoliter injector (WPI) to 2-3 sites in V1 with 0.25-0.5ul per site. Once the injections were complete, craniotomies were sealed using Vetbond and dental cement with dual-layered windows made from coverglass (4mm glued to 5mm, Warner Instruments), and covered between imaging sessions.

Common to all functional imaging procedures, isoflurane levels were titrated to maintain a lightly anesthetized state while the mouse laid on a water-circulating heat pad. However, mapping the V1 retinotopy using ISI is typically more tolerant to high isofluorane levels. For ISI, isoflurane levels were set to 0.25-0.8% (typically 0.5%). For two-photon imaging, each mouse was given chlorprothixene (1.25 mg/kg) intramuscularly and isoflurane levels were adjusted to 0.25-0.5%. Silicone oil (12,500 cst) was periodically applied to the stimulated (left) eye to prevent any dryness or damage to the eye. Silicone oil allows optical transmission of near-UV (nUV) and visible wavelengths. Additionally, the unstimulated (right) eye was coated with Vaseline and covered with black tape during imaging.

Imaging

For ISI, images were captured using a Pco Panda 4.2 camera with Matlab's image acquisition toolbox, at 10-to-20 frames/sec, and 2x2 binning. An X-Cite 110LED (Excelitas) was used for illumination. A longpass colored glass filter (590nm) was placed over the light guides to illuminate the brain, and a bandpass filter (650 +/- 25 nm) was on the camera lens.

Two-photon calcium imaging was performed with a Neurolabware microscope and Scanbox acquisition software. The scan rate varied between 10-15 frames/sec, scaling with the number of rows in the field-of-view. A Chameleon Ultra laser was set to 920 nm to excite GCaMP6f. A Nikon 16x (0.8NA 3mm WD) or Olympus 10x (0.6NA 3mm WD) objective lens was used for all imaging sessions.

Visual stimuli

Setup overview: Two monochrome LED projectors (Keynote Photonics) with spectral peaks at 405 nm (nUV) and 525 nm (green) were used to generate drifting sinewaves with contrast in the

S- and R-opsin plane (Fig. 2xx). The refresh rate was 60 Hz. The rear projector screen was made of Teflon, which provided a near-Lambertian surface[15]. One projector was aligned to the other with an affine transformation in Matlab. The transformation was identified by a user's manual selection of several points on the screen. Stimuli were coded using the Psychophysics Toolbox extension for Matlab[33,34].

The mouse was positioned so that its left eye was vertically centered on the projector screen, and the perpendicular bisector from the mouse's eye to the screen was between 8 and 10 cm. Next, the screen was angled at 30° from the mouse's midline. The final constraint was to align the front edge of the screen to the mouse's midline. The screen size was 43 cm high x 23.5 cm wide, which gives approximately 135° x 105° of visual angle.

Retinotopic mapping stimulus: The same drifting bar stimulus was used to map retinotopy for both ISI and two-photon imaging. This method is an extension of the one used in Kalatsky and Stryker [18], and described in detail in Marshel et al.[17]. To summarize, a periodic drifting bar on black background was modulated by a contrast-reversing checkerboard, and shown in four cardinal directions. The bar's speed and width were dependent on thes screen location such that the temporal phase of cortical responses, at the stimulus frequency, could be directly mapped to altitude and azimuth coordinates in a spherical coordinate system. The speed of the bar was ~5°/s and its width was ~12°. The checkerboard squares were 5° wide and reversed contrast at a rate of 2 Hz. Both green and nUV projectors were shown at maximum contrast to produce the checkerboard. Each drift direction was shown at ~0.055 HZ for wide-field imaging and ~0.042 HZ for the 2-photon calcium imaging setup. There were 2-to-3 120 sec trials, for each drift direction.

Color-SF stimulus: Full-field drifting sinewave gratings at variable SF, color, and direction were shown. The ensemble contained five spatial frequencies (0.0125, 0.025, 0.05, 0.1, 0.2 cyc/deg), 4 axes of S and R contrast space (S, R, S-R, S+R), and 8 equally-spaced drift directions. Five repeats of all combinations were shown, totaling 5x4x8x5 = 800 trials. The experiment was split in half. One half was limited to the S and M axes of color space, and the other half was limited to the S-R and S+R axes of color space. Stimuli were otherwise randomly interleaved. Each grating was shown for 3 sec, followed by 2 sec of a uniform "gray" screen that matched the mean spectral radiance of the gratings.

At least 10 minutes before recording, the pupil was fully dilated with 1% tropicamide, and the retina was adapted to the gray background of the stimuli. The gray background is estimated to yield approximately 200×10^5 photoisomerizations/rod/sec, calculated below. In a prior study, we found that this rod isomerization rate yields V1 responses that are mediated by both rods and cones [15]. As shown in Figure 2, data yield was limited to cells that receive similar response to both S- and R-opsin isolating stimuli.

Calculation of rod isomerization rate

In a recent study, we described the calculation of the rod isomerization in this same preparation [15]. We summarize these calculations here. We start with the calculation of spectral irradiance at the retina as follows:

$$I(\lambda) = T(\lambda) * Rad(\lambda) * \mathrm{A}_{pupil}/\mathrm{D}^2_{retina} \quad ; \quad \text{Retinal irradiance (Watts/ } \mu\mathrm{m}^2_{retina}/\Delta\lambda)$$

$T(\lambda)$ is transmittance through the mouse lens at each wavelength [35]. $Rad(\lambda)$ is the spectral radiance of the gray background, measured with a PR655 spectroradiometer. A_{pupil} is the diameter of the fully dilated mouse pupil, and D_{retina} is the diameter of the retina [36]. Next, we convert Joules to quanta as

$$I_Q(\lambda) = I(\lambda) * \lambda * c/h \quad ; \quad \text{photons/s/}\mu\text{m}^2/\Delta\lambda;$$

where c is the speed of light and h is Planck's constant.

To compute isomerization rate, $R^*/receptor/s$, we take the dot-product between the quantal retinal irradiance, $I_Q(\lambda)$, and the absorption spectrum, $a_c(\lambda)$, and scale by the sample period of the spectrum $\Delta\lambda$:

$$R^*/receptor/s = \Delta\lambda * \Sigma_\lambda I_Q(\lambda) a_c(\lambda) \quad ;$$

To get $a_c(\lambda)$ in μm^2, we started with the unitless absorption spectra of rods from Govardovskii et al.[37] and then scaled by the end-on collection area at the peak wavelength which may be approximated as 0.85 μm^2 for rods [38].

Calibration of cone-opsin contrast

S- and M-opsin isolating gratings: Green and nUV LED sinewaves were combined to produce contrast along one of four axes in cone-opsin space (Fig. 1f). As shown below, S-opsin and R-opsin contrast can be described as a function of the rear-projected spectral radiance of the nUV ($R_{UV}(\lambda)$) and Green ($R_G(\lambda)$) LEDs, the amplitude of their drifting sinewaves (a_{UV}, a_G), and the S- and R-opsin sensitivity functions ($h_S(\lambda)$ and $h_R(\lambda)$):

$$S\ contrast = \frac{\sum_\lambda h_S(\lambda)[a_{UV}R_{UV}(\lambda) + a_G R_G(\lambda)]}{\sum_\lambda h_S(\lambda)[R_{UV}(\lambda) + R_G(\lambda)]}$$

$$R\ contrast = \frac{\sum_\lambda h_R(\lambda)[a_{UV}R_{UV}(\lambda) + a_G R_G(\lambda)]}{\sum_\lambda h_R(\lambda)[R_{UV}(\lambda) + R_G(\lambda)]}$$

Spectral radiance was measured using a PhotoResearch PR655. Opsin sensitivity functions are taken from Govardovskii et al.[37] The above equations were used to solve for a_{UV} and a_G, for each of the four color axes in this study: S, R, S-M, and S+M. However, the total opsin contrast along each axis must first be constrained. Total opsin contrast, $\sqrt{(S\ contrast)^2 + (R\ contrast)^2}$, was set at 60% for each axis. For example, along the S-axis of color space, *S contrast* = 60% and *R contrast* = 0%. For the S+R and S-R axes, each opsin had a contrast of 43%. 100% total contrast was not used because it was not achievable in all axes. The formulae above consist of two equations and two unknowns (a_{UV} and a_G), which can be solved for each color axis.

Modeling the inseparability of color and form

Each simulated neuron in Figure 6 has a RF with an ON and an OFF subfield. The ON and OFF subfields are 1-dimensional Gaussians with the same σ. The peaks of the ON and OFF components are separated by 2σ. The peak SF, f_o , of a RF under these constraints is approximately equal to $(\pi\sigma)^{-1}$. To simulate the population of spatial RFs, we sampled from a gamma distribution of f_o, which yields the spatial RF of ON and OFF Gaussians according to $\sigma = (\pi f_o)^{-1}$. The gamma distribution of f_o was comparable to the data, with a K (i.e. "shape") of 10, and a "scale" of 0.05/K, which sets the mean to 0.05 cyc/deg.

The ON and OFF subfields were each composed of a mixture of "input" from both R-opsin and S-opsin. Specifically, each subfield sampled from a binomial distribution (n = 10, p = 0.5) to create the relative strength of input from S- and R-opsin. The two independent binomial draws, one for each subfield, gives the Gaussian amplitudes of the S-opsin subfields, $A_{S\text{-}ON}$ and $A_{S\text{-}OFF}$. Next, $A_{R\text{-}OFF} = n - A_{S\text{-}OFF}$ and $A_{R\text{-}ON} = n - A_{S\text{-}ON}$. The result is two bandpass RFs, one for each opsin. However, the subfields tend to be unbalanced in amplitude. We define the two RFs as "RF_S" and "RF_R".

To simulate SF tuning curves in response to S+R and S-R sinewave gratings, we first took the Fourier transform magnitude of $RF_S + RF_R$ and $RF_S - RF_R$, respectively. Next, the Fourier transform magnitude at each SF is used to generate a sample from a Poisson distribution in order to better represent the noise in the scatter plots (Fig. 5). Next, we quantified SF tuning with identical procedures used on the real data; viz. a Gaussian function was fit, and the peak SF (Fig. 6d) and the CoM (Fig. 6e) was taken from the fit.

Modeling spatio-chromatic RFs from the cone opsin gradient

Figure 8a,b contains examples of a model to illustrate the logic behind the data analysis and interpretation of Figure 8c,d. To generate these examples, we started with the same achromatic 1D difference-of-Gaussian (DoG) model described in the previous section. However, in this case, we only used $f_o = 0.05$ cyc/deg. To model the case of horizontal preference for orientation (Fig. 8a), the DoG was scaled by %M(v) and 100-%M(v) from Wang et. al. [39] to obtain the spatial RF of M-opsin and S-opsin, respectively. To obtain %M(v) from Wang et al., the domain was converted from millimeters on the retina into degrees of visual field ('v'). The SF tuning curves were then calculated from Fourier transforms. To model the case of vertical preference for orientation (Fig. 8b), the DoG was uniformly scaled by a coefficient for %M in the ventral retina (i.e. < 50%)

Data analysis

ISI imaging: The preferred phase at each pixel, at the frequency of the drifting bar, was calculated for each of the four cardinal drift directions. The opposing directions (e.g. up and down) were then combined to yield retinotopic position as described in Kalatsky and Styker[18]. This gave a boundary of V1 that could be manually outlined.

Generating neuronal tuning curves from two-photon movies: Neurons were identified using the local cross-correlation image, whereby the time course of each pixel was cross correlated with the weighted sum of its neighbors[15,40]. This is given as

$$LocalX(x,y) = \sum_t F(x,y,t)[F(x,y,t) \circledast DoG(x,y)]$$

where $F(x,y,t)$ are the fluorescence values at each pixel and timepoint, $DoG(x,y)$ is a difference-of-Gaussian image, and $\circledast$ indicates convolution in x and y. The DoG rewards pixels that are correlated in time with their immediate neighborhood, but also penalizes them if they are correlated in time with the broader surround. The central Gaussian of the DoG was near the size of a neuron body (σ = 3 μm), whereas the outer "suppressive" Gaussian had σ = 20 μm to capture the surrounding neuropil. The integral of each Gaussian was normalized. Groups of pixels about the size of a cell body could be more clearly disambiguated from the surround, which is important for the subsequent manual selection of cells. The puncta were manually selected using a "point-and-click" GUI. After clicking on a location, a local threshold was applied to $LocalX$ in order to identify the cell ROI. The pixels inside the ROI were then averaged to give the time course of each cell.

On each trial of the drifting grating stimulus, a neuron's raw fluorescence timecourse was converted into units of $\Delta F/F$. Specifically, $\Delta F/F = [F(t)-p]/p$, where $F(t)$ is the raw signal and p is the mean fluorescence during a blank screen preceding the trial. Gratings were presented for 3 sec. p was the average response over the preceding 1000 ms, and the mean of $\Delta F/F$ was taken between 500 ms to 3250 ms after stimulus onset. For the retinotopy stimulus, $\Delta F/F$ was not calculated since we were only interested in the phase.

Quantifying spatial frequency tuning: To get the SF tuning curve of each cell, we averaged responses over repeats, and then drift direction. Next, a Gaussian function with a domain of $\log_2$(SF) was fit in order to quantify tuning curves. The peak location, center-of-mass, high-pass cutoff, and bandpass factor were taken from the fit.

Center-of-mass was determined in the logarithmic domain as $\mathrm{CoM}_{\log} = \sum_f G(\log_2(f)) * \log_2(f) \ / \ \sum_f G(\log_2(f))$, where $G()$ is the Gaussian fit, and f is the stimulus domain in cyc/deg. $\mathrm{CoM}_{\log}$ is not necessarily the peak location of the Gaussian because the sum over f is limited to the stimulus domain. Then, $\mathrm{CoM}_{\log}$ was converted into the same units as f by taking $2^{\mathrm{CoM}_{\log}}$. Next, peak location and high-pass cutoff were both determined from a finely sampled version of the Gaussian fit. Note that the peak location differed from the parameter μ in the Gaussian fit when the tuning was low-pass. High-pass cut-off frequency was taken at 70% of the peak's height, to the right of the peak. Bandpass factor was a function of the response magnitude at the peak SF (R_{peak}), along with lowest SF of 1/80 cyc/deg ($R_{1/80}$). Bandpass factor equals $(R_{peak} - R_{1/80})/(R_{peak} + R_{1/80})$. If the peak is at 1/80 cyc/deg, the bandpass factor equals zero. If $R_{1/80} = 0$, it equals 1.

Orientation tuning: The sampling of orientation in our stimulus was spaced by 45°, which is sufficient to capture the preferred orientation, but insufficient to capture orientation bandwidth. Figure 8c,d compared color-form inseparability against the deviation of preferred orientation (θ_{pref}) from horizontal. $\theta_{pref} = \left[\angle \sum_\theta T(\theta)\, e^{i2\theta}\right]/2$, where $T(\theta)$ is the orientation tuning curve. S+R contrast was used for $T(\theta)$, and responses were averaged over repeats and SF.

Two-photon data yield: All data analysis followed an initial preprocessing pipeline for quality control. To be included in the analysis, each neuron had to exhibit tuning for retinotopy and SF.

If the variance of the SF tuning curve accounted for by the Gaussian fit was above 70%, it was used to estimate SF tuning parameters. We required that the SF tuning along each of the four axes of opsin contrast pass this criterion. 31% of the neurons passed this SF tuning criterion. For retinotopy, we used a metric for the normalized amount of Fourier energy at the stimulus frequency. Specifically, we first subtracted a linear fit to the response of each cell. Next, we calculated the Fourier magnitude at the stimulus frequency, and divided by the sum of all Fourier magnitudes. If this fraction was below 0.2, the cell was excluded. 82% of the population passed this retinotopy criterion. Exclusion from SF tuning "OR" retinotopy gave a net yield of 388 neurons, which was 29% of the population originally selected in the ROIs.

Next, we excluded based on their RF position and their response magnitude to the S- and R-opsin isolating color directions (Fig. 3). The goal was to limit analyses to neurons with exclusive and balanced input from rods and cone S-opsin (Fig. 3). To calculate the balance of rod and S-opsin input, we took the response to R contrast, and divided by the sum of responses to R and S contrast. This metric was defined as "%R" in Figure 3. Of the neurons that passed the quality control for SF tuning and retinotopy described above, 97% had a RF above the midline, and 38% had "balanced" S- and R-opsin input. 37% passed both criteria, giving a total yield of 143 neurons.

ACKNOWLEDGEMENTS

This work was supported by the Whitehall Foundation and the NIH R01EY028657.

We are grateful to Gabriela Coello-Reyes for performing surgeries and intrinsic imaging.

AUTHOR CONTRIBUTIONS

IN and IR designed the experiments. IR performed surgical preparations. IR collected the two-photon data. IN and IR analyzed the data and wrote the paper.

References

1. Szatko, K. P. *et al.* Neural circuits in the mouse retina support color vision in the upper visual field. *Nature Communications* **11**, 3481 (2020).
2. Joesch, M. & Meister, M. A neuronal circuit for colour vision based on rod-cone opponency. *Nature* **532**, 236–239 (2016).
3. Nadal-Nicolás, F. M. *et al.* True S-cones are concentrated in the ventral mouse retina and wired for color detection in the upper visual field. *eLife* **9**,.
4. Denman, D. J. *et al.* Mouse color and wavelength-specific luminance contrast sensitivity are non-uniform across visual space. *eLife* **7**, e31209 (2018).
5. Qiu, Y. *et al.* Natural environment statistics in the upper and lower visual field are reflected in mouse retinal specializations. *Current Biology* (2021) doi:10.1016/j.cub.2021.05.017.
6. Hubel, D. H. & Wiesel, T. N. Receptive fields of single neurones in the cat's striate cortex. *J Physiol* **148**, 574–591 (1959).
7. De Valois, R. L., Albrecht, D. G. & Thorell, L. G. Spatial frequency selectivity of cells in macaque visual cortex. *Vision Res.* **22**, 545–559 (1982).
8. Niell, C. M. & Stryker, M. P. Highly selective receptive fields in mouse visual cortex. *J. Neurosci.* **28**, 7520–7536 (2008).
9. Solomon, S. G. & Lennie, P. The machinery of colour vision. *Nat. Rev. Neurosci.* **8**, 276–286 (2007).
10. Johnson, E. N., Hawken, M. J. & Shapley, R. The spatial transformation of color in the primary visual cortex of the macaque monkey. *Nat Neurosci* **4**, 409–416 (2001).

681 11. Ts'o, D. Y. & Gilbert, C. D. The organization of chromatic and spatial interactions in the
682 primate striate cortex. *J. Neurosci.* **8**, 1712–1727 (1988).
683 12. Lennie, P., Krauskopf, J. & Sclar, G. Chromatic mechanisms in striate cortex of macaque. *J*
684 *Neurosci* **10**, 649–669 (1990).
685 13. Livingstone, M. S. & Hubel, D. H. Anatomy and physiology of a color system in the primate
686 visual cortex. *J. Neurosci.* **4**, 309–356 (1984).
687 14. Wool, L. E. *et al.* Nonselective Wiring Accounts for Red-Green Opponency in Midget
688 Ganglion Cells of the Primate Retina. *J. Neurosci.* **38**, 1520–1540 (2018).
689 15. Rhim, I., Coello-Reyes, G. & Nauhaus, I. Variations in photoreceptor throughput to mouse
690 visual cortex and the unique effects on tuning. *Sci Rep* **11**, 11937 (2021).
691 16. Chen, T.-W. *et al.* Ultrasensitive fluorescent proteins for imaging neuronal activity. *Nature*
692 **499**, 295–300 (2013).
693 17. Marshel, J. H., Garrett, M. E., Nauhaus, I. & Callaway, E. M. Functional specialization of
694 seven mouse visual cortical areas. *Neuron* **72**, 1040–1054 (2011).
695 18. Kalatsky, V. A. & Stryker, M. P. New paradigm for optical imaging: temporally encoded maps
696 of intrinsic signal. *Neuron* **38**, 529–545 (2003).
697 19. Field, G. D. *et al.* Functional connectivity in the retina at the resolution of photoreceptors.
698 *Nature* **467**, 673–677 (2010).
699 20. Szél, A. *et al.* Unique topographic separation of two spectral classes of cones in the mouse
700 retina. *J. Comp. Neurol.* **325**, 327–342 (1992).
701 21. Chang, L., Breuninger, T. & Euler, T. Chromatic coding from cone-type unselective circuits in
702 the mouse retina. *Neuron* **77**, 559–571 (2013).

703 22. Shapley, R. & Hawken, M. Color in the Cortex—single- and double-opponent cells. *Vision*
704 *Res* **51**, 701–717 (2011).
705 23. Spitschan, M., Aguirre, G. K., Brainard, D. H. & Sweeney, A. M. Variation of outdoor
706 illumination as a function of solar elevation and light pollution. *Sci Rep* **6**, 26756 (2016).
707 24. Gouras, P. & Ekesten, B. Why do mice have ultra-violet vision? *Experimental Eye Research*
708 **79**, 887–892 (2004).
709 25. Li, H., Liu, W. & Zhang, H. F. Investigating the influence of chromatic aberration and optical
710 illumination bandwidth on fundus imaging in rats. *JBO* **20**, 106010 (2015).
711 26. Girman, S. V., Sauvé, Y. & Lund, R. D. Receptive Field Properties of Single Neurons in Rat
712 Primary Visual Cortex. *Journal of Neurophysiology* **82**, 301–311 (1999).
713 27. Thorell, L. G., De Valois, R. L. & Albrecht, D. G. Spatial mapping of monkey V1 cells with pure
714 color and luminance stimuli. *Vision Res* **24**, 751–769 (1984).
715 28. Wang, Q. & Burkhalter, A. Area map of mouse visual cortex. *J. Comp. Neurol.* **502**, 339–357
716 (2007).
717 29. Garrett, M. E., Nauhaus, I., Marshel, J. H. & Callaway, E. M. Topography and areal
718 organization of mouse visual cortex. *J. Neurosci.* **34**, 12587–12600 (2014).
719 30. Juavinett, A. L., Nauhaus, I., Garrett, M. E., Zhuang, J. & Callaway, E. M. Automated
720 identification of mouse visual areas with intrinsic signal imaging. *Nat Protoc* **12**, 32–43
721 (2017).
722 31. Rhim, I., Coello-Reyes, G., Ko, H.-K. & Nauhaus, I. Maps of cone opsin input to mouse V1
723 and higher visual areas. *J. Neurophysiol.* jn.00849.2016 (2017) doi:10.1152/jn.00849.2016.

32. Chen, T.-W. *et al.* Ultrasensitive fluorescent proteins for imaging neuronal activity. *Nature* **499**, 295–300 (2013).

33. Brainard, D. H. The Psychophysics Toolbox. *Spat Vis* **10**, 433–436 (1997).

34. Pelli, D. G. The VideoToolbox software for visual psychophysics: transforming numbers into movies. *Spat Vis* **10**, 437–442 (1997).

35. Lei, B. & Yao, G. Spectral attenuation of the mouse, rat, pig and human lenses from wavelengths 360nm to 1020nm. *Experimental Eye Research* **83**, 610–614 (2006).

36. Remtulla, S. & Hallett, P. E. A schematic eye for the mouse, and comparisons with the rat. *Vision Res.* **25**, 21–31 (1985).

37. Govardovskii, V. I., Fyhrquist, N., Reuter, T., Kuzmin, D. G. & Donner, K. In search of the visual pigment template. *Vis. Neurosci.* **17**, 509–528 (2000).

38. Naarendorp, F. *et al.* Dark light, rod saturation, and the absolute and incremental sensitivity of mouse cone vision. *J. Neurosci.* **30**, 12495–12507 (2010).

39. Wang, Y. V., Weick, M. & Demb, J. B. Spectral and temporal sensitivity of cone-mediated responses in mouse retinal ganglion cells. *J. Neurosci.* **31**, 7670–7681 (2011).

40. Smith, S. L. & Häusser, M. Parallel processing of visual space by neighboring neurons in mouse visual cortex. *Nat Neurosci* **13**, 1144–1149 (2010).