bioRxiv preprint doi: https://doi.org/10.1101/2021.04.08.438953; this version posted July 29, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Natural Image Statistics for Mouse Vision

Authors: Luca Abballe[1] and Hiroki Asari[2*]

Affiliations:

1. Department of Biomedical Engineering, Sapienza University of Rome, Rome RM, 00185, Italy

2. Epigenetics and Neurobiology Unit, EMBL Rome, European Molecular Biology Laboratory, Monterotondo RM, 00015, Italy

∗**Corresponding author:** Hiroki Asari, E-mail: asari@embl.it

Acknowledgments: This work was supported by research grants from EMBL (H.A.). We thank all the members of the Asari lab at EMBL Rome for many useful discussions.

Author contributions: L.A. collected and processed the image data sets. H.A. designed the study, analyzed the results, and wrote the manuscript.

Competing interests: The authors declare no competing financial interests.

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.08.438953; this version posted July 29, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Abstract

The mouse has dichromatic color vision based on two different types of opsins: short (S)- and middle (M)-wavelength-sensitive opsins with peak sensitivity to ultraviolet (UV; 360 nm) and green light (508 nm), respectively. In the mouse retina, cone photoreceptors that predominantly express the S-opsin are more sensitive to contrasts and denser towards the ventral retina, preferentially sampling the upper part of the visual field. In contrast, the expression of the M-opsin gradually increases towards the dorsal retina that encodes the lower visual field. Such a distinctive retinal organization is assumed to arise from a selective pressure in evolution to efficiently encode the natural scenes. However, natural image statistics of UV light remain largely unexplored. Here we developed a multi-spectral camera to acquire high-quality UV and green images of the same natural scenes and examined the first- and second-order image statistics. We found that the local contrast and the spatial correlation were both higher in UV than in green for images above the horizon, but lower in UV than in green for those below the horizon. This suggests that the dorsovental functional division of the mouse retina is not optimal for maximizing the bandwidth of information transmission. Factors besides the coding efficiency, such as visual behavioral requirements, will thus need to be considered to fully explain the characteristic organization of the mouse retina.

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.08.438953; this version posted July 29, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Introduction

Sensory systems have been considered to be adapted to the statistical properties of the environment through evolution [1]. Animals encounter different types of sensory signals depending on their natural habitats and lifestyles, and this can serve as an evolutionary driving force for each species to optimize its sensory systems for processing those signals that appear more frequently and are relevant for survival [2]. The optimality of the sensory processing has been broadly supported from an information theoretic viewpoint of coding efficiency [3, 4]. In particular, various physiological properties of sensory neurons can be successfully derived from learning efficient codes of natural images or natural sounds, such as separation of retinal outputs into ON and OFF channels [5], Gabor-like receptive fields of visual cortical neurons [6], and cochlear filter banks [7]. Such computational theories and statistical models are, however, often limited to generic features of the sensory processing, and fail to account for species-specific fine details partly due to a lack of proper data sets of natural sensory signals.

In the past decade, the mouse has become a dominant model for studying the visual system mainly because of the wide availability of experimental tools [8]. Compared to other mammalian model animals such as cats and primates, however, the mouse vision has certain distinctive properties. For example, mice are dichromats as many other mammals are, but their retina expresses ultraviolet (UV)-sensitive short (S)-wavelength sensitive opsins and green-sensitive middle (M)-wavelength sensitive opsins [9–11]. While UV vision is common in amphibians, birds and insects, it has not been identified in mammals except for a few species including rodents [12–14]. Moreover, the mouse retina has no fovea but a prominent dorsoventral gradient in the expression pattern of the two opsins [10, 15–17]. A vast majority of the mouse cone photoreceptors ($\sim$95%) co-express the two opsins but with a dominant expression of S- and M-opsins in the ventral and dorsal parts of the retina, respectively [9, 10, 18, 19]. This makes the upper visual field more sensitive to UV than green, and vice versa for the lower visual field [20]. It is natural to assume that this functional segregation of the mouse vision has evolved due to an adaptation to the natural light distribution as the sunlight is the major source of UV radiation. It remains unclear, though, how optimal the mouse visual system is to natural scene statistics *per se*.

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.08.438953; this version posted July 29, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

While natural image statistics have been extensively studied thus far [1, 21], those outside the spectral domain of human vision remain to be fully explored [2, 18, 22]. Here we thus developed a multi-spectral camera system to sample high-quality images that spectrally match the mouse photopic vision, and analyzed the first- and second-order statistics of the UV and green image data sets. We identified distinct statistical properties in the UV and green channels between the upper and lower visual-field images; however, these image statistics were not necessarily consistent with what the efficient coding hypothesis would predict from the functional organization of the mouse retina.

Materials and Methods

All data and codes are available upon request.

Multi-spectral camera

Design

We built a multi-spectral camera system based on a beam-splitting strategy [23, 24] to acquire images of the same scenes with ultraviolet (UV)- and green-transmitting channels that match the spectral sensitivity of the mouse photopic vision (Fig 1A) [9–11]. The light coming from a commercial camera lens (Nikon, AF Nikkor 50 mm f/1.8D) was collimated with a near-UV achromatic lens (effective focal length, 50 mm; Edmund Optics, 65-976) and split with a dichroic filter (409 nm; Edmund Optics, 34-725). The reflected light, on the one hand, passed through a UV-selective filter set (HOYA U-340 and short-pass filter at 500 nm; Edmund Optics, 84-708) and formed the UV images focused on the first global-shutter camera (Imaging Source, DMK23UX174) with a near-UV achromatic lens (effective focal length, 50 mm; Edmund Optics, 65-976). The transmitted light, on the other hand, passed through a band-pass filter (500±40 nm; Edmund Optics, 65-743) and a lens (Edmund Optics, 65-976), and formed the green images sampled by the second camera (Imaging Source, DMK23UX174). To maximize the dynamic range of the two camera sensors (used with the same settings), we attenuated

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.08.438953; this version posted July 29, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

the light intensity of the green channel using an absorptive neutral density (ND) filter (optical density: 1.0, 1.3, 1.5, 1.8, or 2.0) on a filter wheel (Thorlabs, LTFW6) because the sunlight has much higher power in green than in UV (Fig 1B). The optical components are all mounted with standard light-tight optomechanical components (Thorlabs, 1-inch diameter lens tubes).

A recent study employed a similar design but with a fisheye lens to study the "mouse-view" images [22]. Our design has the following advantages to sample high-quality images suitable for image statistics analysis. First, we chose a small field of view (11.3 degrees horizontally and 7.3 degrees vertically; 0.006 degrees/pixel) to minimize image distortion, and a large field of depth (the smallest aperture size on the Nikon lens, $f/22$) to maximize areas in focus. This also allowed us to adjust camera settings (exposure length) to fully capture the dynamic range of individual scenes. Second, we chose a high-performance camera sensor (Sony, IMX174 complementary metal-oxide-semiconductor; CMOS) that has high quantum efficiency ($\sim 30\%$ at 365 nm; $\sim 75\%$ at 510 nm), high dynamic range (73 dB; 12 bit depth), high pixel resolution (1920-by-1200 pixels), and linear response dynamics (Fig 1A, inset) [25–27]. Third, our design has fewer optical elements to ensure transmission of near UV light (Fig 1B; see below spectral analysis).

Spectral analysis

The spectral sensitivity of the multi-spectral camera system (Fig 1B) was calculated by convolving the relative transmission spectra of the optics for each channel with the spectral sensitivity of the camera sensor (Sony, IMX174 CMOS) [27]. The relative transmission spectra were measured with a spectrometer (Thorlabs, CCS200/M; 200–1000 nm range) by taking the ratio of the spectra of a clear sunny sky (indirect sunlight) with and without passing through the camera optics.

For a comparison, we modelled the spectral sensitivity of the mouse visual system by convolving the transmission spectra of the mouse eye [28] with the absorption spectra of the mouse cone photoreceptors (Fig 1B). We used a visual pigment template [29] with the center frequency at 360 nm and 508 nm to simulate the short (S)- and middle (M)-wavelength-sensitive opsins in the mouse retina, respectively [9–11].

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.08.438953; this version posted July 29, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Image acquisition

In total, we collected 232 images of natural scenes without any artificial object in the suburbs of Lazio/Abruzzo regions in Italy from July 2020 to May 2021. All the images were acquired using a custom-code in Matlab (Image Acquisition Toolbox) without any image correction, such as gain, contrast, or gamma adjustment. The two cameras were set with the same parameter values adjusted to each scene, such as the exposure length, and a proper ND filter was chosen for the green channel so that virtually all the pixels were within the dynamic range of the camera sensors (see examples in S2 Fig). Thus, our image data sets have no underexposed pixels and only a negligible number of overexposed pixels (0.0011% of pixels in 2 UV images and 0.0007% of pixels in 6 Green images). This is critical because the presence of under- or over-exposed pixels will skew the image statistics.

When acquiring images, the camera system was placed on the ground to follow the viewpoint of mice. The following meta-data were also recorded upon image acquisition: date, time, optical density of ND filter in the green channel, weather condition (sunny; cloudy), distance to target object (short, within a few meters; medium, within tens of meters; or long), presence/absence of specific objects (animals; plants; water), and camera elevation angle (looking up; horizontal; looking down). We also took a uniform image of a clear sunny sky (indirect sunlight) as a reference image for vignetting correction (see below Eq.(1)).

All the images were taken under ample natural light during the day. Although we did not measure the exact illuminance Φ of the environment, we expect that the lighting condition was on the order of 10^3–10^5 lux (i.e., $\Phi = 10^7$–10^9 photons/μm^2/s). Assuming the mouse pupil diameter $d_{\text{pupil}} = 0.5$ mm, the eye diameter $d_{\text{eye}} = 2$ mm, the transmittance of the eye optics $T = 0.5$, and the light collection area of a photoreceptor $A_{\text{photoreceptor}} = 0.5\ \mu\text{m}^2$, the photon flux on individual photoreceptors can then be estimated as $\Phi \cdot A_{\text{pupil}}/A_{\text{retina}} \cdot T \cdot A_{\text{photoreceptor}} = 10^4$–$10^6$ photons/photoreceptor/s, where $A_{\text{pupil}} = \pi \left(d_{\text{pupil}}/2\right)^2$ is the pupil area and $A_{\text{retina}} = 4\pi \left(d_{\text{eye}}/2\right)^2/2$ is the total area of the retina internally covering a half of the eye. Here we cannot then exclude a possible activation of rods in the mouse retina because they have similar absorption spectra to the M-opsin expressing cones (peak sensitivity at 498 and 508 nm, respectively) [9, 30] and may escape from saturation even at 10^7 R*/rod/s [31]. However, the rod

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.08.438953; this version posted July 29, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

system is likely optimized to work in the scotopic condition, and thus less affected by the natural image statistics in the photopic condition. In the mouse retina, rods are indeed distributed more densely ($\sim$97% of all photoreceptors) and rather uniformly [32].

Image registration

The raw images from the two cameras (12 bit depth saved in the 16 bit grayscale Portable Network Graphic format, 1920-by-1200 pixels each) were pre-processed to form a registered image in Matlab (Image Processing Toolbox). First, we corrected the optical vignetting by normalizing the pixel intensity of the raw image $I_{\text{raw}}(x, y)$ for each channel by the ratio of the pixel and the maximum intensities of the reference image $I_{\text{ref}}(x, y)$:

$$I_{\text{corrected}}(x, y) = I_{\text{raw}}(x, y) \cdot \frac{\max\left[I_{\text{ref}}(x, y)\right]}{I_{\text{ref}}(x, y)}. \tag{1}$$

We next applied a two-dimensional median filter (3-by-3 pixel size) to remove salt-and-pepper noise from the corrected images for each channel. Then we applied a projective transformation based on manually selected control points to register the UV image to the green image. Finally, we manually cropped the two images to select only those areas in focus. The cropped images resulted in the pixel size ranging from 341 to 1766 pixels (2.0–10.6 degrees) in the horizontal axis and from 341 to 1120 pixels (2.0–6.7 degrees) in the vertical axes (see examples in Fig 2). We never changed the image resolution.

Image analysis

We analyzed the first- and second-order image statistics of the obtained natural scenes in UV and green channels. Here we excluded a small set of the horizontal images ($N = 15$) from the analysis, and focused on the following two major image groups: 1) looking-up images taken with a positive camera elevation angle ($N = 100$), presumably falling in the ventral retina and thus perceived in the upper part of an animal's visual-field; and 2) looking-down images with a negative camera elevation angle ($N = 117$) perceived in the lower visual field (i.e., the dorsal retina). To ensure the separation between the image categories, we calculated the relative light

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.08.438953; this version posted July 29, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

intensity along the horizontal and vertical axes of each image category (S1 Fig). Specifically, we first corrected the pixel values of each image with the exposure length and the ND filter attenuation, and then normalized them by the mean pixel intensity value of all images. For the population analysis, the images were then aligned to the center in horizontal axes for all images, while to the top edge, center, or bottom edge in vertical axes for the lower, horizontal, upper visual field image categories, respectively. For each image data set, we used a sign-test to compare the image statistics parameter values between the UV and green channels (Figs 3–6; significance level, 0.05). All image analysis was done in Matlab (Mathworks).

Normalized light intensity

The visual system adapts its sensitivity to the range of light intensities in each environment [33, 34]. We thus first normalized the pixel intensity of each UV and green image to have the intensity value ranging from zero to one (by subtracting the minimum value of the image, followed by the division by the maximum value), and then calculated the histogram (bin size, 0.01) to compare the normalized intensity distributions of the UV and green images for the upper and lower visual fields (Fig 3A,B).

Local contrast

To calculate the local statistical structure of the normalized intensity images (Fig 3C,D and S3 Fig), we used the second-derivative (Laplacian) of a two-dimensional Gaussian filter:

$$\mathrm{LoG}(x,y) = \frac{1}{\pi\sigma^4}\left(1 - \frac{x^2+y^2}{2\sigma^2}\right)\exp\left[-\frac{x^2+y^2}{2\sigma^2}\right], \tag{2}$$

with the standard deviation $\sigma = 5, 10, 20, 40$ pixels for the spatial range $x, y \in [-3\sigma, 3\sigma]$. Here we chose a rather arbitrary size of the filter width (0.18-–1.44 degrees) because natural image statistics are scale invariant (S3 Fig) [1, 21]. The local contrast distribution was then fitted to the two-parameter Weibull distribution:

$$w(x) = \beta\gamma|x|^{\gamma-1}\exp\left[-\beta|x|^{\gamma}\right], \tag{3}$$

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.08.438953; this version posted July 29, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

where x is the local contrast value, $\beta > 0$ is the scale parameter (width) of the distribution, and $\gamma > 0$ is the shape parameter (peakedness). In particular, larger β and smaller γ values indicate wider and more heavy-tailed distributions, respectively, hence higher contrast in the images. Sign-tests were used to compare these parameter values between UV and green images (Fig 3E–H).

Achromatic and chromatic contrast

To analyze the achromatic contrast of our image data sets (Fig 4), we calculated the root mean square (RMS) contrast $\mathrm{C}^*_{\mathrm{RMS}}(x, y)$ for each channel of normalized intensity images [22]:

$$\mathrm{C}^*_{\mathrm{RMS}}(x, y) = \frac{\sigma^*(x, y)}{\mu^*(x, y)}, \tag{4}$$

where $\mu^*(x, y)$ and $\sigma^*(x, y)$ are the mean and standard deviation of a circular image patch (radius, 30 pixels) centered at location (x, y), respectively; and the asterisk "$*$" is either "UV" or "Green" indicating the channel identity (Fig 4A,B). Chromatic contrast $\mathrm{C}(x, y)$ was then defined as a difference of the RMS contrasts between the two channels (Fig 4C,D):

$$\mathrm{C}(x, y) = \mathrm{C}^{\mathrm{UV}}_{\mathrm{RMS}}(x, y) - \mathrm{C}^{\mathrm{Green}}_{\mathrm{RMS}}(x, y). \tag{5}$$

For quantification, we fitted the Weibull distribution (Eq.(3)) to the left ($\mathrm{C} < 0$) and right ($\mathrm{C} > 0$) sides of the chromatic contrast distributions separately (Fig 4E,F).

Power spectral density

The power spectral density of the normalized intensity image $I(x, y)$ was computed with the fast Fourier transform (FFT; Fig 5):

$$F(\omega_x, \omega_y) = \mathrm{FFT}\left[I(x, y)\right] \tag{6}$$

$$S(\omega_x, \omega_y) = F(\omega_x, \omega_y)F^*(\omega_x, \omega_y), \tag{7}$$

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.08.438953; this version posted July 29, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

where the superscript $*$ denotes complex conjugate, and ω_x and ω_y represent the horizontal and vertical spacial frequency (ranging from -0.5 to 0.5 cycles/pixel), respectively. As the average power spectrum of natural images generally falls with a form $1/f^{\alpha}$ over the spatial frequency f with a slope $\alpha \sim 2$ [1, 35, 36], we fitted the power function b/ω^{a} to $S(\omega_x, 0)$ and $S(0, \omega_y)$, where a and b indicate the slope and Y-intercept in the log-log space. We used a sign-test to compare these parameter values between UV and green channels (Fig 5I–P).

Spatial autocorrelation

Following the Wiener–Khinchin theorem, the spatial autocorrelation $R(x, y)$ was computed with the inverse FFT of $S(\omega_x, \omega_y)$ in Eq.(7):

$$R(x, y) = \text{IFFT}\left[S(\omega_x, \omega_y)\right], \tag{8}$$

where x and y represent horizontal and vertical distances of the two pixel points in the target image, respectively (Fig 6). Sign-tests were used to compare the $R(d_{\text{h}}, d_{\text{v}})$ values at representative data points: $[d_{\text{h}}, d_{\text{v}}] = [0, 50], [50, 0]$ (Fig 6I–L).

Results

Multi-spectral camera for the mouse vision

The mouse retina expresses short (S)- and middle (M)-wavelength sensitive opsins that are maximally sensitive to ultraviolet (UV; ∼360 nm) and green (∼508 nm) wavelengths of light, respectively [9–11]. Existing public databases of natural scenes contain a diverse set of images including both natural and artificial objects in both gray and color scales visible to humans, but do not cover any UV images [e.g., 37–40]. To examine the natural image statistics of the mouse vision, we thus set out to build a multi-spectral camera system for acquiring images of the same scenes in both UV and green spectral domains (Fig 1).

We first modelled the spectral sensitivity of the mouse dichromatic vision to determine the center wavelengths of the two channels. Because the lens and cornea absorb shorter wavelength

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.08.438953; this version posted July 29, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

light (e.g., UV rays) more than longer wavelength light, we corrected the absorption spectra of the mouse cone photoreceptors [29] with the transmission spectra of the whole eye optics [28]. This resulted in a slight shift of the center wavelengths to a longer wavelength by several nanometers: from ∼360 nm to ∼365 nm for the S-cone and from ∼508 nm to ∼512 nm for the M-cone (Fig 1B). Thus, the ocular transmittance had only minor effects on the spectral sensitivity of the mouse vision, reassuring its sensitivity to near-UV light [20, 41].

We then designed a multi-spectral camera system accordingly using a beam-splitting strategy [Fig 1A; see Methods for specifications; 23, 24]. By convolving the measured transmission spectrum of the camera optics with the sensitivity spectrum of the camera sensors [27], we identified that our imaging device had the sensitivities to ∼368±10 nm and ∼500±30 nm (center wavelength ± half-width at half maximum; HWHM) for the UV and green channels, respectively (Fig 1B). This confirms that the UV and green channels of our device were spectrally well isolated, and that the two channels largely matched to the spectral sensitivity of the mouse vision [9–11].

Ultraviolet and green image collection

To collect images that mice would encounter in their natural habitats, we went out to natural fields and wild forests in the countryside and mountain area of Lazio/Abruzzo regions in Italy across different seasons. We placed the multi-spectral camera on the ground at about a height of the mouse eye, and acquired images of natural objects alone at various distances (e.g., clouds, trees, flowers, and animals), excluding any artificial objects. These images were taken with different camera angles in the presence of ample natural light (S1 Fig). The images were preprocessed to correct optical vignetting and remove salt-and-pepper noise, and cropped to exclude areas out of focus on the edges (see Methods for details). This led to a set of 232 pairs of UV and green images of various “mouse-view” natural scenes.

Besides well-known facts that UV light is reflected well by open water and some plants [13, 14], we noticed several distinct features between the UV and green images (see examples in Fig 2). First, clouds often appeared dark and faint in the UV images than in the green ones. In some cases, even negative contrast was formed for the clouds in UV while positive contrast

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.08.438953; this version posted July 29, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

in green. Second, fine textures were more visible in the green images than in the UV ones. In particular, objects in the upper field UV images were often dark in a nearly uniform manner due to back-light, whereas fine details of the objects were nevertheless visible in the corresponding green images despite a high contrast against the sky. For the lower field images, in contrast, distinct brighter spots stood out in UV due to reflections of shiny leaves and cortices, while more shades and shadows were visible in green. These qualitative observations already suggest that the UV and green images have distinct statistical properties.

Normalized intensity and contrast distributions of UV and green images

To analyze the image statistics more formally, we first calculated the normalized intensity distribution of the UV and green channels for the upper and lower visual-field images (Fig 3A,B). Because the visual system adapts its sensitivity to the range of light intensities in each environment [33, 34], we normalized the pixel intensity of each UV and green image to be within the range from zero to unity. We then found that, for the upper visual-field images, the probability distributions of both UV and green intensity values were bimodal (Fig 3A). The two peaks of the UV intensity distribution, however, were higher and more separated than those of the green intensity distribution, suggesting that luminance contrast is higher in UV than in green when animals look up. In contrast, the normalized intensity distributions of the lower field images were unimodal and skewed to the right for both color channels. The distribution was more strongly heavy-tailed for the green than for the UV images (Fig 3B), indicating higher contrast in green than in UV when animals look down.

To better examine the contrast in the two different spectral domains, we calculated the local image contrast using the second derivative (Laplacian) of a two-dimensional Gaussian filter (Eq.(2) in Methods). This filter follows the antagonistic center-surround receptive fields of early visual neurons [e.g., retinal ganglion cells; 42, 43] that are sensitive to local contrast, and is commonly used for edge detection in computer vision [44–46]. Consistent with what was implicated by the intensity distributions (Fig 3A,B), we found that 1) the probability distribution of local contrast was generally wider for the upper visual-field images than for the lower visual field images; and 2) the local contrast distribution was wider for the upper

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.08.438953; this version posted July 29, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

visual-field UV images than for the corresponding green images (Fig 3C and S3A,C,E Fig), but narrower for the lower visual-field UV images than for the green counterparts (Fig 3D and S3B,D,F Fig). To quantify these differences, we fitted a two-parameter Weibull function (Eq.(3) in Methods) to the local contrast distribution of each image in each channel [47, 48], where the first scale parameter (β) describes the width of the distribution, hence a larger value indicating higher contrast; and the second shape parameter (γ) relates to the peakedness, with a smaller value indicating a heavier tail and thus higher contrast in the image. For the images above the horizon, the UV channel had significantly smaller shape parameter values than the green channel (Fig 3G) with comparable scale parameter values (Fig 3E). In contrast, for the images below the horizon, the green channel had significantly larger scale parameter values than the UV channel (Fig 3F), with no difference in the shape parameter values (Fig 3H). Thus the image statistics showed distinct characteristics between the upper and lower visual-field image data sets, with higher contrast in UV than in green for the upper visual-field images, and vice versa for the lower visual-field images.

Importantly, such differences in the local contrast distributions do not agree well with what the efficient coding hypothesis implies from the physiological and anatomical properties of the mouse retina [3, 4]. Solely from an information theoretic viewpoint, a narrower contrast distribution is better encoded with a more sensitive cone type to maximize its bandwidth [49]. In the mouse retina, the functional S-cones are more sensitive to contrast than the functional M-cones [17–20, 41]; and the functional S-cones are denser towards the ventral part of the retina, preferentially sampling the upper part of the visual field, while the functional M-cones towards the dorsal retina, sampling the lower visual field [15, 16, 18]. Therefore, this particular retinal organization is optimal if the upper visual-field images had lower contrast in UV than in green, and the lower visual-field images had higher contrast in UV than in green. Our image analysis, however, showed the opposite trend in the “mouse-view” visual scenes (Fig 3).

Achromatic and chromatic contrast of “mouse-view” images

To examine achromatic and chromatic contrast of our image data sets, we next measured the root mean square (RMS) contrast (Eqs.(4) and (5) in Methods) that is commonly used in

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.08.438953; this version posted July 29, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

psychophysical studies [22]. We found that the achromatic RMS contrast (Eq.(4)) was higher in UV than in green channels, especially for the upper visual field images (Fig 4A,B). The upper visual field images then had an asymmetric chromatic contrast distribution (Eq.(5); Fig 4C), where pixels with higher contrast in UV than in green were more abundant than those with higher contrast in green than in UV (Fig 4E,F). In contrast, the chromatic contrast distribution was rather symmetric for the lower visual field images (Fig 4D), and it was overall wider than that for the upper visual field images (Fig 4E,F).

This indicates that UV-green chromatic information exists across the visual field, even though the exact shape of the chromatic contrast distribution may depend on the image contents [22]. We indeed identified UV-green chromatic objects in both lower and upper visual field images (see examples in Fig 2 and S2 Fig) and thus cannot explain why the mouse retina has chromatic circuitry preferentially on the ventral side (upper visual field) [50–52]. In principle, mice could retrieve UV-green chromatic information across the visual field, given that 1) genuine S-cones and rods are distributed rather uniformly across the mouse retina [32]; 2) rods have similar absorption spectra to M-cones (peak sensitivity at 498 and 508 nm, respectively; Fig 1B) [9, 30]; and 3) rods can escape from saturation even under photopic conditions [31]. Larger image datasets sampled under more diverse conditions are required to assess the optimality of the chromatic circuitry in the mouse retina, especially because the rod system plays a role not only in the color vision but also in the scotopic vision.

Power spectrum and autocorrelation of UV and green images

We next analyzed the second-order statistics of the acquired images. Specifically, we computed the power spectrum (Fig 5) and spatial autocorrelation that describes the relationship of the two pixel intensity values as a function of their relative locations in the images (Fig 6; see Methods for details). As expected [1, 21], the power spectra generally followed $1/\omega^a$ on the spatial frequency ω for both UV and green channels irrespective of the camera angles (in log-log axes; Fig 5A–H); and were higher for the vertical direction than for the horizontal direction (Fig 5A–H)—i.e., the spatial autocorrelation was elongated in the vertical direction (Fig 6A–D).

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.08.438953; this version posted July 29, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

There are, however, several distinct properties between the UV and green channels for the upper and lower visual-field images. First, the slope of the power spectra a was larger for the lower visual-field images than for the upper visual-field images (Fig 5I–L); equivalently, the spatial autocorrelation was narrower for the lower visual-field images (Fig 6E–H), indicating the presence of more fine textures in those images. Second, for the upper visual-field images, the UV power spectra were higher than the green ones in both vertical and horizontal directions (e.g., the Y-intercept b, indicating the log-power at the spatial frequency of 1 cycle/pixel; Fig 5M,N). In contrast, for the lower visual-field images, the UV power spectra were lower with a larger slope than the green counterparts (Fig 5K,L,O,P). Equivalently, the spatial autocorrelation was wider in UV than in green for the upper visual-field images, and vice versa for the lower visual-field images (Fig 6E–H).

Under an efficient coding hypothesis, a higher spatial autocorrelation implies that less cones are needed to faithfully encode the scenes [3, 4, 49]. One would then expect from the "mouse-view" image statistics that the functional S- and M-cones should be denser on the dorsal and ventral parts of the mouse retina, respectively, to achieve an optimal sampling. However, the opposite is the case with the mouse retina [15, 16, 18], suggesting that the cone distribution bias in the mouse retina cannot be simply explained by the optimality principle from an information theoretic viewpoint.

Discussion

To study the natural image statistics for the mouse vision, here we collected a set of 232 "mouse-view" two-color images of various natural scenes across different seasons using a custom-made multi-spectral camera (Figs 1 and 2). We identified distinct properties in the first- and second-order image statistics for the two channels between the images above and below the horizon (Figs 3–6). Specifically, both the local contrast and the spatial autocorrelation were higher in UV than in green for the upper visual-field images, while they were both lower in UV than in green for the lower visual-field images. This disagrees with what the efficient coding hypothesis implies [3, 4] from the functional division of the mouse retina along the dorsoventral axis [15,

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.08.438953; this version posted July 29, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

16, 18]. We thus suggest that the given retinal organization in mice should have evolved not only to efficiently encode natural scenes from an information theoretic perspective, but likely to meet some other ethological demands in their specific visual environments [22].

How faithful are our images to what mice actually see in their natural habitats? This is a critical question because image statistics depend on the quality and contents of the images. Our camera system was designed to collect high-quality UV-green images (Figs 1 and 2) comparable to the existing natural image datasets for human vision [37–40]. However, caveats include that 1) the effects of the mouse eye optics were not considered in the image acquisition or analysis; 2) no motion dynamics were considered; 3) images were taken under ample light during the day, while mice are nocturnal; and 4) our image datasets were still relatively small. It is a future challenge to address these questions, for example, by measuring the properties of the mouse eye optics, simulating images projected onto the mouse retina, and analyzing the statistics of these images.

"Mouse-view" natural image database

We employed a beam-splitting strategy to simultaneously acquire UV and green images of the same scenes (Fig 1) because it has certain advantages over other hyper- or multi-spectral imaging techniques [23, 24]. First, a previous study used a hyperspectral scanning technique where a full spectrum of each point in space was measured by a spectrometer [18]. While the photoreceptor response could be better estimated by using its absorption spectra, the scanned images through a pinhole aperture inevitably had lower spatial and temporal resolutions than the snapshot images acquired with our device. Second, a camera array can be used for multi-spectral imaging with each camera equipped with appropriate filters and lenses [53]. This is easy to implement and will perform well for distant objects; however, because angular disparity becomes larger for objects at a shorter distance, one would have a difficulty in taking close-up images that small animals such as mice would normally encounter in their everyday lives. Finally, our single-lens-two-camera design is simple and cost-effective compared to other snapshot spectral imaging methods [24]. In particular, commercially available devices are often

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.08.438953; this version posted July 29, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

expensive and inflexible, hence not suitable for our application to collect images that spectrally match the mouse vision.

There are several conceivable directions to expand the "mouse-view" natural image database. First, we could take high dynamic range images using a series of different exposure times. This works only for static objects, but can be useful to collect images at night during which nocturnal animals such as mice are most active. Second, we could take a movie to analyze the space-time statistics of natural scenes [22]. It would be interesting to miniaturize the device and mount it on an animal's head to collect time-lapse images with more natural self-motion dynamics [54, 55]. Expanding our "mouse-view" natural image datasets will be critical to better understand the visual environment of mice and develop a theoretical explanation on species-specific and non-specific properties of the mouse visual system.

Optimality of the mouse retina

What selective pressures have driven the mouse retina to favour UV sensitivity over blue and evolve the dorsovental gradient in the opsin expression? Our image analysis suggests that the coding efficiency alone with respect to the natural image statistics cannot fully explain the distinctive organization of the mouse retina (Figs 3–6). For example, we argued from an information theoretic viewpoint that, for equalizing the bandwidth within the system, high contrast images in the upper visual field (Fig 3C) should be encoded with less sensitive photoreceptors (M-cones), while low contrast images in the lower visual field (Fig 3D) with more sensitive photoreceptors (S-cones) [18]. In contrast, one could also argue from an ethological viewpoint that more sensitive S-cones should be driven more strongly by high contrast images in the upper visual field and thus better suited to process biologically relevant information, such as aerial predators [2, 22].

To understand in what sense the moues retina's organizations are optimal, one then needs to clarify visual ethological demands that are directly relevant for survival and reproduction. For example, fresh mouse urine reflects UV very well, and this has been suggested to serve as a con-specific visual cue for their territories and trails besides an olfactory cue [56]. The UV sensitivity can also be advantageous for the hunting behavior of mice because many nocturnal insects are

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.08.438953; this version posted July 29, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

attracted to UV light. Furthermore, increased UV sensitivity in the ventral retina may improve the detection of tiny dark spots in the sky, such as aerial predators [57]. Indeed, the S-opsin-dominant cones in mice have higher sensitivity to dark contrasts than the M-opsin-dominant ones [18], and turning the anatomical M-cones into the functional S-cone by co-expressing the S-opsin will dramatically increase the spatial resolution in the UV channel because the mouse retina has only a small fraction of the uniformly distributed genuine S-cones (∼5 %) compared to the co-expressing cones [∼95 %; 11, 16, 17, 58].

These arguments, however, are difficult to generalize because each species has presumably taken its own strategy to increase the fitness in its natural habitat, leading to convergent and divergent evolution. On the one hand, UV sensitivity was identified in some mammals that live in a different visual environment than mice, including diurnal small animals such as the degu and gerbil [56, 59, 60] and even large animals such as the Arctic reindeer [61]. On the other hand, some species showing a similar behavioral pattern as mice do not have the dorsoventral division of the retinal function [12–14]. For example, even within the genus *Mus*, some species do not have the dorsoventral gradient of the S-opsin expression, and others completely lack the S-cones [62]. It is even possible that the cone distribution bias may have nothing to do with the perception of the color vision, but may arise just because of the developmental processes. Indeed, the center of the human fovea is generally devoid of S-cones [63, 64], and there is a huge diversity in the ratio of M- and L-cones in the human retina across subjects with normal color vision [65, 66]. Behavioral tests across species will then be critical for validating the ethological arguments to better understand the structure and function of the visual system [2]. We expect that the "mouse-view" natural image datasets will contribute to designing such studies.

Figure legends

Fig 1: Multi-spectral camera system for the mouse vision

(A) Schematic diagram of the camera optics. Incoming light was split into UV and Green channels by a dichroic mirror and further filtered to match the spectral sensitivity of the mouse

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.08.438953; this version posted July 29, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

visual system (see panel B). A neutral density filter with the optical density value from 1.0 to 2.0 was used for the Green channel to maximize the dynamic range of the camera sensor to be used with the same parameter settings as the UV channel. The inset shows the pixel intensity values as a function of the exposure time (mean $\pm$ standard deviation; N=2,304,000 pixels), supporting the linearity of the camera sensor (Sony, IMX174 CMOS).

(B) Relative spectral sensitivity of the camera system (UV channel, violet area; Green channel, green area). For comparison, the spectral sensitivity of the mouse rod and S- and M-cone photoreceptors [29] corrected with the transmission spectrum of the mouse eye optics [28] was shown in black, violet and green lines, respectively, as well as typical sunlight spectrum in gray.

Fig 2: Representative images of the natural scenes in UV and green channels

See S2 Fig for the UV-Green pixel intensity distribution of these example images.

(A) Upper visual field images taken with positive camera elevation angles (UV, Green, and pseudo-color merged images from left to right). These images typically contain trees and branches with sky backgrounds.

(B) Lower visual field images taken with negative camera elevation angles, often containing a closer look of grasses and flowers.

Fig 3: Light intensity and local contrast distributions of the "mouse-view" natural images

(A,B) Normalized light intensity distributions of the upper (A) and lower (B) visual field images for UV (violet) and Green (green) channels (median and interquartile range).

(C,D) Local contrast distributions computed with the Laplacian-of-Gaussian filter ($\sigma = 10$ in Eq.(2); see S3 Fig for contrast distributions computed with different σ values). The distribution

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.08.438953; this version posted July 29, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

of the UV channel is more strongly heavy-tailed than that of the Green channel for the upper visual field images (C), but the Green channel's distribution is wider than the UV channel's for the lower visual field images (D).

(E–H) Scale (β; E,F) and shape (γ; G,H) parameters from the Weibull distribution fitted to each image (Eq.(3); see Methods for details). For the upper field images (E,G), the UV channel has significantly smaller γ (G) but comparable β (E) values than the Green channel. In contrast, for the lower field images (F,H), the Green channel has significantly larger β (F) but comparable γ (H) values than the UV channel. P-values are obtained from sign-tests.

Fig 4: Achromatic and chromatic contrast of the "mouse-view" natural images

(A,B) Root mean square (RMS) contrast of the upper (A) and lower (B) field images, computed independently for the UV (violet) and Green (green) channels of each image (local patch size, 30 pixel radius; Eq.(4) in Methods). The UV channel has higher achromatic contrast, especially for the upper visual field images (median $\pm$ interquartile range).

(C,D) Chromatic contrast distributions (median $\pm$ interquartile range) computed as a difference of the RMS contrasts between the UV and Green channels (Eq.(5) in Methods). The distribution was asymmetric for the upper field images (C) but rather symmetric for the lower field images (D).

(E,F) Scale (β; E) and shape (γ; F) parameters from the Weibull distribution (Eq.(3)) fitted to each side of the chromatic contrast distribution of each image. The box plot shows the median $\pm$ interquartile range. The upper field images contain fewer pixels that have higher contrast in Green than in UV (rank-sum test: three stars "$\star\star\star$" indicating $p < 0.001$; $\star\star$, $p < 0.01$; and $\star$, $p < 0.05$).

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.08.438953; this version posted July 29, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Fig 5: Power spectrum of the "mouse-view" natural images

(A–D) The average power spectra of the upper (A,B) and lower (C,D) visual-field images for the UV (A,C) and Green (B,D) channels.

(E–H) The power spectra in the vertical (E,G) and horizontal (F,H) directions (median and interquartile range) for the upper (E,F) and lower (G,H) visual-field images.

(I–M) The slope (a; I–L) and Y-intercept (b; M–P) parameters of the power function b/ω^a in the log-log space fitted to the power spectra of each image in the vertical (I,K,M,O) and horizontal (J,L,N,P) directions. For the upper visual-field images (I,J,M,N), the UV channel has significantly larger b (M,N) but comparable a (I,J) values than the Green channel. For the lower field images (K,L,O,P), in contrast, the Green channel has significantly larger b (O,P) and smaller a (K,L) values than the UV channel. P-values are obtained from sign-tests.

Fig 6: Spatial autocorrelation of the "mouse-view" natural images

(A–D) The average spatial autocorrelation of the upper (A,B) and lower (C,D) visual-field images for the UV (A,C) and Green (B,D) channels, respectively.

(E–H) The spatial autocorrelation in the vertical (E,G) and horizontal (F,H) directions (median and interquartile range). The UV channel has a higher and wider spatial correlation for the upper visual-field images (E,F), while the Green channel has a higher and wider spatial correlation for the lower visual-field images (G,H).

(I–L) Representative spatial correlation values of the pixels horizontally (I,K) or vertically (J,L) separated by 50 pixels for the upper (I,J) and lower (K,L) visual-field images. P-values were obtained from sign-tests.

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.08.438953; this version posted July 29, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Supporting information

S1 Fig: Relative pixel intensities along horizontal and vertical axes

Relative pixel intensities (median $\pm$ interquartile range; UV and green channels in violet and green, respectively) were computed along horizontal (A,C,E) and vertical (B,D,F) axes for three different image categories based on the camera angle: lower (A,B; $N = 117$), horizontal (C,D; $N = 15$), and upper (E,F; $N = 100$) visual field images. Pixel intensity did not change much horizontally but was generally lower in the lower field images (A,B) than in the upper field images (E,F). Discontinuity between the top edge of the lower field images (B, X-axis value of 0) and the bottom edge of the upper field images (F, X-axis value of 0) supports a good separation of the two image categories.

S2 Fig: UV-Green pixel intensity distributions of representative "mouse-view" images

Each scatter plot shows the distribution of the UV-Green pixel values from the corresponding image shown in Fig 2 (A, upper visual field images; B, lower visual field images). Virtually all pixels were within the dynamic range of the camera sensor (Sony, IMX174 CMOS; 12-bit depth saved in a 16-bit format).

S3 Fig: Local contrast distributions of the natural scenes are scale invariant

Local contrast distributions computed with different Laplacian-of-Gaussian filter sizes (A,B, $\sigma = 5$; C,D, $\sigma = 20$; E,F, $\sigma = 40$; Eq.(2)) are shown in the same format as Fig 3C,D ($\sigma = 10$). The upper visual field images (A,C,D) generally showed higher contrast than the lower visual field images (B,D,F), especially for the UV channel (violet). The filter size (0.18–1.44 degrees) used in this study is smaller than the receptive field size of mouse retinal ganglion cells (3-13 degrees) [67, 68]. Given the scale invariance [1, 21], however, we expect that our analysis results should hold for larger filters as well [22].

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.08.438953; this version posted July 29, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

References

[1] Simoncelli EP, Olshausen BA. Natural image statistics and neural representation. Annu Rev Neurosci. 2001;24:1193–1216.

[2] Baden T, Euler T, Berens P. Understanding the retinal basis of vision across species. Nat Rev Neurosci. 2020;21:5–20.

[3] Attneave F. Some informational aspects of visual perception. Psychol Rev. 1954;61:183–193.

[4] Barlow HB. Possible principles underlying the transformation of sensory messages. In: Rosenblith WA, editor. Sensory Communication. MIT Press; 1961. p. 217–234.

[5] Gjorgjieva J, Sompolinsky H, Meister M. Benefits of pathway splitting in sensory coding. J Neurosci. 2014;34:12127–12144.

[6] Olshausen BA, Field DJ. Emergence of simple-cell receptive field properties by learning a sparse code for natural images. Nature. 1996;381:607–609.

[7] Smith E, Lewicki M. Efficient auditory coding. Nature. 2006;439:978–982.

[8] Huberman AD, Niell CM. What can mice tell us about how vision works? Trends Neurosci. 2011;34:464–473.

[9] Jacobs GH, Neitz J, Deegan II JF. Retinal receptors in rodents maximally sensitive to ultraviolet light. Nature. 1991;353:655–656.

[10] Lyubarsky A, Falsini B, Pennesi M, Valentini P, Pugh Jr E. UV- and midwave-sensitive cone-driven retinal responses of the mouse: a possible phenotype for coexpression of cone photopigments. J Neurosci. 1999;19:442–455.

[11] Nikonov S, Kholodenko R, Lem J, Pugh Jr E. Physiological features of the S-and M-cone photoreceptors of wild-type mice from single-cell recordings. J Gen Physiol. 2006;127:359–374.

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.08.438953; this version posted July 29, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

[12] Peichl L. Diversity of mammalian photoreceptor properties: adaptations to habitat and lifestyle? Anat Rec A Discov Mol Cell Evol Biol. 2005;287:1001–1012.

[13] Cronin TW, Bok MJ. Photoreception and vision in the ultraviolet. J Exp Biol. 2016;219:2790–2801.

[14] Hunt DM, Wilkie SE, Bowmaker JK, Poopalasundaram S. Vision in the ultraviolet. Cell Mol Life Sci. 2001;58:1583–1598.

[15] Szél A, Röhlich P, Caffé AR, Juliusson B, Aguirre G, Van Veen T. Unique topographic separation of two spectral classes of cones in the mouse retina. J Comp Neurol. 1992;325:327–342.

[16] Applebury ML, Antoch MP, Baxter LC, Chun LL, Falk JD, Farhangfar F, et al. Cone photoreceptor: a single cone type expresses both S and M opsins with retinal spatial patterning. Neuron. 2000;27:513–523.

[17] Röhlich P, Van Veen T, Szél A. Two different visual pigments in one retinal cone cell. Neuron. 1994;13:1159–1166.

[18] Baden T, Schubert T, Chang L, Wei T, Zaichuk M, Wissinger B, et al. A tale of two retinal domains: Near-optimal sampling of achromatic contrasts in natural scenes through asymmetric photoreceptor distribution. Neuron. 2013;80:1206–1217.

[19] Wang YV, Weick M, Demb JB. Spectral and Temporal Sensitivity of Cone-Mediated Responses in Mouse Retinal Ganglion Cells. J Neurosci. 2011;31:7670–7681.

[20] Jacobs GH, Williams GA. Contributions of the mouse UV photopigment to the ERG and to vision. Doc Ophthalmol. 2007;115:137–144.

[21] Hyvärinen A, Hurri J, Hoyer PO. Natural Image Statistics: A Probabilistic Approach to Early Computational Vision. Springer-Verlag London; 2009.

[22] Qiu Y, Zhao Z, Klindt D, Kautzky M, Szatko KP, Schaeffel F, et al. Mouse retinal specializations reflect knowledge of natural environment statistics. Curr Biol. 2021;31:1–15.

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.08.438953; this version posted July 29, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

[23] Grahn H, Geladi P. Techniques and Applications of Hyperspectral Image Analysis. John Wiley and Sons; 2007.

[24] Hagen N, Kudenov MW. Review of snapshot spectral imaging technologies. Opt Eng. 2013;52:090901.

[25] Jähne B, Herrmann H. EMVA 1288 datasheet Basler acA1920-155um, 0 dB gain; 2015. Available from: `https://doi.org/10.5281/zenodo.33040`.

[26] Jähne B, Herrmann H. EMVA 1288 datasheet Basler acA1920-155um, 12 dB gain; 2015. Available from: `https://doi.org/10.5281/zenodo.33041`.

[27] FLIR White Paper: Sony Pregius Global Shutter CMOS Imaging Performance; 2015 (accessed 28-July-2021). Available from: `https://www.flir.com/globalassets/support/iis/whitepaper/sony-pregius-global-shutter-cmos-imaging-performance.pdf`.

[28] Henriksson JT, Bergmanson JPG, Walsh JE. Ultraviolet radiation transmittance of the mouse eye and its individual media components. Exp Eye Res. 2010;90:382–387.

[29] Govardovskii VI, Fyhrquist N, Reuter T, Kuzmin DG, Donner K. In search of the visual pigment template. Vis Neurosci. 2000;17:509–528.

[30] Bridges CDB. Visual Pigments of Some Common Laboratory Mammals. Nature. 1959;184:1727–1728.

[31] Tikidji-Hamburyan A, Reinhard K, Storchi R, Dietter J, Seitter H, Davis KE, et al. Rods progressively escape saturation to drive visual responses in daylight conditions. Nat Commun. 2017;8:1813.

[32] Jeon CJ, Strettoi E, Masland RH. The Major Cell Populations of the Mouse Retina. J Neurosci. 1998;18:8936–8946.

[33] Shapley R, Enroth-Cugell C. Visual adaptation and retinal gain controls. Prog Retinal Res. 1984;3:263–346.

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.08.438953; this version posted July 29, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

[34] Gollisch T, Meister M. Eye Smarter than Scientists Believed: Neural Computations in Circuits of the Retina. Neuron. 2010;65:150–164.

[35] Burton GJ, Moorhead IR. Color and spatial structure in natural scenes. Appl Opt. 1987;26:157–170.

[36] Tolhurst DJ, Tadmor Y, Tang C. The amplitude spectra of natural images. Ophthalmic Physiol Opt. 1992;12:229–232.

[37] van Hateren JH, van der Schaaf A. Independent Component Filters of Natural Images Compared with Simple Cells in Primary Visual Cortex. Proc Biol Sci. 1998;265:359–366.

[38] Geisler WS, Perry JS. Statistics for optimal point prediction in natural images. J Vis. 2011;11:14.

[39] Tkačik G, Garrigan P, Ratliff C, Milčinski G, Klein JM, Seyfarth LH, et al. Natural Images from the Birthplace of the Human Eye. PLOS ONE. 2011;6:1–12.

[40] Zhou B, Lapedriza A, Xiao J, Torralba A, Oliva A. Learning Deep Features for Scene Recognition using Places Database. In: Ghahramani Z, Welling M, Cortes C, Lawrence N, Weinberger KQ, editors. Advances in Neural Information Processing Systems. vol. 27. Curran Associates, Inc.; 2014.

[41] Jacobs GH, Williams GA, Fenwick JA. Influence of cone pigment coexpression on spectral sensitivity and color vision in the mouse. Vision Res. 2004;44:1615–1622.

[42] Rodieck RW, Stone J. Analysis of Receptive Fields of Cat Retinal Ganlion Cells. J Neurophysiol. 1965;28:833–849.

[43] Enroth-Cugell C, Robson JG. The contrast sensitivity of retinal ganglion cells of the cat. J Physiol. 1966;187:517–552.

[44] Marr D, Hildreth E. Theory of edge detection. Proc R Soc Lond Ser B. 1980;207:187–217.

[45] Ghosh K, Sarkar S, Bhaumik K. Understanding image structure from a new multi-scale representation of higher order derivative filters. Image Vis Comput. 2007;25:1228–1238.

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.08.438953; this version posted July 29, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

[46] Lindeberg T. A computational theory of visual receptive fields. Biol Cybern. 2013;107:589–635.

[47] Geusebroek JM, Smeulders AWM. A six-stimulus theory for stochastic texture. Int J Comput Vis. 2005;62:7–16.

[48] Gijsenij A, Gevers T. Color Constancy Using Natural Image Statistics and Scene Semantics. IEEE Trans Pattern Anal Mach Intell. 2011;33:687–698.

[49] Laughlin S. A simple coding procedure enhances a neuron's information capacity. Z Naturforsch C Biosci. 1981;36:910–912.

[50] Joesch M, Meister M. A neuronal circuit for colour vision based on rod–cone opponency. Nature. 2016;532:236–239.

[51] Nadal-Nicolás FM, Kunze VP, Ball JM, Peng BT, Krishnan A, Zhou G, et al. True S-cones are concentrated in the ventral mouse retina and wired for color detection in the upper visual field. eLife. 2020;9:e56840.

[52] Szatko KP, Korympidou MM, Ran Y, Berens P, Dalkara D, Schubert T, et al. Neural circuits in the mouse retina support color vision in the upper visual field. Nat Commun. 2020;11:3481.

[53] Genser N, Seiler J, Kaup A. Camera Array for Multi-Spectral Imaging. IEEE Trans Image Process. 2020;29:9234–9249.

[54] Meyer AF, Poort J, O'Keefe J, Sahani M, Linden JF. A head-mounted camera system integrates detailed behavioral monitoring with multichannel electrophysiology in freely moving mice. Neuron. 2018;100:46–60.e7.

[55] Sattler NJ, Wehr M. A Head-Mounted Multi-Camera System for Electrophysiology and Behavior in Freely-Moving Mice. Front Neurosci. 2021;24:592417.

[56] Chàvez AE, Bozinovic F, Peichl L, Palacios AG. Retinal spectral sensitivity, fur coloration and urine reflectance in the genus *Octodon* (Rodentia): implications for visual ecology. Invest Ophthalmol Vis Sci. 2003;44:2290–2296.

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.08.438953; this version posted July 29, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

[57] Yilmaz M, Meister M. Rapid innate defensive responses of mice to looming visual stimuli. Curr Biol. 2013;23:2011–2015.

[58] Haverkamp S, Wässle H, Duebel J, Kuner T, Augustine GJ, Feng G, et al. The primordial, blue-cone color system of the mouse retina. J Neurosci. 2005;25:5438–5445.

[59] Govardovskii VI, Röhlich P, Szél A, Khokhlova TV. Cones in the retina of the Mongolian gerbil, *Meriones unguiculatus*: an immunocytochemical and electrophysiological study. Vision Research. 1992;32:19–27.

[60] Jacobs GH, Calderone JB, Fenwick JA, Krogh K, Williams GA. Visual adaptations in a diurnal rodent, *Octodon degus*. J Comp Physiol A. 2003;189:347–361.

[61] Hogg C, Neveu M, Stokkan KA, Folkow L, Cottrill P, Douglas R, et al. Arctic reindeer extend their visual range into the ultraviolet. J Exp Biol. 2003;214:2014–2019.

[62] Szél A, Röhlich P, Caffé AR, Van Veen T. Distribution of cone photoreceptors in the mammalian retina. Microsc Res Techn. 1996;35:445–462.

[63] Wald G. Blue-Blindness in the Normal Fovea. J Opt Soc Am. 1967;57:1289–1301.

[64] Curcio CA, Allen KA, Sloan KR, Lerea CL, Hurley JB, Klock IB, et al. Distribution and morphology of human cone photoreceptors stained with anti-blue opsin. J Comp Neurol. 1991;312:610–624.

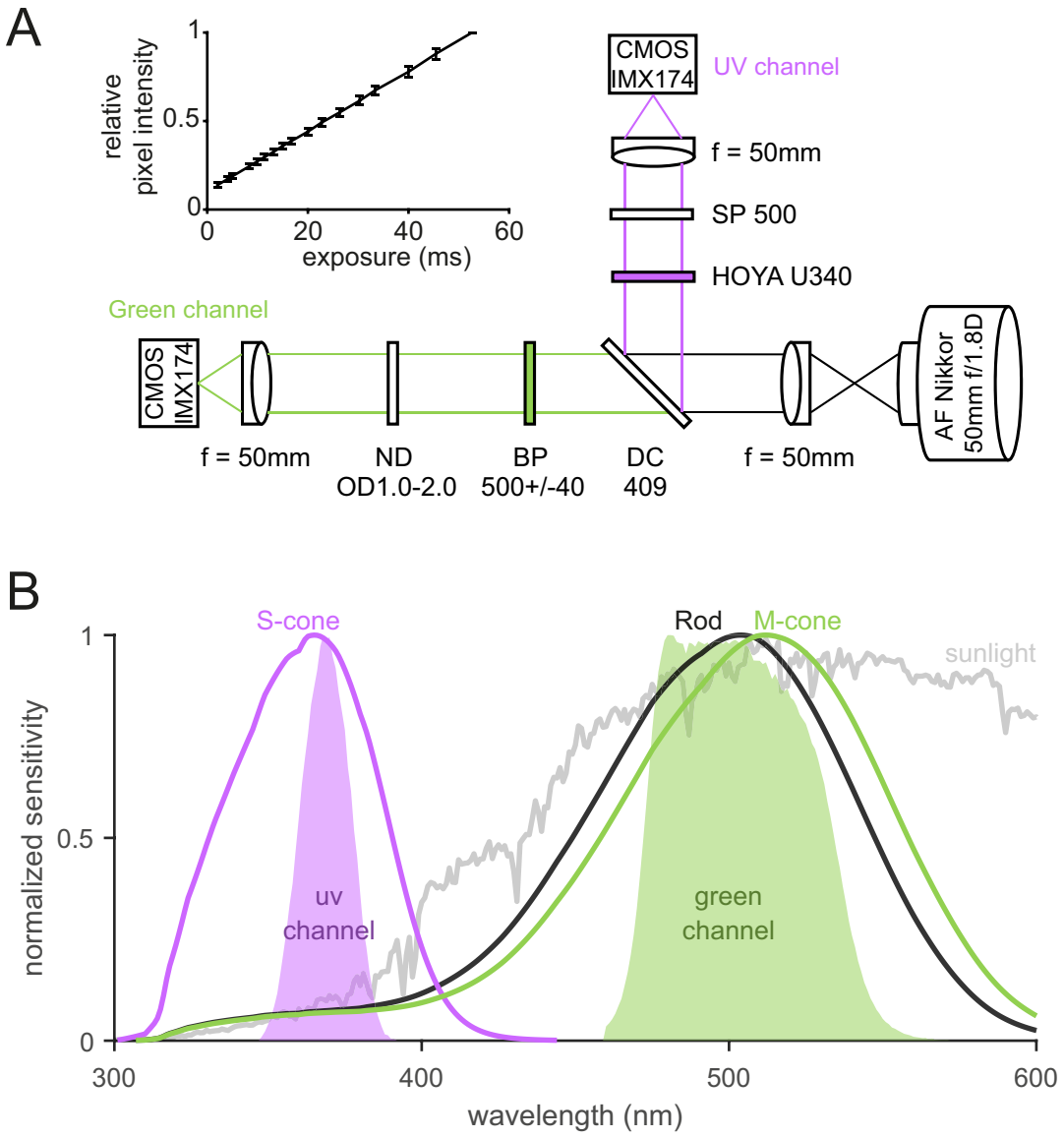
[65] Roorda A, Williams DR. The arrangement of the three cone classes in the living human eye. Nature. 1999;397:520–522.

[66] Hofer H, Carroll J, Neitz J, Neitz M, Williams DR. Organization of the Human Trichromatic Cone Mosaic. J Neurosci. 2005;25:9669–9679.

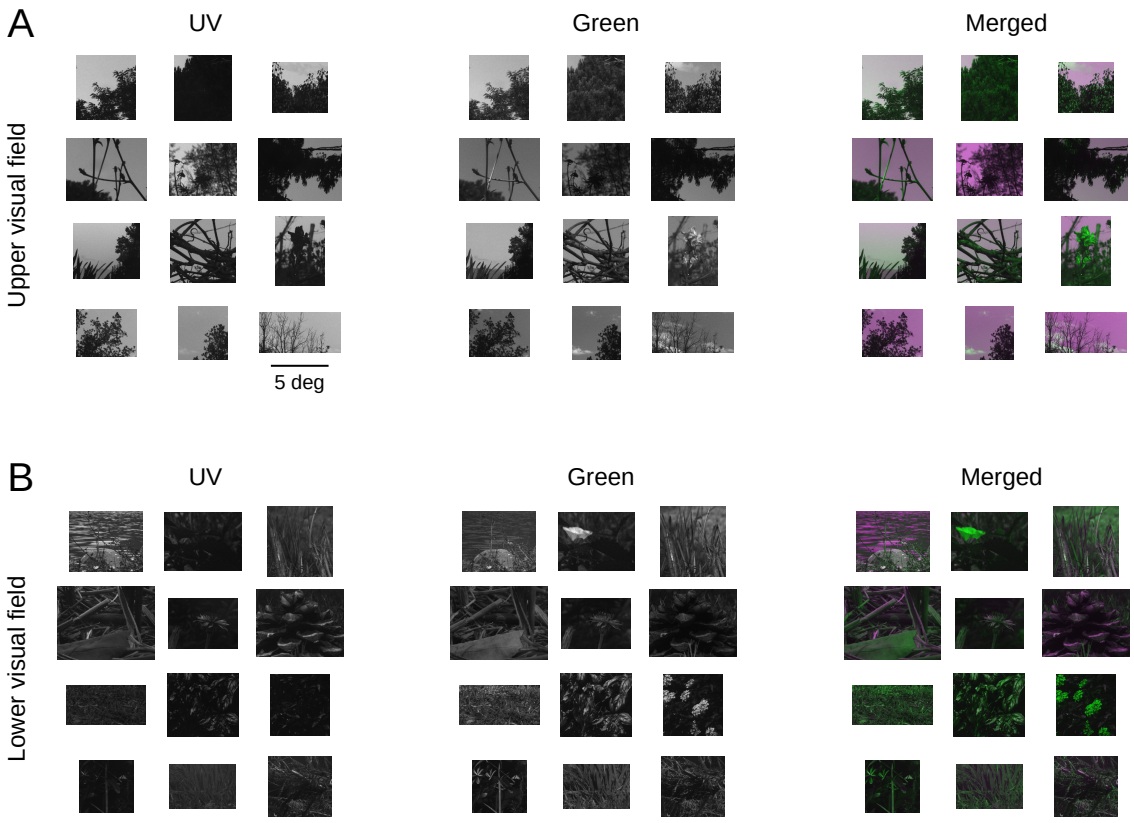
[67] Bleckert A, Schwartz GW, Turner MH, Rieke F, Wong ROL. Visual space is represented by nonmatching topographies of distinct mouse retinal ganglion cell types. Curr Biol. 2014;24:310–315.

[68] Jacoby J, Schwartz GW. Three Small-Receptive-Field Ganglion Cells in the Mouse Retina Are Distinctly Tuned to Size, Speed, and Object Motion. J Neurosci. 2017;37:610–625.

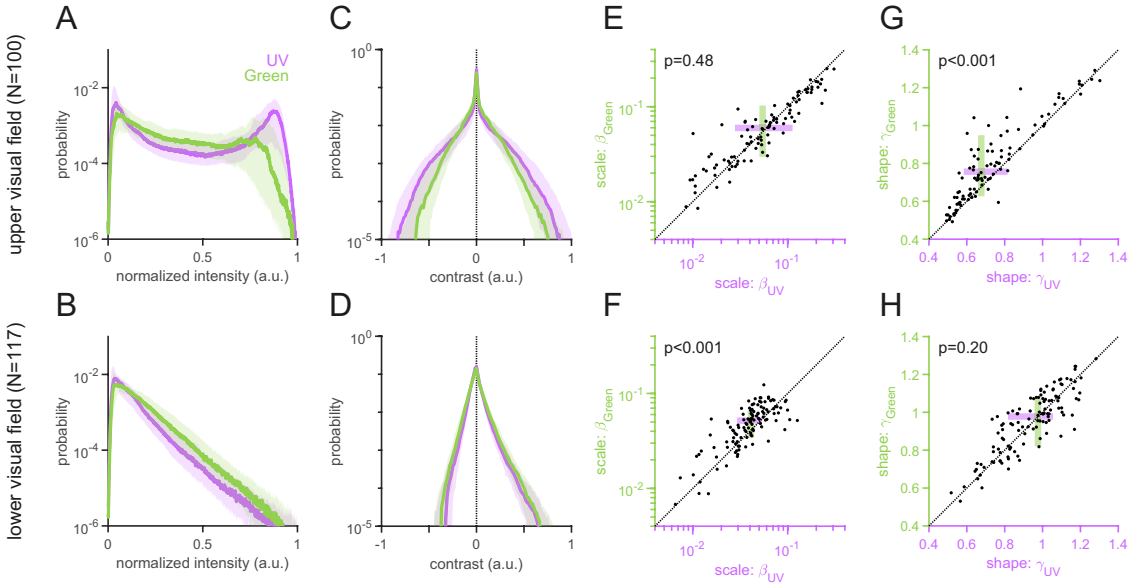
bioRxiv preprint doi: https://doi.org/10.1101/2021.04.08.438953; this version posted July 29, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.



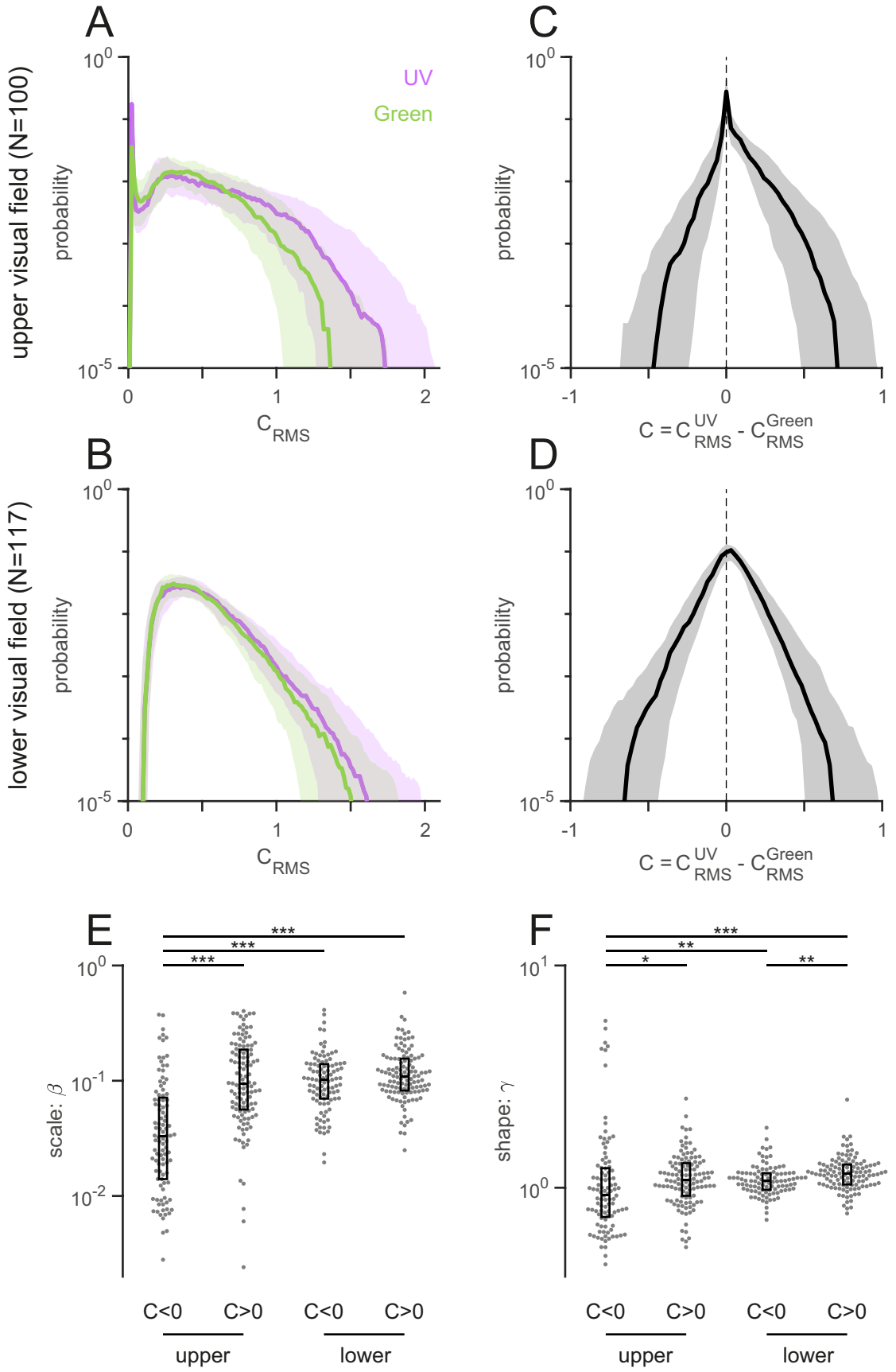
bioRxiv preprint doi: https://doi.org/10.1101/2021.04.08.438953; this version posted July 29, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.



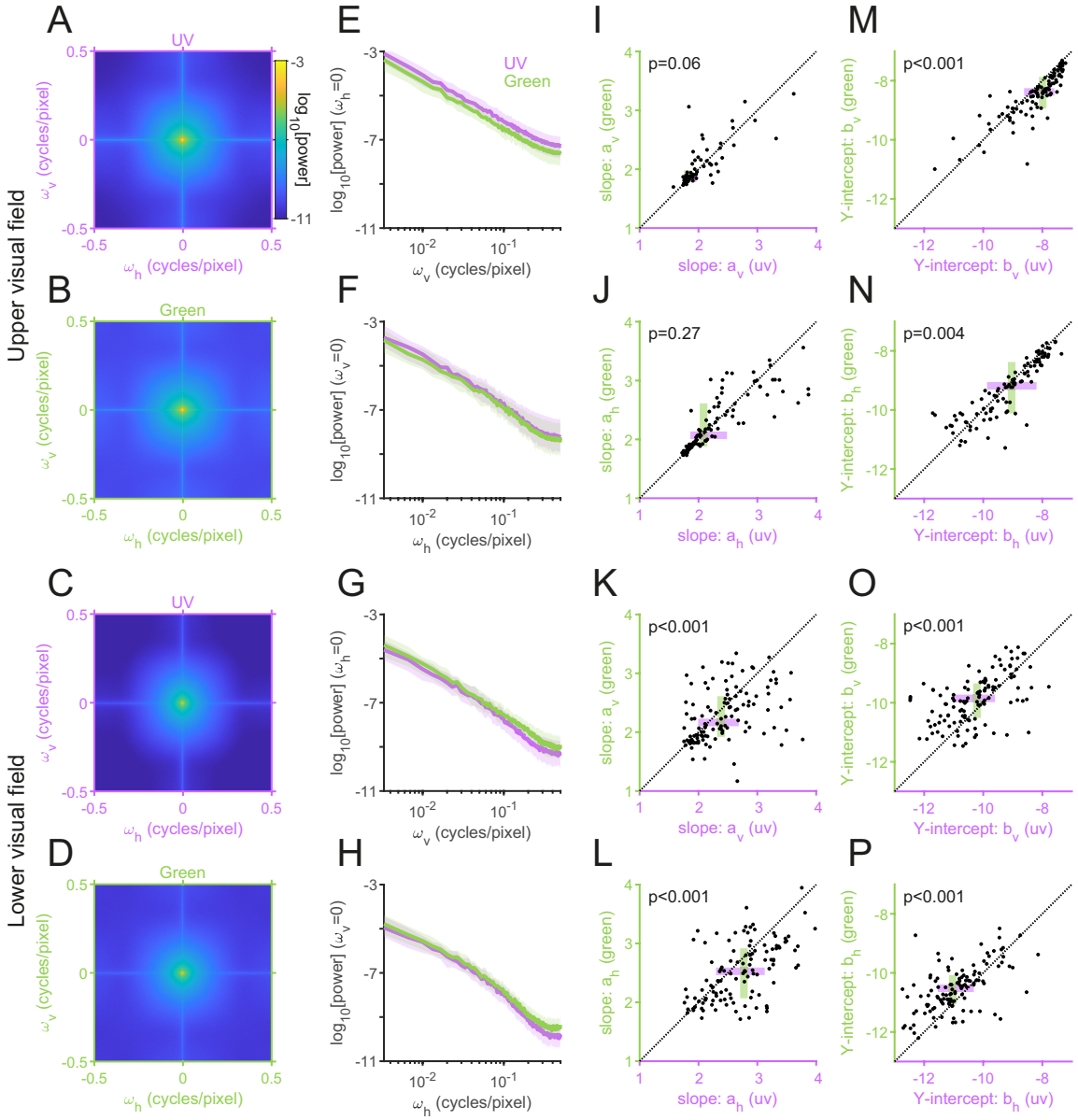
bioRxiv preprint doi: https://doi.org/10.1101/2021.04.08.438953; this version posted July 29, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.



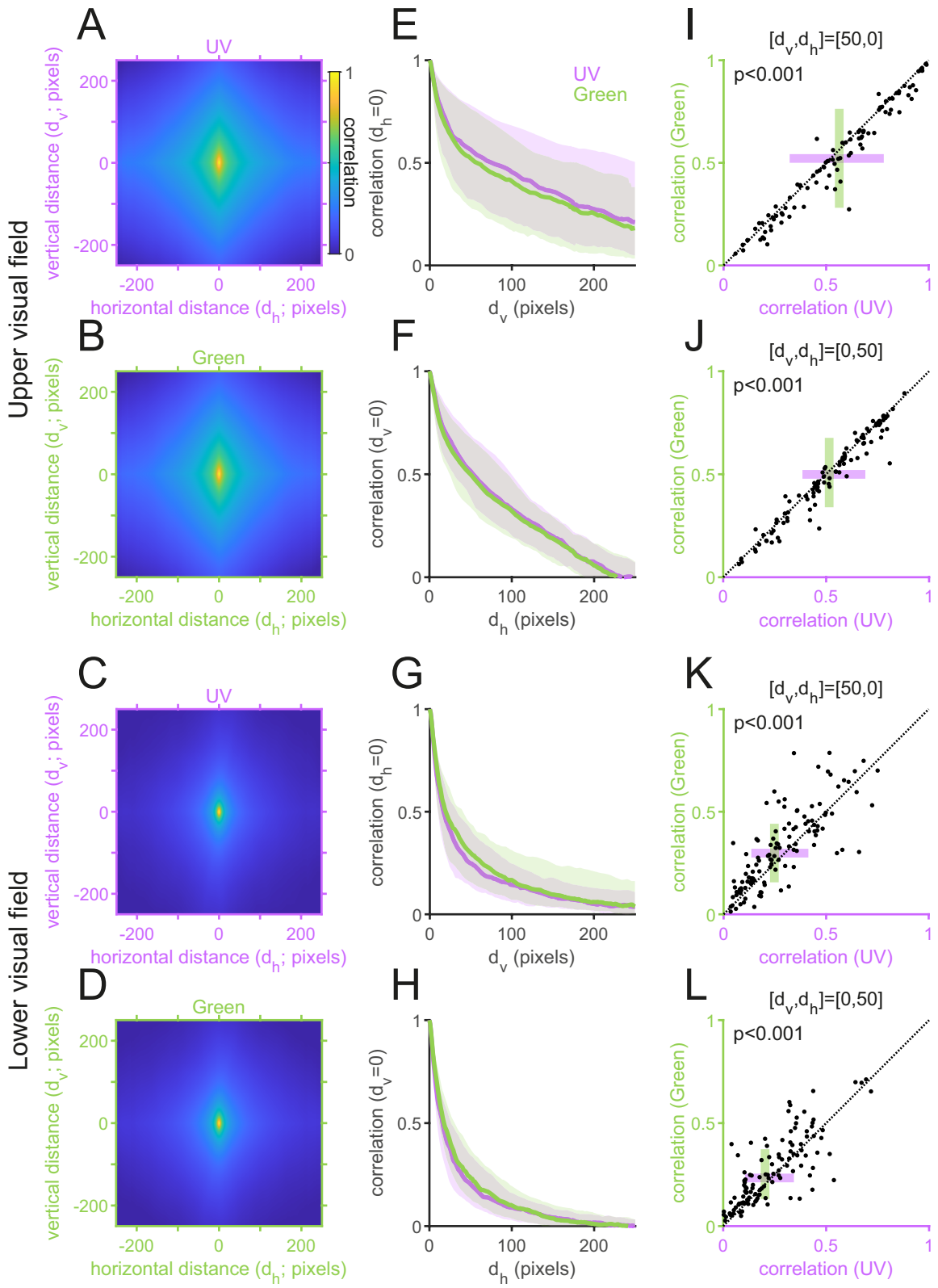
bioRxiv preprint doi: https://doi.org/10.1101/2021.04.08.438953; this version posted July 29, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.



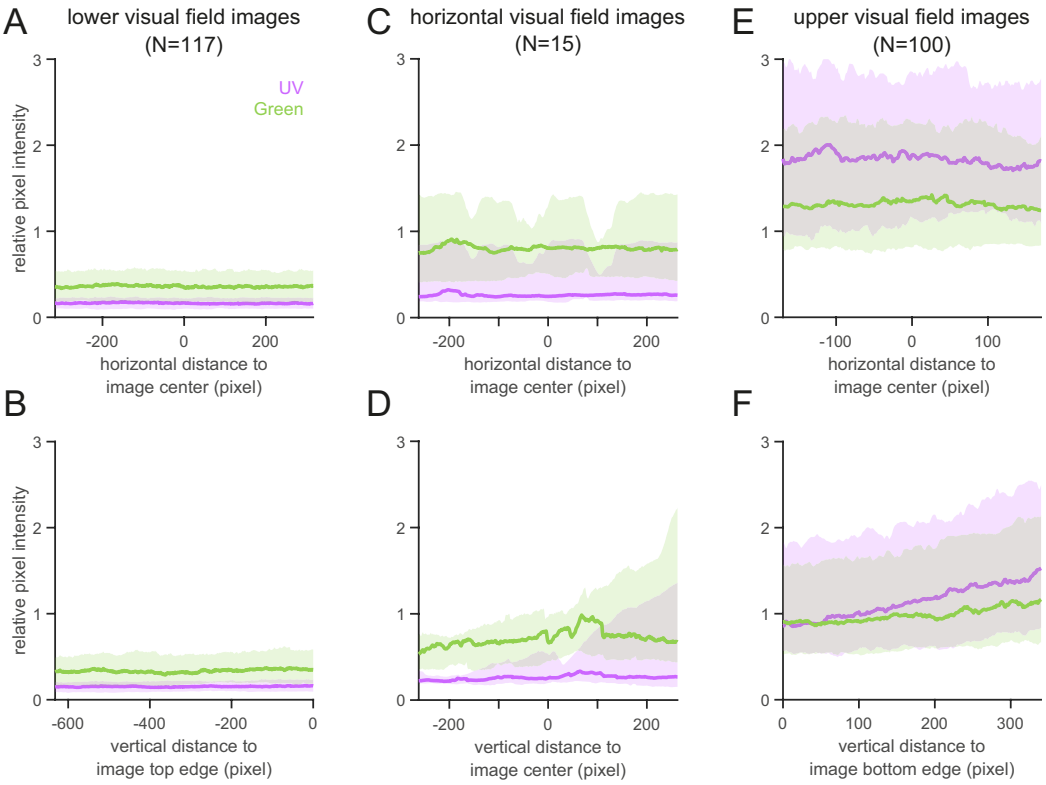
bioRxiv preprint doi: https://doi.org/10.1101/2021.04.08.438953; this version posted July 29, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.



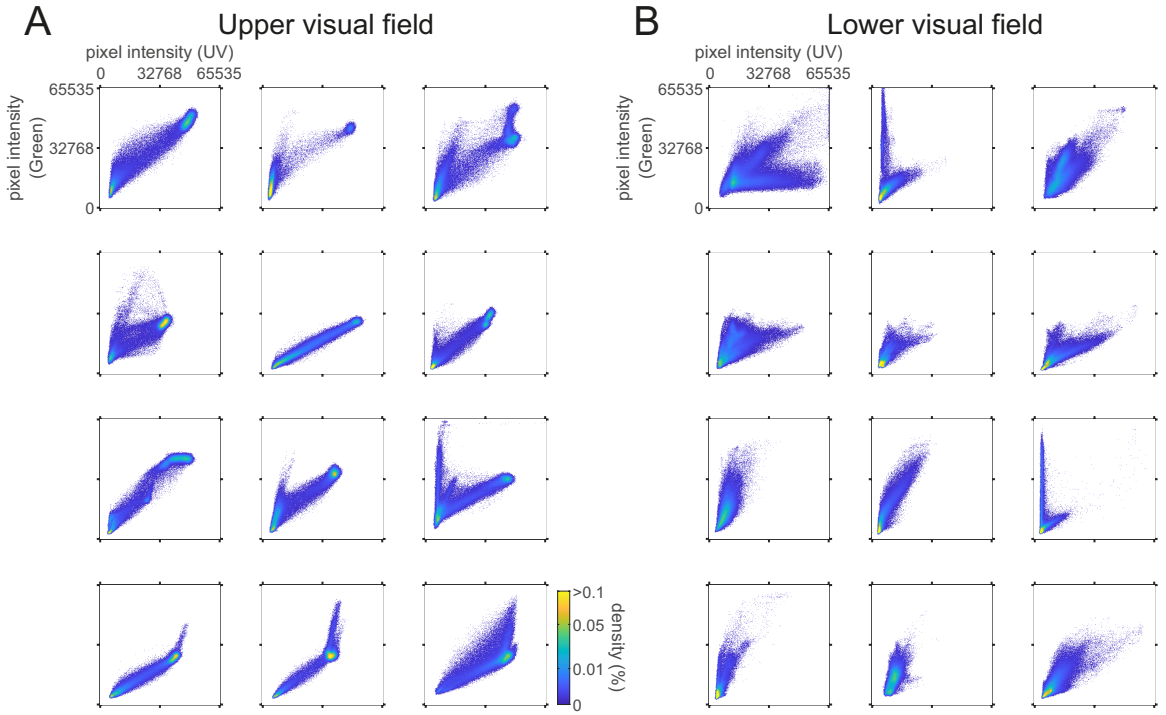
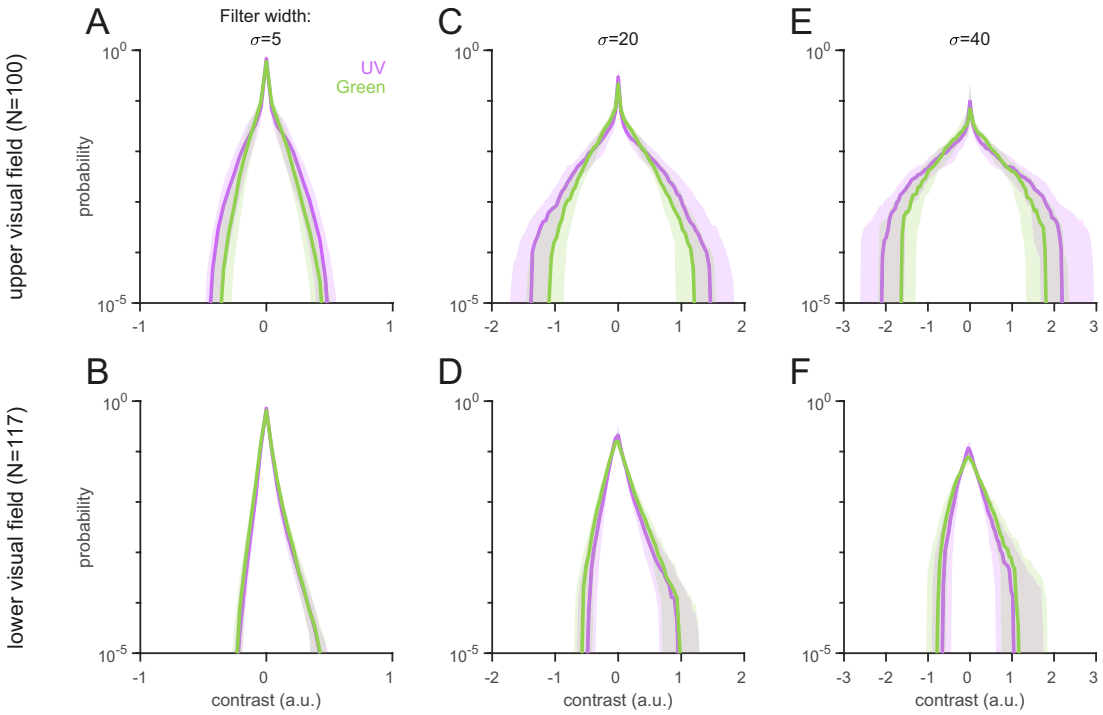
bioRxiv preprint doi: https://doi.org/10.1101/2021.04.08.438953; this version posted July 29, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.



bioRxiv preprint doi: https://doi.org/10.1101/2021.04.08.438953; this version posted July 29, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.



bioRxiv preprint doi: https://doi.org/10.1101/2021.04.08.438953; this version posted July 29, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.



bioRxiv preprint doi: https://doi.org/10.1101/2021.04.08.438953; this version posted July 29, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.