https://doi.org/10.2352/ISSN.2470-1173.2021.11.HVEI-151

Neurocomputational model explains spatial variations in perceived lightness induced by luminance edges in the image

Michael E. Rudd, Department of Psychology and Center for Integrative Neuroscience, University of Nevada, Reno, NV 89557 USA Email: mrudd@unr.edu

Abstract: Computer simulations of an extended version of a neural model of lightness perception [1,2] are presented. The model provides a unitary account of several key aspects of spatial lightness phenomenology, including contrast and assimilation, and asymmetries in the strengths of lightness and darkness induction. It does this by invoking mechanisms that have also been shown to account for the overall magnitude of dynamic range compression in experiments involving lightness matches made to real-world surfaces [2]. The model assumptions are derived partly from parametric measurements of visual responses of ON and OFF cells responses in the lateral geniculate nucleus of the macaque monkey [3,4] and partly from human quantitative psychophysical measurements. The model's computations and architecture are consistent with the properties of human visual neurophysiology as they are currently understood. The neural model's predictions and behavior are contrasted though the simulations with those of other lightness models, including Retinex theory [5] and the lightness filling-in models of Grossberg and his colleagues [6].

Introduction

This paper reports some new theoretical results that are part of an ongoing project to model achromatic color percepts ("lightness," for short) with a computational theory constrained by both psychophysical and neurophysiological data. Because the theory conforms to known visual physiology, I refer to it as a "neurocomputational" model of lightness. The results reported here extend work that was reported at the 2020 Electronic Imaging Conference and later published in extended form [2]. Other related modeling can be found in [1,7,8] as well in preceding papers cited in those works.

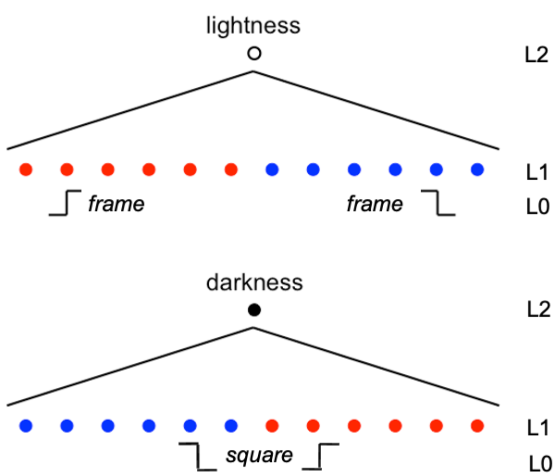

Fig. 1. Diagram of the neural lightness model showing its response to the stimulus used in the experiments modeled. See text for details.

The basic framework of the neural lightness model is illustrated in Fig. 1. The model assumes the existence of populations of "lightness" and "darkness" neurons at Level 2 of the model—possibly corresponding to documented "brightness" and "darkness" neurons located in area V4 of primate visual cortex [9]—that spatially integrate the outputs of local, directionally-selective contrast detectors (red and blue nodes) at Level 1 of the model. The contrast detectors half-wave rectify their spatially-oriented odd-symmetric receptive field outputs. At model Level 3 (not shown), achromatic color is computed at each retinal position by subtracting the output of the darkness neuron located at that position in the Level 2 network from the output of the lightness neuron located at the same position. Level 0 is the input.

In this one-dimensional depiction of the model, only two types of these odd-symmetric receptive fields exist: those that detect luminance increases in the direction of the receptive field centers (in retinal coordinates) of the lightness neurons that spatially integrate their outputs, and those that detect a decrease in luminance in the direction of the receptive field centers of the darkness neurons that spatially integrate their outputs. However, in the 2D model described here, the local contrast detectors exist at many different orientations, and spatial scales.

A critical assumption of the model is that the local contrast detector neurons that detect local luminance *increments* have different sensory gains than the local contrast detector neurons that detect luminance *decrements*. An explanation of exactly how this is accomplished is one of the contributions of the work reported here and will be explained below. The receptive field shapes of the lightness and darkness neurons at Level 2 that separately integrate the outputs of the local contrast detectors at Level 1 are assumed to be identical.

It follows from the model assumptions that the achromatic color assigned to any point in the spatial color map at Level 3 of the model (i.e. in the model output, which models the achromatic color percept) will be computed from a long-range spatial sum of local spatially-filtered luminance steps or gradients (with luminance measured in log units) [1]. Furthermore, the influence of a luminance step or gradient that is located at a distance z from the point to which the color is assigned will depend on the product of two independent factors: 1) the distance z of the step or gradient from the point in the image to which the final achromatic color is assigned at Level 3; and 2) the contrast polarity of the step or gradient measured along a vector directed from the local contrast element to the image point to which color is assigned. This fact is expressed by the following equation, which specifies the *total* weight given to a small luminance step or gradient element in the computation of achromatic color as a product of a spatial weighting factor $\omega(z)$ that depends only on the distance z of the step or gradient element from the point to which achromatic color is assigned, and another weighting factor n_p that depends only on the contrast polarity of the step or gradient element:

$$w(z,p) = \omega(z)n_p\,. \qquad (1)$$

On the basis of data on the neural response of ON- and OFF-cells in the lateral geniculate nucleus of macaque monkey [3,4], I have argued [2] that $n_+ = 0.27$ and $n_- = 1.0$, where "+" denotes the visual response to an increment and "-" denotes the visual response to a decrement.

Development of the 2D Neural Model

In this section, I present the main new results, the goal of which is to specify plausible neural computations that could instantiate the model and to do this in a way that could at least potentially predict pointwise lightness judgments. Although the developments reported will result in a model that computes lightness at every point in the image, I will here apply the model only to a single simple psychophysical experiment described below. It will be left to future work to apply the model to a larger array of psychophysical data. Since the results of the single experiment modeled here consist of unitary lightness judgments that apply to whole surfaces, I will take the computed achromatic color at the center of the target patch or surface This assumption might be questioned and requires empirical verification, but it suffices for the purposes at hand. I will return to this issue in the discussion.

2D Model Description

Level 1: Half-wave Rectified and Spatially-oriented "Barbell" Filters

Neurons at Level 1 of the model consist of directional contrast detectors with spatially-oriented receptive fields that detect local changes in luminance within some spatial frequency band. These neurons can detect either sharp luminance steps (edges) or gradual luminance changes (gradients). In the computer simulations reported here, pairs of either ON cells, or OFF cells, were separately combined for this purpose to create filters that I will refer to as "barbell" filters (BBFs) because of their shape (Fig. 2). Each BBF was created by half-wave rectifying the responses of two 2D difference-of-gaussians (DOG) filters, both of which modeled *either* an ON or an OFF cell response. To model each individual ON cell in a pair, a surround Gaussian in the pre-rectification filter was subtracted from the center Gaussian. The standard deviation σ_s of the surround Gaussian was 6 times as large as the standard deviation σ_c of its corresponding center gaussian. To model each individual OFF cell, a center Gaussian was subtracted from the surround Gaussian. The dimensions of the center and surround Gaussians were the same as for the center and surround Gaussians associated with the ON cells. In both cases, the area under each Gaussian was normalized to 1.0. The DOGs that models the receptive fields of the ON and OFF cells were multiplied by the neural gain factors assumed for that cell type: $n_+ = 0.27$ in the case of ON cells, and $n_- = 1.0$ in the case of OFF-cells, prior to half-wave rectification.

To model ON-BBFs, the response of an ON cell—call it ON2—whose receptive field center was located 1.808 x $\sqrt{2}(\sigma_c)$ the receptive field center of another ON cell—ON1—was subtracted from the response of ON1. The difference in the two separately rectified filter responses was then itself rectified. This produced a filter response selective only for luminance *increments* in the direction of the receptive field center of ON1 within the shared spatial frequency band of the two ON cells, whose DOG receptive fields were equal in size. The receptive field center of ON1 defined the location in the Level 1 map of the resulting ON-BBF constructed in this way (Fig. 2). The offset distance 1.808 x $\sqrt{2}(\sigma_c)$ was chosen because it produced the optimum degree of overlap in the receptive fields of the two ON or OFF cells comprising a BBF. To model OFF-BBFs, the response of a spatially offset OFF cell, OFF2, was similarly subtracted from the response of another OFF cell OFF1. The resulting difference filter was then rectified to produce an OFF-BBF selective for luminance *decrements* in the direction of the receptive field center of OFF1, which defined the location of the OFF-BBF in the Level 1 map (Fig. 2).

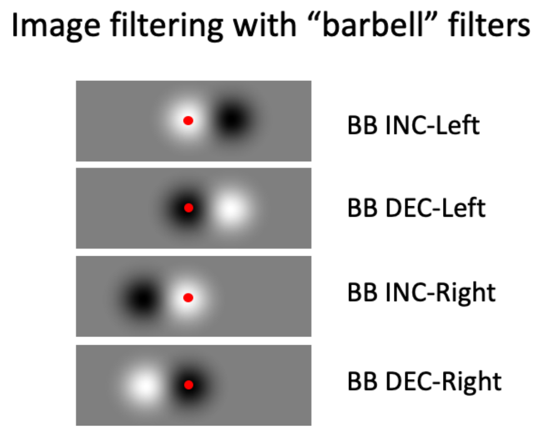

Fig. 2. Four example barbell filters. The filter labelled "BB INC-Left" responds only when luminance increments to the left in the input image, and similarly for the other filters shown.

The computer simulations described below were based on a bank of 16 such ON-BBFs and 16 OFF-BBFs that encoded 16 different angular directions of luminance increments and decrements, spaced at intervals of 22.5 deg in orientation angle.

Level 2: Directional Integration of BB filter Outputs by Lightness and Darkness Neurons

As illustrated in Fig. 1, separate populations of parallel pathway lightness (*L*) and darkness (*D*) at Level 2 of the model spatially integrate the outputs of directionally-sensitive local contrast neurons existing at Level 1. Each *L* neuron spatially integrates the outputs of ON-BBFs that encode luminance *increments* in the input image (Level 0) that are directed towards the *L* neuron's receptive field center, while each *D* neuron spatially integrates the outputs of OFF-BBFs that encode luminance *decrements* in the image that are directed towards the *D* neuron's receptive field center.

The ON-BBFs and OFF-BBFs described above model the local contrast detectors that encode spatially-directed luminance increments or decrements at Level 1. To understand how the *L* and *D* neurons at Level 2 spatially integrate the ON- and OFF-BBFs responses in 2D, it may help to envision this spatial integration in reverse: that is, in terms of how the outputs of the ON- and OFF-BBFs project to *L* and *D* neurons, respectively, at Level 2. It was assumed here that the neural projection of an ON-BBF or OFF-BBF to an *L* or *D* neuron at Level 2 neurons is *strongest* along an axis defined by a vector directed from the ON2 or OFF2 cell (defined above) to the ON1 or OFF1 cell that defines the location of the ON-BBF or OFF-BBF in the Level 1 map (as indicated by the red dots in Fig. 2). Level 2 neurons whose receptive field centers do not lie directly along this axis can also be activated by a BBF output, but their level of activation is reduced as a function of the angle of the vector from the BBF location to the L or D neuron's location relative to vector from ON2 or OFF2 to ON1 or OFF 1 described above. In particular, the neural activation produced by a particular BB-filter input to a particular *L* or *D* neuron was modeled by the equation:

$$c_p(x_2,y_2) = e^{-a\sqrt{(x_2-x_1)^2+(y_2-y_1)^2}} cos^k\theta_{\bar{v}} \int BB_p(\theta_{\bar{v}},x_1,y_1) \otimes I(x,y)dxdy, \qquad (2)$$

where $c_p(x_2,y_2)$ represents the contribution to the output of either an *L* neuron (subscript *p* = +) or a *D* neuron (subscript *p* = -) located at point (x_2,y_2) in the retinal image; *a* is a constant that determines the rate of spatial falloff in the exponential receptive fields of the Level 2 neurons; (x_1,y_1) is the location in the retinal image of the BBF whose output is integrated; *k* is a constant that determines the directional specificity of the population of network projections from Level 1 to Level 2 (the larger is *k*, the more directional the activations); $\theta_{\bar{v}}$ is the angle of a vector directed from (x_1,y_1) to (x_2,y_2) relative to a vector directed from the ON2 or OFF2 receptive field center to the ON1 or OFF2 receptive field center (x_1,y_1) of the BBF whose output is being integrated; and $BB_p(x_1,y_1)$ represents the activity of the ON-BBF (*p* = +) or OFF-BBF (*p* = -) at (x_1,y_1) whose response is being spatially integrated by either an *L* or *D* neuron at (x_2,y_2). Note that the symbol ⊗ in Eq. (3) denotes "operates on." It is used here in place of multiplication because the BBFs are not linear filters due to the fact that the filter itself, and the separate ON- or OFF- filters on which they are based, are each half-wave rectified.

It follows from Eq. (2) that the total activation of an *L* or *D* neuron located at (x_2,y_2) is given by

$$C_p(x_2,y_2) = \sum_{\theta_{\bar{v}},x_1,y_1} e^{-a\sqrt{(x_2-x_1)^2+(y_2-y_1)^2}} cos^k\theta_{\bar{v}} \int BB_p(\theta_{\bar{v}},x_1,y_1) \otimes I(x,y)dxdy. \qquad (3)$$

Level 3: Combining L and D neuronal activations to determine the percept

At model Level 3 (output level), neurons whose activities represent the perceived achromatic color corresponding to a location (x_3,y_3) in the retinal image combine the outputs of L and D neurons at that location according to the formula

$$P(x_3, y_3) = C_+(x_2, y_2) + C_-(x_2, y_2)\,. \qquad (4)$$

To avoid potential confusion due to ambiguities surrounding the definitions of lightness, brightness, and achromatic color in the literature, in Eq. (4) I have used the symbol "*P*," for *percept*, to represent the pointwise function that represents the perceptual output of the model.

Computer Simulation of a Psychophysical Lightness Matching Experiment

Background change experiment

Rudd [7] performed a lightness matching experiment in which two 1.06 deg X 1.06 deg squares, each with a luminance of 0.5 log cd/m^2, were presented on an LCD monitor, separated by a distance of 8.5 deg center-to-center. Each square was surrounded by 1.0 log cd/m^2 frame but the frame surrounding the right square was much narrower (0.19 deg) than the frame surrounding the left square (1.78 deg). Two observers adjusted the luminance of the left square to match the right square in lightness. This task was performed at 12 different levels of background luminance, ranging from a luminance well below that of the two squares, to a luminance well above the common luminance of the two frames.

The neural model asserts that the lightness of each square should depend on a sum of weighted steps in log luminance at the inner and outer edges of each frame with the weights given by Eq. (1). When the square luminances on the two sides are identical—as they were at the start of each experimental trial—the only thing that is different in the formula for the square lightness on the two sides of the display is the contribution of the *outer* edges of two frames. That edge was closer to the square on the right side of the display so, according to the model assumptions, a change in the background luminance should have a greater influence on the lightness of the right square than on the lightness of the left square. To achieve a lightness match between the two squares, an observer in the experiment must therefore compensate for a change in the background luminance by adjusting the left square luminance.

As the background luminance is *increased*, we expect the observer to decrease the luminance of the left (matching) square to compensate for the fact that the incremental luminance at both outer frame edges is decreasing, which will darken right square at a faster rate than the left square. Furthermore, the rate of change in the left square luminance adjustment as a function of the background luminance is predicted to be smaller when the outer frame edges both increment in the direction of the squares (that is, when the background luminance less than the frame luminance) than when the outer frame edges both decrement in the direction of the squares (background luminance greater than the frame luminance). This is because any incremental luminance directed towards a target region is assumed by model to be encoded by ON cells, while any decremental luminance directed towards a target region whose lightness is computed is encoded by OFF cells. Since the contrast polarity-dependent gain factor n_p is assumed here to be only 0.27 times as large for ON cell mediated lightness computation as for OFF cell mediated lightness computation, the neural model predicts that the rate of change in the left-square luminance settings required for a match to be about 0.27 as large when the background intensity is less than the common frames luminance as when it is greater than the common frames luminance. Fig. 3 illustrates the stimuli used in the background change experiment and the lightness matches made by the two observers. The average results for the two observers are consistent with the model predictions.

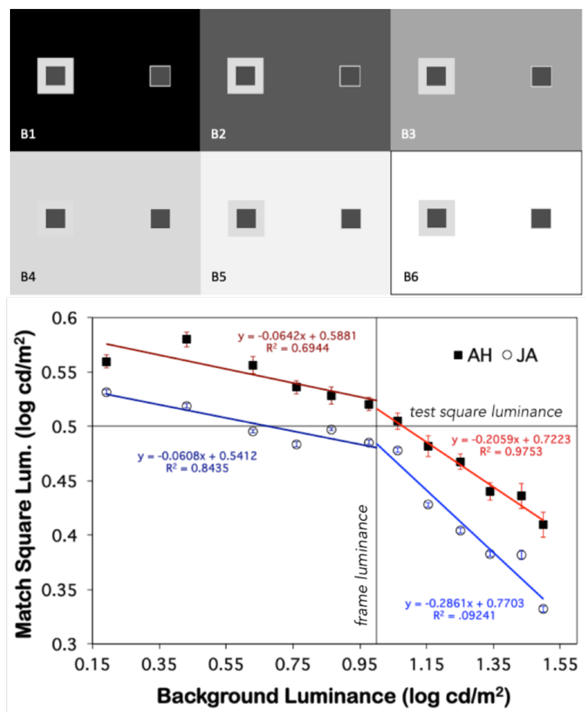

Fig. 3 (top). Stimuli used in the background change experiment. Six sample background luminances are shown for illustration purposes, but twelve were used in the actual experiment. (bottom) Lightness matches plotted against the background luminance.

Details of the Simulation

The simulation was performed in the MATLAB programming language (Version 9.9, R2020b) on a 16-inch Apple MacBook Pro laptop computer (2019) with an 8-Core Intel Core i9 processor.

The system of equations described in the previous section were simulated using discrete approximations in which the functions $I(x_0, y_0)$, $BB_p(x_1, y_1)$, $C_p(x_2, y_2)$, and $P(x_3, y_3)$, were represented by discrete two-dimensional arrays, each of size 905 x 905. These sizes were chosen such that 100 pixels represented 1 deg of visual angle. A (small) exception was made for the sizes of the central squares, which were 105 x 105 pixels (1.05 deg x 1.05 deg) instead of 1.06 deg x 1.06 deg, as in the actual experiment. This change was made so that the simulation output (i.e. the percept) could be evaluated at a pixel corresponding to the square centers.

The experimental data were modeled by separately simulating the model's response to the left and right square-and-frame stimuli in the experiment for 12 values of background luminance, which was varied in the simulation from 0.2-1.52 log cd/m^2 in steps of 0.2 log cd/m^2. The luminance of both frames was fixed at the value 1.0 log cd/m^2, and the luminance of the right square was fixed at the value 0.5 log cd/m^2, as in the original experiment. The width of the right frame in the simulation was 19 pixels (0.19 deg) and the width of the left frame was 178 pixels (1.78 deg).

To model each barbell filter, the center-to-center distance between the receptive field centers of the ON1 (OFF1) and ON2 (OFF2) cells from which the ON-BBFs (OFF-BBFs) were constructed was set to the value $1.808\sqrt{2}\sigma_C$, where σ_C is the standard deviation of the Gaussian function defining the center mechanism of the DOG function that modeled the receptive field of each ON or OFF cell. That standard deviation was set to the value 1.0 and only one scale of DOG function was used in the simulations reported here (though the model, more generally, has been programmed to include multiple spatial scales of DOG filters, spaced in octaves).

The only model parameter that was free to vary in the simulations was the spatial decay constant *a* of the exponential receptive fields of the *L* and *D* neurons at model Level 2 (Eqs. (2) and (3)). By appropriately adjusting *a*, a successful lightness match was achieved by the model at all background levels.

To model the perceptual response to the right square-and-frame configuration, the background luminance was varied and the Level 3 activity at the location of the right square center ($P_{RS}(453,453)$) was recorded at each of the 12 simulated background levels. To model the perceptual response to the left square-and-frame configuration, the left square luminance was varied along with the background luminance according to the average lightness matches of the two observers in the original experiment. This variation corresponded to the equations: log S_{left} = 0.56465 – 0.0625 log *B* when the background luminance *B* was less than the frames luminance, and log S_{left} = 0.7463 – 0.2460 log *B* when the background luminance *B* was greater than the frames luminance (see regression models in Fig. 3). Over repeated trials of the simulation, the space constant *a* of the Level 2 neurons was adjusted until the neural activity at model Level 3 at the location of the left square center ($P_{LS}(453,453)$) matched the neural activity at the right square center (that is, until $P_{LS}(453,453) = P_{RS}(453,453)$ at all values of the background luminance). A satisfactory match was achieved when a = 0.7866 deg^{-1} (see Fig. 4).

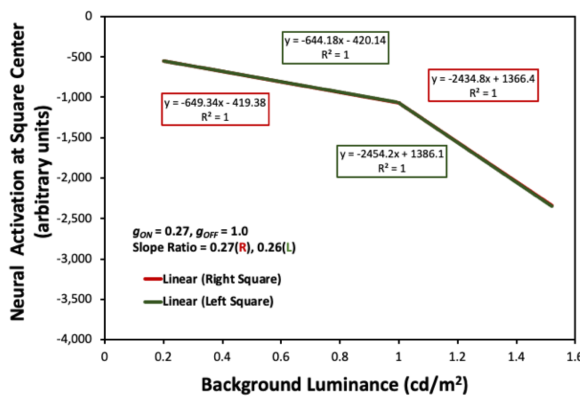

Fig. 4. Simulated results for the background change experiment.

General Discussion

It has been shown here that a neurally plausible model of human lightness computation can account for key features of lightness matching data from a psychophysical experiment utilizing simple square-and-frame stimuli presented on a computer monitor. The two primary mechanisms that the model employs to explain the perceptual data are: 1) separate encoding of luminance increments and decrements by parallel ON and OFF channels characterized by differing inherent neural gains, and: 2) independent spatial summation of neural responses to increments and decrements by large-scale receptive fields of separate populations of lightness- and darkness-encoding neurons which are subsequently combined to produce a unitary achromatic color percept.

The model was able to fit the average lightness matches from two observers made to a square-and-frame stimulus with by varying on one parameter: the common spatial decay constant of the exponential receptive fields of the lightness and darkness neurons.

Setting this parameter to the value $a = 0.79$ resulted in an excellent fit to the perceptual data.

I am currently adapting the neural model described here to account a more extensive data set of lightness matches made to Staircase Gelb and related stimuli. These data were previously fit with a less fully elaborated version of the neural model [2]. To fit the Gelb data with the current model requires that the form of the distance-dependent component $\omega(d)$ of the edge weighting function be more complex than the simple exponential function used here. Motivating this more sophisticated model of $\omega(d)$ and applying it to the Gelb results requires considerably more space than is available in this short *Proceedings* paper, so a full report of this work will be saved for a subsequent paper.

One advantage of the present approach over the one taken in previous work on the model [2] is that the current version of the model can, in principle, make predictions not only about the perceived lightness of surfaces, but also about the pointwise lightness of small surface elements: hence the title of this paper. An important assumption made in the simulations reported here was that the lightness matches made by the actual psychophysical observers in the background change experiment were equivalent to matching the pointwise lightnesses at the centers of the two squares. Whether this is what human observers actually do when they match surfaces lightnesses is an open question. But a theory that makes quantitative predictions regarding pointwise lightness computations provides a valuable tool for helping to answer this important research question.

References

[1] M. E. Rudd, “Lightness computation by the human visual system,” J. Electron. Imaging, 26(3), 031209, 2017; doi: 10.1117/I JEI.26.3031209.

[2] M. E. Rudd, “Neurocomputational lightness model explains the appearance of real surfaces viewed under Gelb illumination. J. Percept. Imaging, 2020; doi: https://doi.org/10.2352/J.Percept.Imaging.2020.3.1.010502.

[3] R. L. De Valois, I. Abramov, and G. H. Jacobs, “Analysis of response patterns of LGN cells,” J. Opt. Soc. Am., 56, 966-977, 1966.

[4] V. A. Billock, ”Hue opponency: chromatic valence functions, individual differences, cortical winner-take-all opponent modeling, and the relationship between spikes and sensitivity,” J. Opt. Soc. Am. A, 35, B267-B277, 2018.

[5] E. Land and J. J. McCann. “The retinex theory of vision,” J. Opt. Soc. Am, Vol. 61, 1-11,1971.

[6] S. Grossberg and D. Todorović, “Neural dynamics of 1-D and 2-D brightness perception: a unified model of classical and recent phenomena,” Percept. Psychophys., 43, 241-277, 1988

[7] M. E. Rudd, “How attention and contrast gain control interact to regulate lightness contrast and assimilation,” J. Vision, 10(14) 40, 2010; doi:10.1167/10.14.40.

[8] M. E. Rudd, “Edge integration in achromatic color perception and the lightness-darkness asymmetry,” J. Vision, 13(14)18, 2013; doi: 10.1167/13.14.18.

[9] Bushnell, B. N., Harding, P. J., Kosai, Y., Bair, W., and Pasupathy, A. “Equiluminance cells in visual cortical area V4.” J. Neurosci., 31, 12398–12412, 2011.