A Neural Model of Surface Perception: Lightness, Anchoring, and Filling-In

Stephen Grossberg and Simon Hong*

Department of Cognitive and Neural Systems
And
Center for Adaptive Systems
Boston University
677 Beacon Street, Boston, MA 02215, USA

Abbreviated title: A Neural Model of Surface Lightness Perception

August 2003


All correspondence should be addressed to
Professor Stephen Grossberg
Department of Cognitive and Neural Systems
Boston University
677 Beacon Street
Boston, MA 02215
Phone: 617-353-7858
Fax: 617-353-7755
Email: steve@bu.edu

* Authorship in alphabetical order. SG and SH were supported in part by the Air Force Office of Scientific Research (AFOSR F49620-01-1-0397), the Defense Advanced Research Projects Agency and the Office of Naval Research (ONR N00014-95-1-0409), and the Office of Naval Research (ONR N00014-01-1-0624).
Abstract

This article develops a neural model of how the visual system processes natural images under variable illumination conditions to generate surface lightness percepts. Previous models have clarified how the brain can compute the relative contrast of images from variably illuminated scenes. How the brain determines an absolute lightness scale that "anchors" percepts of surface lightness to use the full dynamic range of neurons remains an unsolved problem. Lightness anchoring properties include articulation, insulation, configuration, and area effects. The model quantitatively simulates these and other lightness data such as discounting the illuminant, the double brilliant illusion, lightness constancy and contrast, Mondrian contrast constancy, and the Craik-O’Brien-Cornsweet illusion. The model also clarifies the functional significance for lightness perception of anatomical and neurophysiological data, including gain control at retinal photoreceptors, and spatial contrast adaptation at the negative feedback circuit between the inner segment of photoreceptors and interacting horizontal cells. The model retina can hereby adjust its sensitivity to input intensities ranging from dim moonlight to dazzling sunlight. At later model cortical processing stages, boundary representations gate the filling-in of surface lightness via long-range horizontal connections. Variants of this filling-in mechanism run 100-1000 times faster than diffusion mechanisms of previous biological filling-in models, and shows how filling-in can occur at realistic speeds. A new anchoring mechanism called the Blurred-Highest-Luminance-As-White (BHLAW) rule helps simulate how surface lightness becomes sensitive to the spatial scale of objects in a scene. The model is also able to process natural images under variable lighting conditions.

Keywords: Surface perception, Lightness, Anchoring, Filling-in, Retinal adaptation
1. Introduction

The human visual system perceives surface reflectance (percent of light reflected by a surface in each wavelength) with remarkable fidelity even under greatly varying illumination conditions. The retina receives luminance signals, which are a product of reflectances and illumination levels (Hurlbert, 1989), from objects in the world, rather than the reflectances that are a property of object surfaces. From these luminance signals, the visual system needs to discount the illuminant to discover the reflectances themselves by using contextual cues, including cues of illumination (Figure 1). Discounting the illuminant is not sufficient, however, because the illuminant-discounted signals characterize only the relative amounts of light that each object surface reflects to the eyes. For effective perception, the brain also needs to discover an absolute lightness that can represent the full-range of experience from dim moonlight to dazzling sunlight. The present article describes a neural model that contributes to understanding how such an absolute lightness is constructed by the brain from the illuminant-contaminated signals that are received at our retinas.

Retinal preprocessing of visual signals contributes greatly to discovering an absolute lightness scale. These processes include two mechanisms of gain control: Light adaptation and contrast adaptation. Human vision adapts to ten orders of magnitude of daily variations of ambient illumination (Martin, 1983). For example, if the brain gets an input like the one in Figure 2A, it would “see” it like the one in Figure 2B. This property, termed light adaptation, depends in part on retinal circuitry (Werblin, 1971). Figure 2C shows the model response to varying background illumination. The graph illustrates how the range of maximal sensitivity of an early stage of model adaptation shifts with background illumination without undergoing compression, as also occurs in the retina (Werblin, 1971). Another dimension of adaptation is spatial contrast adaptation. For example, if there is a big contrast in the visual field such as the one in Figure 2D, the brain can, under a wide range of viewing conditions, properly rescale input signals to see the dark side as well as the bright side of the scene, as in the model simulation of Figure 2E. Since retinal ganglion cells, which are the sole output units of the retina, have firing rates that vary over less than three orders of magnitude, the visual system needs to compress the

![Figure 1. What the visual system sees is luminance, a product of reflectance and illumination. The visual system attempts to estimate the reflectance using available illumination cues.](image-url)
dynamic range of input at the retinal level, without a loss of sensitivity. Currently, the mechanisms of spatial as well as temporal component of contrast adaptation are still undergoing intensive experimental investigation (Demb, 2002; Baccus & Meister, 2002). Some of the gain-control mechanisms of the retina contributing to these adaptations may include: (1) Ca$^{2+}$ ion-mediated negative feedback occurring at the photoreceptors (Koutalos & Yau, 1996) and bipolar cells (Nawy, 2000); (2) bleaching of photopigments (Dowling, 1987; Fain, 2001); (3) surround negative feedback by the horizontal cell (HC) network (McMahon et al., 2001; Thibos & Werblin 1978; Werblin, 1974); and (4) a circuitry switch from cones to rods (Mills & Massey, 1995; Ribelayga, Wang & Mangel, 2002). Such mechanisms enable cells to dynamically change their operating range to adapt to varying lighting situations.
Surface lightness percepts cannot, however, fully be explained by such low-level mechanisms. For example, visual percepts depend upon appropriate interactions between both ON and OFF channel signals that seem to be largely segregated up until cortical area V1 (Schiller, Sandell & Maunsell, 1986; Schiller, 1992). Attempts to explain surface lightness range from the classic inference theory of Helmholtz (1866) to recent theories that Gilchrist and his colleagues classify as intrinsic image theories (Arend 1994; Gilchrist et al., 1999). While several theories propose that lightness is derived from luminance ratios among surfaces in a display,
these computations can, at best, recover relative reflectances. Thus there still remains the problem of systematically mapping these relatively defined lightness values to the absolute lightness values that are experienced during visible percepts. One proposed possible solution is the average luminance rule suggested by Helson (1943). This hypothesis proposes that the average luminance of the display, defined as middle gray, acts as a standard “anchoring” point for other luminances. For example, higher luminances than the average luminance will be assigned higher values of lightness than middle gray. Figures 3A and 3B show an example where this rule makes a correct prediction. However, when the rule meets a situation like the one in Figure 3C, it make the error shown in Figure 3D: The whiteboard becomes middle gray. As the example shows, the average luminance rule does not explain lightness data quantitatively.

In one of the first attempts to quantify lightness, Wallach introduced an anchoring hypothesis which became known as highest-luminance-as-white (HLAW) rule (Horn 1977; Land & McCann, 1971; Wallach, 1948, 1976). This rule assumes that the perceptual quantity “white” is assigned to the highest luminance in a given scene as the standard, and that lower luminant surfaces are assigned to gray values relative to it. According to this rule, the whiteboard in Figure 3E should look white, as in Figure 3F. In cases where there is a highest luminance like the one in Figure 3G, however, the HLAW rule makes a wrong prediction, as shown in Figure 3H. The white curve in Figure 3H that is superimposed on the image shows the profile of the predicted lightness along the horizontal section of the image that crosses the light source. The value “w” on the right side of Figure 3H marks the lightness value “white” along the vertical axis. By converting the intense illumination source into “white,” the HLAW rule drives all other lightness values to unacceptably small levels.

To overcome these shortcomings of previous hypotheses, the current model, which was briefly reported in Hong and Grossberg (2003), proposes how brain dynamics may instantiate a new rule called the blurred-highest-luminance-as-white (BHLAW) rule. The blurring part, which is spatial integration, makes the model sensitive to the area subtended by the highest luminance, thus introducing spatial-scale into the assessment of surface lightness (Figure 4A). This mechanism also enables this model to explain the self-luminosity of certain surface regions (Figure 4B). See Section 2.5 for further explanation.

Figures 3I, 3J show model’s property that is similar to HLAW rule. Figures 3K and 3L show the distinct property of the model that correctly predicts the lightness of the surface, despite the light source in the input. The curve on Figure 3L shows the profile of the simulated lightness of the horizontal section of the image that crosses the light source. The peak of the curve going above white “w” predicts that the light source will look self-luminous. By incorporating a BHLAW process into a multi-stage model of boundary and surface processing, the model also explains, among other lightness data, the four sets of data that Gilchrist and his colleagues (1999)
have proposed should be explained by any quantitative lightness theory. These four factors in lightness assignment are Articulation, Configuration, Insulation, and the Area Effect.

Figures 5A to 5E illustrate the procedure and the percepts of the Articulation Effect: A black patch (reflectance 3%) is fixed in front of a homogenous dark background (Figure 5A). When the patch gets illumination 30 times that of the dark background resulting in the luminance of 1.4 ftL (foot Lambert), it looks white (Figure 5B). (This 30-to-1 foreground-background illumination setting is also used in the following Configuration, Insulation, and the Area effects). When a real white patch (reflectance 90%) appears near the white-looking black patch, the black patch appears gray (Figure 5C). In the experiment, the subjects indicated the perceived reflectance by selecting a match from a Munsell chart of 16 examples near the subject. The Munsell chart was illuminated with a different light source so that the luminance of the whitest white, Munsell 9.5, was 160 ftL. The phenomenon illustrated in Figures 5B and 5C is called Gelb effect (for further discussion, see Cataliotti & Gilchrist, 1995). As more gray patches are added, the dark ones look darker and darker (Figures 5D and 5E). This darkening effect does not affect the highest luminance surface, which remains “anchored” to white. The graph in Figure 5F shows data that summarize this effect. Figure 5G shows the model simulation of these data. It should be noted that even in the two-Mondrian case in Figure 5C, the reflectances of these patches range from black to white covering the full span of reflectance used in the experiment. Thus the process of adding different luminance patches is just a process of “articulation”. This effect may not necessarily need patches of many different levels of luminance. For example, one
large white surface and one gray surface will give a smaller perceived lightness difference than in the situation where the two large surfaces are divided into small pieces and intermingled. See the text below for further explanation.

Figures 6A and 6B illustrate the procedure and the percepts of the Configuration Effect: A Mondrian display in Figure 6B—namely, a 2-D arrangement of juxtaposed gray patches—widens the range of perceived reflectance compared to the linear arrangement of patches shown in Figure 6A. Said more simply, the dark patches in Figure 6A appear lighter than the corresponding dark patches in Figure 6B. Graphs in Figures 6C and 6D show data that summarize this effect. Comparison of the graphs 6C and 6D shows that this effect becomes
Greater with more local articulation. Figures 6E and 6F show model simulation results. See the text below for further details.

Figures 7A to 7C show the procedure and the percepts of the Insulation Effect: When the staircase arrangement is surrounded by a white insulating region, the range of perceived reflectance widens (Figure 7B). The widening of the range of perceived lightness in a frame does not occur when the staircase is insulated by a black border (Figure 7C). The graph in Figure 7D
summarizes data showing this effect. Figure 7E shows the model simulation of these data. See the text below for further explanation.

The lower part of Figure 8A shows the experiment setting for the Area Effect. The head of the subject is covered by a dome that is divided into two regions. The upper part of Figure 8A illustrates the stimuli and the corresponding percepts. When the highest luminance area occupies more than half of the visual field, it appears white while the darker part looks gray. As the darker area occupies more than half of the visual field, however, it approaches white, while the lighter area gets pushed above white and appears self-luminous. The data curves in Figure 8B show this effect. Figure 8C shows the simulation results. See the text below for further explanation.

No published models have yet explained how these various data can be explained using an anchoring process, among other stages in the processing of lightness information. This study develops a biologically plausible model that explains and simulates a wide range of lightness data, including these data of Gilchrist and his colleagues.
2. Description of The Model

Figure 9 illustrates the model. The first stage adapts to ambient luminance and spatial contrasts. Using the adapted signal, the next stage generates contrast signals using multiple-scales of antagonistic ON-center OFF-surround and OFF-center ON-surround processes (see the following descriptions). The light-adapted signal itself also goes via a parallel pathway to the next level without change as the luminance signal. The process then branches into two streams: The boundary and surface processing streams, which have previously been modeled as the Boundary
Contour System (BCS) and Feature Contour System (FCS), respectively (Grossberg, 1994, 1997; Grossberg and Kelly, 1999; Grossberg and Mingolla, 1985a, 1985b; Grossberg and Todorovic, 1988; Kelly and Grossberg, 2000). The luminance and contrast signals are pooled at the filling-in stage of the surface system, where their spread is gated, or blocked, by boundary signals. At the final stage, the filled-in signals are rescaled via an anchoring process to assign appropriate lightness values. The anchored signals represent the perceived lightness in the model.

2.1 Retinal Adaptation

This stage of the model calculates the steady-state of retinal adaptation (light adaptation and spatial contrast adaptation) to a given input image. Using an intracellular gating mechanism at the outer-segment of the photoreceptor, the model first shifts the sensitivity curve of the photoreceptor and computes a light-adapted signal (GATED INPUT in Figure 10) at each position of the visual field (Baylor, Hodgkin & Lamb, 1974a, 1974b; Carpenter & Grossberg, 1981; Koutalos & Yau, 1996). This light-adapted signal is further processed at the inner segment
of the photoreceptor where it gets feedback from the horizontal cell (HC) that is connected with other HCs by gap junctions, forming a syncytium (Figure 10). This HC inhibition is hypothesized to further adjust the sensitivity curve of the photoreceptor for spatial contrast adaptation.

It is assumed that the permeability of the gap junctions between HCs decreases as the difference of the inputs to the HCs from the coupled photoreceptors increases. For simplicity only the connections between nearest neighbors are shown. In simulations, long-range connections are also allowed. The gray bidirectional arrows show the mutual influence between connected units. See the text and Appendix for further details.
(GLUTAMATE RELEASE in Figure 10) by modulating the Ca$^{2+}$ influx at the inner segment of the photoreceptor. This feedback prevents the output from saturation by localized high-contrast input signals. Thus this mechanism helps us see the room lit by a light bulb, the light bulb itself, and the label on it. See Appendix A for mathematical details.

2.2 Multiple-Scale Contrast and Luminance Stage

The retinally-adapted signal is processed by the center-surround contrast stage. The separation of the initial stage of retinal adaptation from the following center-surround stages seems to benefit the visual system in several ways: (1) The subsequent stages do not have to handle the light itself as the input anymore. They are cushioned from the impact of the vastly varying external inputs and receive normalized input signals. (2) The contrast stage is assumed to concentrate on spatial frequency-specific processing in multiple scales, extracting salient information for each spatial frequency. In the model, this stage simulates the cell types having concentric receptive fields of on-center off-surround or off-center on-surround found in the retina (Barlow, 1953; Cook and McReynolds, 1998; Kuffler, 1953; Werblin & Dowling, 1969) and the lateral geniculate nucleus (LGN) (Dubin & Cleland 1977; Hubel & Wiesel, 1961; Jones et. al., 2000). The model on-center off-surround (ON) units are excited by signals falling on the central part of their receptive fields, while they become suppressed when light falls on the surround of their receptive fields. This is implemented by a combination of a narrow excitatory Gaussian filter for the center and a broad inhibitory Gaussian filter for the surround. The model off-center on-surround (OFF) units increase their activities when a light stimulus falls on their surround receptive field and decrease their activities when a stimulus falls on the center part of the receptive field (Schiller 1992).

The 1-D cross-sections of the contrast operators are illustrated in Figure 9 between the RETINAL ADAPTATION and CONTRAST stages. Using feed-forward shunting equations (Grossberg 1983; see the equation below), the operation extracts local contrasts. In the case of an on-center off-surround network, for example, this process effectively eliminates the illumination influence on the local input in a scale-specific manner by dividing the input of the center by the local average represented by the surround, thus estimating the local contrast:

$$\text{Contrast} = \frac{L_{\text{spot}} - L_{\text{background}}}{L_{\text{spot}} + L_{\text{background}}}$$

where $L_{\text{spot}}$ and $L_{\text{background}}$ are the luminances of the probe stimulus and background, respectively. Using different sizes of surround, the system extracts small-scale to large-scale contrasts. These various surround sizes simulate the different sizes of lateral inhibition cell types in the retina (for a review, see Masland, 2001). The model uses a fixed narrow center kernel with the different surround scales (Grossberg et al, 1995; Mingolla et al, 1999) and thereby also simulates the output of a sharp center at the ganglion cells due to interactions in the retinal
network (Cook and McReynolds, 1998; Roska et al., 2000). All discussion of multiple spatial scales in this study is restricted to scales of surround kernels. This simplification reflects our minimal approach to designing a model capable of explaining various lightness data. Although the center-surround process is presented separately from the retinal adaptation stage, it also contributes to background adaptation. In the same vein, the adaptation carried by the photoreceptor and HCs is a type of center-surround process with a large surround scale.

Multiple scales, which are defined by the width of the Gaussian filters, contribute with different weights to form a complete representation of the stimulus. Figure 11 illustrates how three different scales respond to luminance inputs. Since the large-scale signal tends to represent the luminance signal more veridically (see LARGE-SCALE RESPONSE in Figure 11), we call this as luminance signal. A scale that is small relative to a given image region may exhibit a property known as brightness bowing (see SMALL-SCALE RESPONSE in Figure 11). In order to preserve the resolution of the image, single-scale models typically use such small scales, while
omitting bigger scales. The brightness bowing property stems from the fact that a small-scale (high frequency) center-surround unit acts like an edge detector of luminance boundaries, and thereby suppresses information from large homogeneous surface parts. Summation of multiple scales naturally compensates for this problem.

This multiple-scale property of the model is consistent with electrophysiological observations in V1 of alert primates (Bartlett & Doty, 1974; Kayama et al., 1979; Kinoshita & Komatsu, 2001; Komatsu, Murakami & Kinoshita, 1996) and anesthetized cats (MacEvoy, Kim & Paradiso, 1998), where cells not only code edge signals but also uniform surface luminance as well. A recent electrophysiology study with alert monkeys by Friedman et al. (2003) also shows that cells in V1 and V2 code uniform color surface information. For the LGN, uniform surface luminance coding units have been found in anesthetized primates (Marrocco, 1972) and cats (Papaioannou & White, 1972) as well as in alert primates (Barlow, Snodderly & Swadlow, 1978; Kayama et al., 1979). When surface luminance was temporally modulated, the cells in the LGN and V1 of anesthetized cats coding the surface region were modulated (Rossi et al, 1996; Rossi & Paradiso, 1999). Bartlett and his colleagues (1980) failed to detect such neurons in visual cortex of the awake rabbit. Their data suggest that there may be some differences between species, and techniques of anesthesia also seem to play an important role.

2.3 Boundary Formation

In surface perception, the boundaries defining a surface are prominent cues. Boundary formation as a factor in surface percept generation has been recognized by the Gestalt psychologists (Koffka, 1935), and used to model psychophysical and neural data about surface perception (Arrington, 1994; Cohen & Grossberg, 1984; Grossberg & Mingolla, 1985a, 1985b; Grossberg & Kelly, 1999; Grossberg, Hwang & Mingolla, 2002; Kelly & Grossberg, 2000; Pessoa, Mingolla & Neumann, 1995). Grossberg & Todorović (1988) developed this concept to simulate psychophysical data about brightness (perceived luminance). In their model, a center-surround network among cells obeying membrane equations discounts the illuminant. The surviving contrast signals are used to fill-in a surface brightness estimate within a region surrounded by boundaries that are themselves derived from the illuminant-discounted contrast signals (see the following Filling-In section for more details). The current model adopts this hypothesis to explain lightness data using contrast and luminance signals together to fill-in a region defined by surrounding boundaries (Figure 9).

Boundary formation begins at model simple cells that simulate orientationally-tuned simple cells in layer 4 of cortical area V1 (Figure 9), which have contrast-polarized and oriented ON (excitatory for luminance) and OFF (excitatory for darkness) regions in their receptive fields (Bullier and Henry 1979; Gilbert 1977; Hubel and Wiesel 1962). The model simple cells pool
model ON cell LGN outputs in their ON region and OFF cell LGN outputs in their OFF region. This is consistent with the observation that the ON and OFF subfield properties of simple cells seem to originate from the projections of ON and OFF cells in the LGN, respectively (Alonso, Usrey & Reid 2001; Lee et al., 2000; Reid & Alonso, 1995). Figure 12A shows the assumed input circuit for the model simple cell. These receptive field properties of a simple cell are modeled by a pair of elongated Gaussian kernels with shifted centers (Grossberg, Mingolla & Williamson, 1995; Mingolla, Ross & Grossberg, 1999; Pessoa, Mingolla & Neumann, 1995). Since the ON and OFF regions are spatially oriented and juxtaposed, the model simple cell is
maximally active when there is a luminance edge aligned with the oriented border between the ON and OFF regions. This property comes from the design of the ON and OFF regions made of Gaussian kernels that interact with each other antagonistically. For example, a simple cell with a vertical orientation and a light-dark polarity from left to right pools excitatory inputs from on-center off-surround contrast signals on the left side of the kernel and off-center on-surround signals from the right side of the kernel, and also pools inhibitory inputs from on-center off-surround contrast signals on the right side of the kernel and off-center on-surround signals from the left side of the kernel. Since the output is a rectified version of the sum of the filtered signals, the model simple cell becomes active only when there is an imbalance of luminance with the correct polarity across the oriented axis.

The boundary signals of an object need to be joined together even in cases where the contrast polarity reverses along the border of the object, such as at the edge of a middle gray object on a white-and-black checkerboard background (Grossberg 1994). The model achieves this requirement using model complex cells that pool a pair of light-dark and dark-light simple cell signals of the same orientation at each position. This pooling process is illustrated in Figure 12B. The classical proposal of such a hierarchical combination of simple cell outputs at complex cells (Hubel and Wiesel 1962; Schiller, Finlay & Volman, 1976) is supported by recent experimental data (Alonso & Martinez, 1998; Dresp & Grossberg, 1997; Martinez & Alonso, 2001), a theoretical analysis (Sakai & Tanaka, 2000) and modeling studies (e.g., Gove, Grossberg & Mingolla, 1995; Grossberg & Mingolla, 1985a). By this process, a model complex cell simulates the known complex cell property in V1 of responding to oriented luminance edges without having clear ON/OFF subfield zones (see Mechler & Ringach (2002) for further discussion). Additional feedback interactions are also known to exist (e.g., see Raizada and Grossberg, 2003), but are not needed for present purposes.

### 2.4 Surface Filling-In

At the filling-in stage, filling-in units first pool signals from multiple scales. Three scales are used: Small-scale and medium-scale contrast signals and large-scale luminance signals. These pooled signals spread during the filling-in process along long-range horizontal connections. The spread is blocked by boundary signals.

The model hereby extends the idea that pooled multiple-scale contrast and luminance signals are filled-in inside boundaries to form a surface percept (Cohen & Grossberg, 1984; Grossberg & Todorović, 1988; Pessoa, Mingolla & Neumann, 1995). This boundary-gated surface filling-in concept has been used to explain many psychophysical data about brightness and color perception and 3D figure-ground perception (Grossberg & Mingolla, 1985b; Grossberg & Kelly, 1999; Grossberg, Hwang & Mingolla, 2002). Consistent with this hypothesis, the
psychophysical study of surface perception by Paradiso and Nakayama (1991) showed that rapidly formed contour signals may gate the spread of surface signals to form a complete percept of a surface.

The filling-in mechanism utilizes two streams of the What cortical visual pathway: The surface stream runs through the blobs and V2 thin strips to V4; and the boundary stream runs through V1 interblobs and V2 interstrips to V4. These two streams have been proposed to compute complementary properties during visual information processing (Grossberg, 2000). Filling-in in the blind spot is an example of surface filling-in (Komatsu et al., 1996, 2000). Surface representations can be formed early in visual processing even without top-down cognition signals (Kamitani & Shimojo, in press). Sasaki et al. (2001) showed using fMRI that when a human subject perceives a transparent illusory region bounded by illusory contours, the V1 region corresponding to the illusory visual field became active. These data, combined with data about illusory contour representations known to exist in V2, are consistent with the possibility that surface representations may start to form in V2. More direct evidence comes from electrophysiology combined with cortical imaging: Hung et al. (2001) reported that the
Craik-O’Brien-Cornsweet Effect can be detected in V1 and is prevalent in V2. In their experiment, the activities of cells having receptive fields inside the homogeneous surfaces were modulated with cusps at the edge of the surfaces. The large spatial scale needed to fully integrate information across visual space (Angelucci et al., 2002) also marks V2 as a processing stage where surface representations start to get formed. Data concerning border ownership representations in V2 and V4 (Zhou et al., 2000) are also consistent with this conclusion. See Grossberg (1994, 1997) for further discussion.

Figure 13 shows a 1-D illustration of the filling-in model network. The round units on top represent the units in the filling-in layer. The units in the contrast and luminance layer feed their signals to the retinotopically corresponding filling-in units. Then the received signals spread between filling-in cells along long-range horizontal connections with Gaussian receptive fields. Signal propagation is gated by boundary signals (represented by a vertical line in Figure 13) coming from the complex cells (for simplicity just one set of gating units is shown). The gated horizontal connections have smaller conductances (thin horizontal lines) than the other ones (thick horizontal lines). The gray levels of the filling-in cells and the contrast, and luminance input cells represent the level of activation of these cells. The FILLING-IN LAYER illustrates two homogeneous filled-in regions, black and middle gray, separated by a boundary signal, representing two surfaces with different relative reflectances.

2.5 Lightness Anchoring
At the lightness anchoring stage, anchoring units receive the filled-in surface signals, while the activities of the anchoring units are modulated by a feedback signal originating from the anchoring module itself (Figures 9 and 14A).

_Anchoring and Blurred Highest Luminance As White (BHLAW) Rule_
While the retinal adaptation and contrast calculation stages generate normalized signals, these early processes do not provide output signals that encode an absolute lightness scale. Without an anchoring process, for example, a large whiteboard covering more than half of the visual field may look middle gray (Figure 3), because the early normalizing center-surround processes compute just the relative luminance of the center with the surround as the standard, or anchoring point. This is true for each spatial scale and is also true for the pooled multiple-scale representation. Anchoring rescales these relatively defined surface signals.

To simulate this rescaling process, the model embodies a new anchoring rule, called the Blurred-Highest-Luminance-As-White (BHLAW) rule. As noted in the Introduction, this revision overcomes problems of traditional HLAW rule; namely, a point-like small bright patch on a visual field will be dealt with the same as a large whiteboard occupying most of the visual field (Figure 3). By introducing a spatial Gaussian averaging mechanism into the anchoring
process, the model solves this problem and also explains the Gilchrist area effect (Figure 8). Figures 4A and 4B illustrate the model’s explanation of the area effect for a two-field Ganzfeld configuration, as in Figure 8. In such a display, there are just two homogeneous surfaces with different luminances, one the target surface, the other the Ganzfeld.

To achieve the anchoring property, the model first makes a blurred version of the anchoring signal, called the BHLAW signal (Figure 14A). The model uses this signal to anchor the highest value of the blurred pattern to white. This rescaling is achieved by using the BHLAW signal to modulate an automatic gain control process, labeled \( \Psi \) (Figures 9 and 14A). Gain \( \Psi \) multiplicatively rescales the filled-in surface signals. The process \( H \), which inhibits \( \Psi \), becomes activated whenever any BHLAW signal exceeds a threshold that determines the value of white.

Figure 14. BHLAW rule and Area Effect in a two-field Ganzfeld configuration. (A) Model circuit of lightness anchoring. The activities of the ANCHORING units are locally pooled by BHLAW units to form a blurred version of the ANCHORING signals. The filter used to generate the blurred signals is shown as a bell-shaped figure between the ANCHORING and BHLAW modules. The BHLAW signals are fed to an inhibitory unit \( H \). The unit \( H \) becomes active when any of the BHLAW unit exceeds its threshold set to WHITE and fires. When active, \( H \) inhibits the tonically active unit \( \Psi \) that modulates the activities of ANCHORING units. This circuit allows the activities of ANCHORING units to grow until at least one of the blurred version of anchoring signals, BHLAW, meets the criterion of WHITE. See Appendix A for mathematical details. One thing to notice is that the inhibition by \( H \) on \( \Psi \) lowers but does not completely shut off the activity of \( \Psi \), leaving a chance to the BHLAW signals to go beyond WHITE when the bottom up signal is strong enough, for example, a bright light source of some size. In such a case, even the BHLAW rule will be violated. (B-D), Two-dimensional simulation of two-field Ganzfeld configuration. The curve on each figure represents the activities of the units along the horizontal midline. This convention applies to all the following figures. The scale for the curve is denoted on the left side of each figure. B.STIMULUS shows the input configuration. D.ANCHORED LIGHTNESS shows the area effect corresponding to the one in Figure 4B. Note that the highest activity of the BHLAW module in Figure C is anchored to white (w).
(WHITE in Figures 4A and 4B). Since it is the highest value of the BHLAW signal that the model uses for anchoring, the anchored lightness (ANCHORED LIGHTNESS), or unblurred pattern, will look self-luminous (Figure 4B) in case the area of the highest filled-in activity is not broad enough to span the blurring kernel, because the blurring kernel then averages lower activities as well. As the area of the highest filled-in activity becomes smaller, this mechanism predicts that the background will approach WHITE because of the small difference between the highest and background BHLAW signals (Figure 4A). In such a case, the ANCHORED LIGHTNESS will grow until the highest BHLAW signal equals the anchoring value WHITE, which will also bring the background up close to WHITE. When the area of the highest luminance is larger than the blurring kernel, the highest BHLAW activity will equal the highest ANCHORED LIGHTNESS. Thus there will be no self-luminosity for that region after anchoring (Figure 4A). Figures 14B to 14D show a 2-D simulation of the two-field Ganzfeld configuration. The curve in each figure shows the activities of the cells along the horizontal middle section of the 2-D image. The labels on the left side of each figure indicate the scale of vertical axis for the curve; in particular, \( w \) denotes white.

### 3. Results

#### 3.1 Background Light Adaptation

Figure 2C shows light adaptation of model cells to ambient illumination. This shift property simulates the cell recording data of Werblin (1971). The leftmost curve of the shift property at lower values of background luminance corresponds to the physical limit of light adaptation observed in retinal ganglion cells (Barlow & Levick, 1969; Enroth-Cugell & Shapley, 1973a). Over a wide range of background luminances the model obeys the Weber law (Grossberg, 1983). Ambient illumination is removed by divisive intracellular negative feedback signals in the photoreceptors. See Appendix B for stimuli used for this simulation.

#### 3.2 Discounting the Illuminant

Figure 15A shows two light patches on a dark background seen in a gradient of illumination. To generate the input, light patches with the same reflectance and a background with a smaller constant reflectance were multiplied by a gradient of illumination. The curves on Figures 15A and 15B show the input intensities and anchored lightnesses along the horizontal midline, respectively. Figure 15B shows the property of illumination discounting: The light patch on the left is almost as light as the one on the right, unlike the one in Figure 15A. This property comes from the ratio-calculating property of the local contrast units. Figure 15B also shows that, when the gradient of illumination is big enough, the model exhibits a lightness bias where the square patch with higher illumination looks slightly lighter than the one with lower illumination. This
property of the model is due to the influence of the large scale that adds a more veridical representation of the stimuli to percepts. This prediction is supported by the observation that, when subjects are asked to decide the perceived reflectance of surfaces, they always give a higher value to the highly illuminated one than the same one with low illumination. (Gilchrist et. al., 1999, p. 826).

3.3 Simultaneous Contrast

Figure 16 shows a simulation of simultaneous contrast. The two middle gray patches in Figure 16A have identical luminance. In this configuration, small and medium scales calculate local ratio contrasts, and their contribution makes the light square on the dark background look lighter
than the one on the bright background even though they have identical luminance (Figure 16B). Since lightness anchoring just rescales the filled-in multiple-scale signals via BHLAW gain control, the anchoring process does not distort the relative lightnesses of the surfaces.

3.4 Evenly and Unevenly Illuminated Mondrians: Contrast Constancy

Figures 17A and 17B show an evenly illuminated Mondrian and the corresponding prediction of the percept by the model, respectively. A part of the square on the upper left of each figure has been cut and pasted to the square on the bottom right of the figure. Since there is no difference in luminance between the two squares, 17A shows no trace of cut patch at the bottom right square. Figure 17B shows that the square on the top left is perceived to be lighter than the bottom right square. The lightness difference between the two squares in 17B derives from the fact that the square on the right bottom is surrounded by lighter surfaces than the above square. For this reason, the square on the right bottom receives more surround suppression than the square on the

![Figure 17](image)

Figure 17. Evenly and unevenly illuminated Mondrians. To facilitate the comparison, a part of the square on the upper left of each figure has been cut and pasted to the square on the bottom right of the figure. (A-B) Evenly illuminated Mondrian. STIMULUS and ANCHORED LIGHTNESS panels show the configuration of an evenly illuminated Mondrian stimulus and the output of the model, respectively. (C-D) Unevenly illuminated Mondrian. The STIMULUS shows the differently illuminated target surfaces because of the illumination gradient. The gradient of illumination is made by a light source located at the bottom-right corner. ANCHORED LIGHTNESS shows the final output of the model. See the text for details.
upper left, which is surrounded by darker surfaces. Figures 17C and 17D show an unevenly illuminated Mondrian and the corresponding percept predicted by the model, respectively. A gradient of illumination from the bottom right to the top left is introduced by a light source located at the bottom right corner. The output of the model (Figure 17D) shows that the square on the upper left looks lighter than the one on the bottom right despite the fact that the luminance at the bottom right is higher. This correct prediction of the reflectances of the squares comes from the fact that the small and medium scales calculate the local contrasts and ignore the global-scale illumination gradient. This contrast constancy calculation by the two scales overrides the prediction by the big scale that picks up the gradient. Grossberg and Todorović (1988) simulated this effect with a single contrast scale.

3.5 Craik-O’Brien-Cornsweet Effect
The model is also capable of explaining the Craik-O’Brien-Cornsweet effect (Cornsweet, 1970), as shown in Figure 18. Figure 18A shows the stimulus with a uniform background luminance with a luminance cusp in the middle. Figure 18B shows the anchored lightness percept in which

![Figure 18. Craik-O’Brien-Cornsweet effect. (A) Two divided identical surfaces with a luminance cusp in the middle. (B) Simulated lightness of the model. The two surfaces are perceived differently. The boundary-gated homogenization of surface signals through a filling-in process makes the surface on the left look slightly lighter than the one on the right. (C) Boundary. (D) Small-scale contrast signals for the two surfaces. Left surface has more activities than the right one explaining the difference at the filled-in surface lightnesses in B.](image-url)
the left half of the image looks uniformly lighter than the right half. The model explains this illusion using the boundary-gated filling-in process, much as in Grossberg and Todorović (1988). At the filling-in stage, the pooled center-surround contrast signals are flattened within areas defined by boundary signals (Figure 18C). This flattening of signals makes the surface on the left lighter than the one on the right because of the larger contrast activities here (Figure 18D). For the illusion to hold, the area-defining boundaries play a critical role. When no boundaries surround the luminance cusp, the illusion does not happen; see Grossberg and Todorović (1988).

3.6 Double Brilliant Illusion

Bressan (2001) presented a lightness illusion called the Double-Brilliant illusion (Figures 19A and 19B), wherein the diamond that has less contrast around it (Figure 19B) looks lighter than the one having a high contrast around it (Figure 19A) even though both diamonds have the same luminance (Figure 19C). The model ascribes this phenomenon to the gated negative feedback in the retina. Because the permeability of gap junctions in the horizontal cell (HC) syncytium decreases only where there is a sharp luminance edge in the input, the gradual change of luminance around the diamond in Figure 19B does not block the diffusion of signals across the

Figure 19. Double Brilliant Illusion. (A-B) Stimuli. A psychophysical experiment shows that the diamond part of the stimulus B looks lighter than that of the stimulus A. (C-D) Stimulus and the output of simulation, respectively. (E-F) Simulated activities of HCs and the steady outputs of photoreceptor inner segments, respectively. See the text for more details. The figures A and B are from Bressan (2001).
HC syncytium. The luminance edges around the diamond in Figure 19A do block the diffusion process and segregate the diamond region from the rest of the figure. This gated-diffusion process is simulated in Figure 19E. The segregated and concentrated high signals shown in the diamond region on the left of 19E suppress the corresponding region of photoreceptor outputs. This results in a less active diamond region on the left in Figure 19F compared to the diamond region on the right. The anchored lightness of the model in Figure 19D reflects this difference and correctly predicts the illusion. This example illustrates that multiple levels of context-sensitive adaptation and recalibration can cooperate to yield lightness percepts. Since the retinal adaptation mechanisms are monocular, dichoptic presentations of different parts of these stimuli to each eye may yield different lightness effects.

3.7 Anchoring Properties
The model explains the four major effects of lightness anchoring (Articulation, Configuration, Insulation, Area Effect) as follows:

**Articulation effect:** The Articulation effect says that, as the display contains more gray surfaces, the range of perceived lightness widens (Figures 5A-5F). One noticeable fact is that even in the two-Mondrian case in Figure 5C, the reflectances of these patches range from black to white covering the full span of reflectance used in the experiment. Thus adding more gray patch does not result in a wider range of reflectance in the experiment. Figure 5G summarizes the model simulation of this effect. As the number of surface patches having different luminances increases in a region, the image contains more high spatial frequencies. In the model, this means that the medium and large spatial scale kernels have less chance to fully activate and suppress the homogeneous area of each patch. Figure 11 illustrates the situation: The divided square luminances on the right cause higher contrast signals in the medium and large-scales compared to the corresponding contrast signals on the left column with a larger square luminance stimulus. The loss of full suppression by each spatial-scale results from the mismatch between the size of the filters and of the patches in the scene. This mismatch at one spatial channel means less suppression, thus more veridical representation for that scale, in turn causing a more veridical percept. The BHLAW process assures that the data remain anchored at white.

**Configuration effect:** The Configuration effect says that, when a display contains gray surfaces arranged in a Mondrian, a wider range of lightnesses is perceived than when the same gray surfaces are arranged in a luminance staircase. Figures 6E and 6F summarize model simulations. The model explains this effect much as it does the Articulation effect: In the Mondrian configuration, since the intermingled luminance patches are arranged in a more radially compact way, the round-shaped surround kernels in the contrast module are influenced by more luminances of surrounding surfaces compared with the staircase arrangement. This
gives the surround kernels more chance to set the local means (surround activities) to be different from the corresponding center activities. Thus, the increased range of differences between the center and the surround activities results in a bigger range of perceived reflectances for the display. Explained otherwise, if all the adaptation and contrast stage surround activities were the same as their center activities, surround inhibition would drive them all to zero. The radially compact arrangement decreases the distance between different levels of gray patches, thereby inducing stronger lateral inhibition. The dependence of the distance between an inducer and test surfaces has been observed in lightness (Newson, 1958) and brightness experiments (Cole & Diamond, 1971; Fry and Alpern, 1953; Leibowitz et al., 1953), where the darker test surface became lighter with increasing distance from the inducer, an effect interpreted to be due to surround inhibition. Again the BHLAW process anchors the perception of white.

**Insulation effect:** The Insulation effect of Figure 7 shows that, when the staircase display is insulated by a white surround, the range of its perceived reflectance widens. Figures 7D and 7E show the data and simulation results, respectively. According to the model, a spatial contrast explanation also helps to explain this effect: Insulation of gray surfaces with a white surround causes bigger surround inhibition by the introduced bright insulation on the gray surfaces, making them look darker. This results in an expansion of the range of lightness due to the newly added suppression on dark patches by the surround. Insulation by a black surround, however, may not cause much difference in lightness assessment. This is because the gray surfaces are under illumination 30 times that of the background. Since the gray patches are already not getting much background inhibition, the introduction of black insulation does not significantly change the amount of surround inhibition, thus hardly changing the percept. Once again, the BHLAW process converts these relative lightness activities to an absolute anchored lightness.

**Area effect:** The Area effect in Figure 8 shows that, in a two-field Ganzfeld situation, as an area other than the area of highest luminance becomes larger than the half of the visual field, its lightness approaches white while the highest luminance area is pushed above white. Figure 8C shows the simulation of this effect. Comparison of data with the simulation shows that the model closely fits the suggested effect. As explained in the Section 2.5, self-luminosity of a small highest luminance area is explained by the BHLAW rule: When the highest luminance area is smaller than the blurring kernel at the anchoring stage, the blurred filled-in surface signals will have shallower highest activities compared to the un-blurred image (Figures 4B). Since the BHLAW mechanism uses the blurred signals to anchor lightness, the anchored lightness of the highest luminance area will look lighter than white. The case in Figure 4B corresponds to the increasing portions of curves in Figures 8B and 8C. The case in Figure 4A corresponds to the flat regions of the curves in Figures 8B and 8C.
4. Discussion
The model developed herein integrates known neuroanatomy, electrophysiology, and psychophysics to clarify how the brain generates a representation of surface lightness. The following discussion analyzes the model’s assumptions and limitations.

4.1 Retinal Adaptation
The model simulates the retinal adaptation using two mechanisms. First, at the outer-segment of the photoreceptor, its sensitivity to light is controlled by concentrations of chemicals, such as Ca\textsuperscript{2+} ions, that reflect the photoreceptor’s spatio-temporal history of experience of visual stimuli (Koutalos & Yau 1996). Second, a simulated inhibition at the inner-segment of the photoreceptor by the horizontal cells (HCs) was used to approximate HC modulation of glutamate release at the synaptic terminal of the photoreceptor (Fahrenfort et al., 1999; Verweij et al., 1996). This second mechanism instantiates the hypothesis that the communication among HCs through gap junction contributes to spatial contrast adaptation. The permeability of HC gap junctions is known to be modulated by various mechanisms, including neurotransmitters (DeVries & Schwartz, 1989, 1992; McMahon, 1994; Xin & Bloomfield, 2000), and transjunctional voltage (Lu et al., 1999; Spray et al., 1979).

The model assumes that the permeability is governed by an intracellular mechanism, which is in turn controlled by the output of the presynaptic photoreceptor. For example, for two HCs connected by a gap junction, the permeability of the junction decreases as the difference increases between the inputs that the HCs receive from the photoreceptors (Figure 10). Such a model retina can properly rescale inputs that have too much contrast, such as the one in Figure 20A. Figure 20D shows the steady-state HC activities for the input 20A. The dark and light image regions hereby deliver different suppressive feedback signals to the photoreceptors. Figure 20C illustrates the two sensitivity curves of the inner-segments of the photoreceptors caused by two different negative feedback levels of the HC network for the image. Using these two sensitivities, the network can properly rescale the response of the output at the inner-segment of the photoreceptor, which could have mapped to be too low or high in response if it used just one sensitivity, as illustrated in Figure 20B. The rescaled steady-state output of the photoreceptor inner-segments are shown in Figure 20E. The output of the model photoreceptor in Figure 20F shows visible dark and light image regions. Figures 20G, 20H and 20I show a simulation without the HC gating mechanism. The adapted signals in 20H and the output 20I show signal distortion (a halo) along the border of the dark and light parts, and the dark part is less visible. Figures 20J, 20K and 20L show a simulation with no diffusion among HCs. The results show a prominent compression of signals.
(A) STIMULUS (B) NO CONTRAST ADAPTATION (C) CONTRAST ADAPTED

(D) HC ACTIVITIES (E) RETINAL ADAPTED (F) LIGHTNESS

(G) HC ACTIVITIES (H) RETINAL ADAPTED (I) LIGHTNESS

(J) HC ACTIVITIES (K) RETINAL ADAPTED (L) LIGHTNESS
In summary, in addition to the adaptation at the outer-segment of the photoreceptor, which shifts the sensitivity curve of the photoreceptor, the HC negative feedback further shifts photoreceptor sensitivity to be sensitive to the spatial context of input contrasts. This HC feedback is not a compressive process; rather, it shifts the sensitivity curve of the inner-segment of the photoreceptor. See the mathematical details in Appendix A.

HC receptive field size change due to negative feedback between the photoreceptor and the HC was proposed by Kamermans et al. (1996). In their model, they emphasize the contribution of the negative feedback on determining the length constant of the HCs. Inclusion of such a mechanism may help further explain the dynamics of retinal adaptation. The model also does not simulate the cone-rod circuitry switch (Mills & Massey, 1995; Ribelayga, Wang & Mangel, 2002) and the pupillary light reflex (Dowling, 1987), which are known to contribute to adaptation. These refinements were not needed to simulate this article’s targeted data.

4.2 Multiple-Scale Filtering

The model interprets the neuroanatomy of the retina as the initial source of multiple-scale representation whereby center-surround processes shape the outputs of ganglion cells having diverse receptive field sizes (Masland, 2001; Roska et al., 2000; Werblin, 2001). Direct evidence of multiple-scale representation within V1 includes the fact that cell optimal spatial frequencies gradually increase as their positions move away from blob centers (De Valois & De Valois, 1988; Edwards et al., 1995). Issa et al. (2000) also show gradual changes in cell spatial frequency preferences that conform to the hypercolumn cortical organization in V1. Data about cell responses in LGN, V1 and V2 to uniform surface luminance and color also support the existence of large spatial scales (Bartlett & Doty, 1974; Friedman et al., 2003; Komatsu, 2001; Marrocco, 1972; Papaioannou & White, 1972).

The model hypothesis that cells in the blob stream pool their multiple-scale representations has not yet been directly tested. The pooling of ON and OFF signals, however, is consistent with the finding that the segregation of ON and OFF channels from the retina and LGN, and their projection to layer 4 in V1 (for a review, see Schiller, 1992), is largely lost in the cortex of the ferret (Chapman & Gödecke, 2002).
4.3 Boundary Representation

For simplicity, the model does not implement the boundary completion property of the visual system (Field, Hayes & Hess, 1993; Grossberg & Mingolla, 1985a, 1985b; Grossberg & Raizada, 2000; Kellman, & Shipley, 1991; Raizada & Grossberg, 2001; von der Heydt, Peterhans & Baumgartner 1984). Incorporation of this property into the model would explain more psychophysical data, such as boundary grouping properties like illusory contours (Gove, Grossberg & Mingolla, 1995; Grossberg & Mingolla, 1985a), 3-D figure-ground separation (Grossberg, 1994; Kelly & Grossberg, 2000), and surface noise suppression (Grossberg et al., 1995; Mingolla et al., 1999). Surface noise suppression exploits long-range boundary completion by bipole cells to group noisy pixels into coherent boundaries. Filling-in across noisy surface signals is contained within these boundaries, thus forming a noise-free surface. For figures that do not require significant boundary completion, surface noise suppression can be achieved by the present model’s simplified processing. Figure 21D uses a smaller boundary-gating parameter than for Figure 21C.

Figure 21. Noise suppression property of the model. (A) Input with Gaussian noise of signal-to-noise ratio of 10 dB. (B) Boundary signals at input A. Despite the disruptive noises, boundary signals still show coherent representation of the edge signals. (C) Output of the model without parameter change. The model does not show much of noise suppression property. (D) Output of the model with a smaller gating parameter ($\varepsilon = 20$). With a smaller gating parameter, the model shows noise suppression property. See the text for further discussion.
4.4 Filling-In
The model proposes a mechanism called Gated Blurring for the filling-in process. Unlike previous filling-in models that use nearest-neighbor diffusion as the cortical mechanism, the current model uses intralaminar propagation via horizontal long-range connections with boundary-gating signals that block signal propagation across luminance edges (Figure 13). This long-range connectivity of the model is consistent with the known horizontal connections in the visual cortices (Angelucci et al., 2002; Gilbert & Wiesel, 1979; Rockland & Lund, 1982; Stettler et al., 2002; Yabuta & Callaway, 1998). The gating mechanism, by selectively allowing communication between only certain connections, can be viewed as a process of dynamic restructuring of neuron receptive fields. One possible mechanism is axo-axonal gating mechanism of the horizontal connections, which is consistent with the report by Kobayashi et al. (2000) of norepinephrine-mediated suppression of horizontal propagation in V1. The mechanism runs at least one hundred times faster than previous nearest-neighbor-based diffusion models, and thus clarifies how filling-in can occur with realistic delays. For example, ten iterations of the non-diffusive filling-in process were used to generate the filled-in image of Craik-O’Brien-Cornsweet effect in Figure 18B. For the long-range diffusive mechanism, 100 iterations were necessary. With the previous diffusion mechanism, which does not have long-range connectivity, about 10,000 iterations were needed to get an output of the same quality. Overall, the model is 100 times faster than the previous models in terms of the filling-in, including the retinal long-range filling-in.

4.5 Lightness Anchoring
The model assumes that lightness anchoring process happens after the filling-in stages in V2 or V4. A recent electrophysiological experiment by MacEvoy and Paradiso (2001) reported lightness constancy in V1. However, their experiment does not provide unequivocal evidence that V1 is the place where anchoring occurs. It demonstrates just one aspect of lightness perception; namely discounting the illuminant, or input normalization, which can be achieved at the adaptation and contrast stages in the model. Even though the model analyses lightness percepts mainly in terms of luminance-based processes, other factors, such as long-range grouping, which take place in V2 (von der Heydt, Peterhans & Baumgartner 1984; Peterhans & von der Heydt, 1989) may also influence the process (see Gilchrist et. al., 1999 for further discussion). The need for extrastriate involvement in the lightness anchoring process also comes from the fact that global integration of information, which is needed for the BHLAW rule, may need a bigger scale of interaction than the one supported by the horizontal connections in V1 (Angelucci et al., 2002). Areas V2 or V4 are probable places for anchoring that satisfy the need for large-scale integration of surface information. V2 provides a rich environment for the
boundary system (interstripes) and luminance and contrast signals (thin stripes) to interact (Roe and Ts'о, 1995) to begin to form surface percepts. The data of Hung et al. (2001) showing a prominent Craik-O’Brien-Cornsweet effect in V2 are also compatible with this assumption.

4.6 Area Effect in Natural Images

The area effect tends to be limited to simple Ganzfeld configurations. Gilchrist and his colleagues (1999, p. 802) note: “Strictly speaking, the rule applies to visual fields composed of only two regions of nonzero luminance. Application of the rule to more complex images remains to be studied.” In the model, it is assumed that when the simple Ganzfeld configuration was tested, the visual system of the subject adapted its multiple scales to compensate for the unusually sparse visual cues. In particular, Sections 2.2 and 4.2 noted that the model incorporates multiple spatial scales which suppress signals that are uniform with respect to each scale. Hence, given the sparse contrasts in the Ganzfeld display, the model would be expected to suppress small scales. Multiple scales were not used in the anchoring module, for simplicity. Instead, two different parameter sets were used to explain the area rule: For simple images having just two regions of non-zero luminance (Figure 8), a bigger Gaussian kernel was used. For all the other, more complex, images with smaller regions, a smaller kernel was applied. See Table 1 for parameters. The sizes of the two anchoring kernels were chosen that best fit the sets of data suggested in the Anchoring theory by Gilchrist et al. (1999). Automatic rescaling of the anchoring process will be incorporated when the model fully exploits its multiple scales for purposes of 3-D vision and figure-ground perception; see Grossberg (1994, 1997) for a discussion of how multiple scale are used in 3-D vision.
APPENDIX A: MODEL EQUATIONS

The model implements 2-D simulations on a 200 x 200 grid that represents the whole visual field.

Retinal adaptation

The potential $s_{ij}$ at position $(i,j)$ of the outer segment of the retinal photoreceptor is simulated by the equation:

$$s_{ij}(t) = I_{ij} \cdot z_{ij}(t), \quad (A1)$$

where $I_{ij}$ is the input and $z_{ij}(t)$ is an automatic gain control term simulating negative feedback mediated by Ca$^{2+}$ ions, among others:

$$\frac{dz_{ij}}{dt} = (B_z - z_{ij}) - z_{ij} \left( C_I I_{ij} + C_T \bar{I} \right), \quad (A2)$$

(cf., Carpenter and Grossberg, 1981; Grossberg 1980). In (A2), parameter $B_z$ is the asymptote which $z_{ij}(t)$ approaches in the absence of input, and term $-z_{ij}(C_I I_{ij} + C_T \bar{I})$ describes the inactivation of $z_{ij}$ by the present input $I_{ij}$ and a spatial average $\bar{I}$ of all inputs that approximates the effect of recent image scanning by sequences of eye movements. The equilibrium response $s_{ij}$ directly follows from (A1) and (A2):

$$s_{ij} = \frac{B_z I_{ij}}{1 + C_I I_{ij} + C_T \bar{I}}. \quad (A3)$$

The inner segment of the photoreceptor receives the signal $s_{ij}$ from the outer segment and also gets feedback $H_{ij}$ from the horizontal cell (HC) at position $(i,j)$, as in Figure 10. HC modulation of the output of the inner segment of the photoreceptor is modeled by the equation:

$$S_{ij} = \frac{s_{ij}}{B_h \exp(H_{ij}) \cdot (B_s - s_{ij}) + 1}, \quad (A4)$$

where $B_h$ is a small constant, and $B_s$ is a constant close to the value $(B_z / C_I)$. When $B_s$ equals the value of $(B_z / C_I)$, perfect shifts of log($I_{ij}$) - $S_{ij}$ curve occur with varying $H_{ij}$ (Figure A1A). When $B_s$ deviates from $(B_z / C_I)$, compression occurs when $B_s > (B_z / C_I)$ (Figure A1C). Expansion occurs when $B_s < (B_z / C_I)$ in addition to the shift. Thus to prevent expansion, which would mean excitation by the HC negative feedback, $B_s$ needs to be bigger or equal to $(B_z / C_I)$. Figure A2 shows the 10-Mondrian Articulation situation (see Figure 5) with two values for $B_s$, one equals to $(B_z / C_I)$, and the other to 1.2$(B_z / C_I)$. This simulation demonstrates that the model is robust under this variation. Compare Figure A2 with the graph in Figure 5G.

The equation (A4) can be generalized as follows.

$$S_{ij} = \frac{s_{ij}}{f(H_{ij}) \cdot (B_s - s_{ij}) + 1}. \quad (A4')$$
Many increasing functions $f(H_{ij})$ will generate the shift property of $S_{ij}$ as a function of $\log(I_{ij})$. Function $f(H_{ij}) = B_h \exp(H_{ij})$ was chosen because $\exp(H_{ij})$ makes the sensitivity curve shift in an accelerating manner with increasing $H_{ij}$, where $H_{ij}$ is the sigmoid output of the HC at $(i, j)$ in response to its potential $h_{ij}$:

$$H_{ij} = \frac{a_H h_{ij}^2}{b_H^2 + h_{ij}^2},$$

where $a_H$ and $b_H$ are constants. This bounded function causes the amount of shift to decrease as $h_{ij}$ becomes large. The combination of the initial acceleration by the exponential function in the equation (A4) and the later saturation by the equation (A5) causes the $S_{ij}$ curve to accelerate.
initially and later decelerate with increasing $h_{ij}$. Figure (A1A) shows an example of this shift property. The leftmost curve represents the $S_{ij}$ curve with $h_{ij} = 0$; the other curves have $h_{ij}$ values of 0.1, 0.2, ..., 0.5, respectively. All these curves have the same average luminance $\bar{I} = 10^2$. The shift property is generated at any average luminance $\bar{I}$. Note that the leftmost curve in Figure (A1A) is the same as the curve with $\bar{I} = 10^2$ in Figure 2C. Figure (A1B) shows what happens when $H_{ij} = h_{ij}$ is used in stead of equation (A5), with all other equations the same; it shows no deceleration. Here, $h_{ij}$ values of 0 to 10 were used with increments of 1. Figure (A1C) shows a situation where the term $(B_s - s_{ij})$ in equation A4 has been replaced by 1; it shows a prominent compression. For this simulation, $h_{ij}$ values of 0 to 0.5 with increments of 0.1 were used. Figure (A1D) shows a situation with $f(H_{ij}) = H_{ij}$ in equation A4'; it does not have the smooth acceleration shown in Figure (A1A). The same $h_{ij}$ values as for Figure (A1C) were used for this simulation.

The potential of an HC connected to its neighbors through gap junctions is defined as follows.

\[
\frac{dh_{ij}}{dt} = -h_{ij} + \sum_{(p,q) \neq h_{ij}} P_{pqij} (h_{pq} - h_{ij}) + S_{ij},
\]

(A6)

where $P_{pqij}$ is the permeability between cells at $(i, j)$ and $(p, q)$; namely,

\[
P_{pqij} = \frac{-1}{1 + \exp\left[-\left(|S_{ij} - S_{pq}| - \beta_p \right) / \lambda_p \right]} + 1
\]

(A7)

Figure A2. Robustness of the model. The curves show ten-Mondrian Articulation situation with two values for $B_s$, one ($B_s/C_i$), the other 1.2($B_s/C_i$). While the deviation of 20% from the optimal value shows a bit of compression, the overall quality of Articulation effect remains robust. This demonstrates that the model tolerates a fair amount of fluctuation in the value of the parameter.
Terms $\beta_p$ and $\lambda_p$ in (A7) are constants, and $N^H_{ij}$ in (A6) is the neighborhood of size $\epsilon_H$ to which the model HC at $(i, j)$ is connected:

$$N^H_{ij} = \{(p, q) : \sqrt{(i - p)^2 + (j - q)^2} \leq \epsilon_H \text{ and } (p, q) \neq (i, j)\}$$  \hspace{1cm} (A8)

**Center - Surround Stage**

The retinally adapted signal $S_{ij}$ is then processed by small-scale and medium-scale on-center off-surface and off-center on-surface networks. In the following, scale subscripts (e.g., $x_s$ and $x_m$ for small and medium scales, respectively) are omitted for simplicity. An on-center off-surface (ON) network of cell activities $x_{ij}^+$ that obey membrane equations is defined as follows:

$$\frac{dx_{ij}^+}{dt} = -Ax_{ij}^+ + (B - x_{ij}^+)C_{ij} - (x_{ij}^+ + D)E_{ij},$$  \hspace{1cm} (A9)

where $A, B$ and $D$ are constants. The on-center input obeys:

$$C_{ij} = \left( \sum_{(p,q) \in N^C_{ij}} S_{pq} C_{pqij} \right) \frac{W_C}{\sum_{(p,q) \in N^C_{ij}} C_{pqij}},$$  \hspace{1cm} (A10)

and the off-surround input obeys:

$$E_{ij} = \left( \sum_{(p,q) \in N^E_{ij}} S_{pq} E_{pqij} \right) \frac{W_E}{\sum_{(p,q) \in N^E_{ij}} E_{pqij}},$$  \hspace{1cm} (A11)

with the excitatory Gaussian on-center kernel:

$$C_{pqij} = C \exp \left\{ -\frac{(p-i)^2 + (q-j)^2}{\alpha^2} \right\}$$  \hspace{1cm} (A12)

and the inhibitory Gaussian off-surround kernel:

$$E_{pqij} = E \exp \left\{ -\frac{(p-i)^2 + (q-j)^2}{\beta^2} \right\}.$$  \hspace{1cm} (A13)

Coefficients $C$ and $E$ in (A12) and (A13), which normalize and make the sums of the center and surround kernels the same, are defined by:

$$C = \frac{W_C}{\sum_{(p,q) \in N^C} \exp \left\{ -\frac{p^2 + q^2}{\alpha^2} \right\}}$$  \hspace{1cm} (A14)

and

$$E = \frac{W_E}{\sum_{(p,q) \in N^E} \exp \left\{ -\frac{p^2 + q^2}{\beta^2} \right\}}.$$  \hspace{1cm} (A15)

Terms $\alpha, \beta, W_C$ and $W_E$ are constants. $N^C_{ij}$ in equation (A10) is the on-center neighborhood to which the cell at $(i,j)$ is connected:
\begin{align*}
N_{ij}^C &= \left\{ (p, q) : \sqrt{(i-p)^2 + (j-q)^2} \leq \varepsilon_C \text{ and } 0 \leq p \leq 199 \text{ and } 0 \leq q \leq 199 \right\}, \\
\text{where } \varepsilon_C \text{ is a constant defining the size of the neighbor. } \quad (A16)
\end{align*}

\text{The only difference between } N_{ij}^C \text{ and } N^C \text{ is that } N_{ij}^C \text{ is constrained by the boundary of the image (200x200), which may cut kernels along the borders, while } N^C, \text{ which defines the whole kernel, is not. For brevity, the same convention between } N_{ij}^C \text{ and } N^C \text{ is used for other equations as well. For example, } N_{ij}^E \text{ in equation (A11) is the neighborhood for the surround kernel with a size } \varepsilon_E \text{ with the same form of definition as equation (A16), and its corresponding standard neighbor is } N^E \text{ with the same form of definition as equation (A17). See Table 1 for parameters.}

\text{For each position, the normalizing factors } W_C / \Sigma C_{pqij} \text{ and } W_E / \Sigma E_{pqij} \text{ in (A10) and (A11) are constants, mostly just 1, except for the positions along the border of the image. Normalization eliminates unwanted boundary effects created by filters with a fixed kernel size. In case of a center-surround filter, for example, without normalization, halos along the border of the image can occur because of the disinhibition caused by cut kernels there.}

\text{The equilibrium activities of (A9) are:}
\begin{equation}
x_{ij}^+ = \frac{BC_{ij} - DE_{ij}}{A + C_{ij} + E_{ij}}. \quad (A18)
\end{equation}

\text{The corresponding equilibrium activities of the off-center on-surround (OFF) network are:}
\begin{equation}
x_{ij}^- = \frac{BC_{ij}^- - DE_{ij}^-}{A + C_{ij}^- + E_{ij}^-}. \quad (A19)
\end{equation}

\text{In (A19),}
\begin{equation}
C_{ij}^- = E_{ij} \quad (A20)
\end{equation}

\text{and}
\begin{equation}
E_{ij}^- = C_{ij} \quad (A21)
\end{equation}

\text{(Grossberg, Mingolla, and Williamson, 1995). The output signals are rectified versions of } x_{ij}^+ \text{ and } x_{ij}^- :}
\begin{equation}
X_{ij}^+ = \left[ x_{ij}^+ \right]^+ \quad (A22)
\end{equation}

\text{and}
\begin{equation}
X_{ij}^- = \left[ x_{ij}^- \right]^+. \quad (A23)
\end{equation}

\text{Luminance signals } L_{ij}, \text{ which constitute the large-scale of the center-surround process, are defined by:}
\begin{equation}
L_{ij} = S_{ij} \quad (A24)
\end{equation}

\text{Through these processes, the initial stage of the model achieves automatic gain control in all its small, medium and large scales.}
Boundary System

Simple cell activities are simulated using a network of units having polarized and oriented receptive fields around a grid of pixel units. Figure (A3A) shows pixel units at \((i, j)\) denoted as small filled circles, and eight surrounding numbered positions at \((i', j')\) where pairs of model simple cells with the same orientation but opposite contrast polarity are located. Each simple cell is represented by a half-filled and half-hollow oriented ellipse (Figure A3B). The eight positions are as follows: \((i + 0.5, j), (i + 0.5, j + 0.5), (i, j + 0.5), (i - 0.5, j + 0.5), (i - 0.5, j), (i - 0.5, j - 0.5), (i, j - 0.5), (i + 0.5, j - 0.5)\). A pair of simulated simple cells has one of 4 orientations: \((0, \pi/4, \pi/2, 3\pi/4)\). The even numbered positions have only two \((0, \pi/2)\) orientations; positions 3 and 7 have three orientations \((0, \pi/4, 3\pi/4)\); and positions 1 and 5 have three orientations \((\pi/4, \pi/2, 3\pi/4)\). The responses of simple cells are modeled using medium-scale contrast signals. This simplification was chosen because it gives relative clean edge signals. The outputs from simple cells having light-dark and dark-light luminance polarities in their receptive fields are simulated as follows:

\[
\begin{align*}
  s_{LD}^{ij} & = \left[ (L_{L}^{+} + R_{L}^{-}) - (R_{L}^{+} + L_{L}^{-}) \right]^{+} \\
  s_{DL}^{ij} & = \left[ (R_{L}^{+} + L_{L}^{-}) - (L_{L}^{+} + R_{L}^{-}) \right]^{+},
\end{align*}
\]  

where the superscripts \(LD\) and \(DL\) indicate light-dark and dark-light luminance polarities of the model simple cell receptive fields, respectively, and \(k\) denotes the orientation. Activation of a model simple cell left and right sub-receptive fields from ON and OFF channels is modeled as follows:

\[
\begin{align*}
  L_{L}^{+} & = \left( \sum_{(p,q) \in N_{ij}^{L}} X^{m+}_{pq} G_{pqi'j',KL} \right) \frac{W_{B}}{\sum_{(p,q) \in N_{ij}^{L}} G_{pqi'j',k}} \\
  R_{L}^{+} & = \left( \sum_{(p,q) \in N_{ij}^{L}} X^{m+}_{pq} G_{pqi'j',KR} \right) \frac{W_{B}}{\sum_{(p,q) \in N_{ij}^{L}} G_{pqi'j',k}} \\
  L_{R}^{+} & = \left( \sum_{(p,q) \in N_{ij}^{R}} X^{m-}_{pq} G_{pqi'j',KL} \right) \frac{W_{B}}{\sum_{(p,q) \in N_{ij}^{R}} G_{pqi'j',k}} \\
  R_{R}^{+} & = \left( \sum_{(p,q) \in N_{ij}^{R}} X^{m-}_{pq} G_{pqi'j',KR} \right) \frac{W_{B}}{\sum_{(p,q) \in N_{ij}^{R}} G_{pqi'j',k}}.
\end{align*}
\]  

Subscripts \(L\) and \(R\) indicate the two sub-receptive fields for the simple cell with \(L\) indicating the left part (to the anticlockwise) of the sub-receptive field, and the \(R\) the right part (to the
clockwise) of the sub-receptive field along the axis of the orientation. Constant $W_B$ is the sum of the standard kernel weights of the simple cell:

$$W_B = \sum_{(p,q)\in\mathbb{N}^2} G_{pq,i',k} \cdot,$$

(A31)

At each position, the normalization factor $WB / \sum G_{pq,i',k}$ is constant, mostly just 1, except for positions along the border of the image where the Gaussian kernel is incomplete. To see the size of the simple cell kernel neighbor, $N^B$, see $\varepsilon_B$ in Table 1.

A pair of oriented Gaussian kernels, indicated as $L$ and $R$, simulates receptive fields for the simple cell:

$$G_{pq,i',k,(L,or,R)} = \kappa \exp\left\{-\left[\frac{(p-i') \cos(\pi k/4) + (q-j') \sin(\pi k/4)}{\gamma_h^2}\right)^2 - \left[\frac{-(p-i') \sin(\pi k/4) + (q-j') \cos(\pi k/4) + \text{Shift}_{(L,or,R)}}{\gamma_v^2}\right)^2\right\},$$

(A32)

where $\text{Shift}_{(L)}$ and $\text{Shift}_{(R)}$, which shift the sub-receptive fields orthogonal to the axis of orientation, are constants $-\gamma_v$ and $\gamma_v$, respectively; $\kappa$ is a constant. $k$ is one of the four numbers (1, 2, 3, 4) that sets the orientation; and $\gamma_h$ and $\gamma_v$ are constants that define the widths of the kernel along and across the axis of orientation, respectively.

The model complex cells are also located at the eight $(i', j')$ positions, and have oriented receptive fields, as illustrated in Figure (A3C). The model complex cell of orientation $k$ at $(i', j')$ pools the outputs of a pair of simple cells as follows:

$$z_{r,j,k} = s_{r,j,k}^{LD} + s_{r,j,k}^{DL}.$$

(A33)

This cell potential goes through an activation function:

$$Z_{r,j,k} = f(z_{r,j,k}),$$

(A34)

where

$$f(x) = \frac{a_B x^{1.7}}{b_B^2 + x^{1.7}}.$$

(A35)

The parameter 1.7 of the power of $x$ was in (A35) used that gave the optimal strength of the boundary signals across simulations. The complex cell gates any horizontal connections that cross its gating field. The effective gating strength at a point $(x, y)$ along a passing horizontal connection is the product of the gating weight $(G^c_{xy,i',k})$ at the point and the activation of the gating complex cell at $(i', j')(Z_{i',j'})$:

$$Z_{x,y,i',j'} = G^c_{xy,i',j'} Z_{i',j'},$$

(A36)

where $x, y$ are continuous variables. The Gaussian kernel of the gating field, which represents the spatial spread of gating weight of complex cell axons at points $(x, y)$ along the line $(i, j) - (p, q)$, is defined as follows:
\[
G^c_{xyij/k} = \exp\left\{-\frac{\left[(x - i')\cos(\pi k / 4) + (y - j')\sin(\pi k / 4)\right]^2}{\gamma_{ch}^2}
- \frac{\left[-(x - i')\sin(\pi k / 4) + (y - j')\cos(\pi k / 4)\right]^2}{\gamma_{cv}^2}\right\}, \quad (A37)
\]

Figures (A3D) and (A3E) show an example of the complex cell gating mechanism for a given input. For a given complex gating field, it is assumed that the gating occurs at just one point for each crossing connection. The gating point \((x, y)\), which lies along the line \((i, j) - (p, q)\), is chosen that gives the maximum value of equation (A37). In the simulation, 10 equidistance points along the cross-section between the ellipse and the crossing line \((i, j) - (p, q)\) were examined to find the approximate inflection (maximum) point as shown in Figure (A3F). The size of each dot in the figure represents the value \(G^c_{xyij/k}\) of equation (A37) for each examined point.

![Diagram](image)

Figure A3. Model boundary system and gating mechanism. (A) Relative positions of model simple and complex cells to pixel points. The model simple and complex cells are poisoned between the pixel points. For example, for a given pixel in the middle (the small gray filled circle in the middle), there are eight surrounding positions (1 through 8) where simple and complex cells are placed. (B) Configuration of simple cell network around a pixel unit in the middle. Just one set of pixel-complex and simple cell relationship is shown for clarity. The same pixel-complex and simple cell relationship applies to other pixels. (D) Example of a stimulus. (E) Illustration of gating mechanism for stimulus D. It illustrates the resulting activations of gating components with the input in figure D. The activated complex cells that surround the disk area gate any connections crossing their gating fields represented as ellipses. The connection between \((i, j)\) and \((p, q)\) is gated (the dotted line) by a gating signal at \((x, y)\) in the gating field of the complex cell centered at \((i', j')\). The other connections are not gated, being allowed to have high conductances (solid lines). For the purpose of illustration, more orientations are shown than the four orientations used for the simulations. (F) Position of the gating point. The figure shows the blown up part of the gated part of the connection in the figure E. In the simulation, 10 equidistance points (5 of them are shown for clarity) along the cross-section between the ellipse and the crossing line \((i, j) - (p, q)\) were examined to find the approximate inflection (maximum) point. The size of each dot represents the value \(G^c_{xyij/k}\) of equation (A37) for each examined point.
**Filling-in**

Cortical filling-in is driven by the inputs $M_{ij}$ which are the pooled luminance and contrast signals as follows:

$$M_{ij} = \left[ w_s(X_{ij}^x - X_{ij}^z) + w_m(X_{ij}^m - X_{ij}^z) + w_iL_{ij} + b_M \right],$$  \hspace{2cm} (A38)

where $w_s$, $w_m$, $w_i$, are weighting constants, and $b_M$ is a tonic bias term. Either of two versions of the filling-in process yield equivalent simulations of the targeted data. A long-range diffusion process, much as in the retinal HC diffusion in (A6), works well with activities $F_{ij}$ instead of the activities $h_{ij}$ in (A6), and inputs $M_{ij}$ instead of the inputs $S_{ij}$ in (A6). This long-range diffusion runs 100 times faster than previous nearest-neighbor diffusions for filling-in. In addition, the conductance $P_{pqij}$ are divisively gated by activated complex cells along its path. They are defined by:

$$P_{pqij} = \frac{\delta \exp[-((i-p)^2 + (j-q)^2)/\sigma^2]}{\prod_{i'j'k}(1 + \varepsilon Z_{i'j'k})},$$  \hspace{2cm} (A39)

where $\sigma$, $\delta$ and $\varepsilon$ are constants. The numerator of (A39) describes the strengths of horizontal connections, assumed to have a Gaussian distribution, such that longer connections have smaller strengths.

Alternatively, a long-range propagation process that does not require diffusion, but is normalized in a different way, generates essentially identical simulations, which are the ones that are shown in this article. This process runs 1000 times faster than nearest-neighbor diffusion processes. The first step of the filling-in is to activate the filling-in units with the pooled multiple-scale input signals $M_{ij}$:

$$F_{ij} = M_{ij}.$$  \hspace{2cm} (A40)

Here, the filling-in activity $F_{ij}(t+1)$ equals:

$$F_{ij}(t+1) = \left( \sum_{(p,q)\in N^F_{ij}} F_{pq}(t)P_{pqij} \right) \frac{W_F}{\sum_{(p,q)\in N^F_{ij}} P_{pqij}},$$  \hspace{2cm} (A41)

where the conductance $P_{pqij}$ shares the same form of equation (A39) with different parameters (see Table 1). The constant $W_F$ in (A41) is a sum of conductances defined as follows:

$$W_F = \sum_{(p,q)\in N^F} \frac{\delta \exp[-((i-p)^2 + (j-q)^2)/\sigma^2]}{\prod_{i'j'k}(1 + 0)}.$$  \hspace{2cm} (A42)

Since $W_F$ is constant, the constant $\delta$ for a fixed $\sigma$ is calculated as follows:

$$\delta = \frac{W_F}{\sum_{(p,q)\in N^F} \exp[-((i-p)^2 + (j-q)^2)/\sigma^2]}.$$  \hspace{2cm} (A43)

The size of the filling-in neighborhood $N^F$ is determined by parameter $\varepsilon_F$ in Table 1. Equation (A41) assumes that the filling-in unit can normalize its conductances. The normalizing factor
\( W_F / \Sigma P_{pqij} \) affects the conductance in two ways. First, at the border of the image, the incomplete kernels get normalized to have the same size as \( W_F \). Second, normalization compensates for the overall lost conductance caused by gating (division by the denominator in equation (A39)). By this normalization process, the sum of the effective conductances equals:

\[
W_F = \sum_{pq} \left( P_{pqij} \frac{W_F}{\sum_{pqj} P_{pqij}} \right).
\]

(A44)

For example, if half of the input connections were totally blocked by gating signals, the unit would try to increase the effective input flow by doubling the efficacy of the remaining connections, keeping the sum of all the incoming conductances the same. Ten iterations of equation (A41) gives satisfactory filled-in results.

**Lightness Anchoring**

At the anchoring stage, the filled-in surface activity \( F_{ij} \) becomes anchored into the activity \( A_{ij} \) using the following equation:

\[
\frac{dA_{ij}}{dt} = -B_A A_{ij} + \Psi (C_A - A_{ij})F_{ij},
\]

(A45)

where \( B_A \) and \( C_A \) are constants. The tonic gain control signal \( \Psi \), which modulates all the anchoring activities \( A_{ij} \), uses the following equation:

\[
\frac{d\Psi}{dt} = \tau_\Psi \{- \Psi + (B_\Psi - \Psi)T_\Psi - \Psi H\}.
\]

(A46)

The term \( \tau_\Psi \) is a time constant that determines the speed of integration of equation (A46). The term \( -\Psi \) is a leakage component. The next term \( (B_\Psi - \Psi)T_\Psi \) is an excitatory component that drives the gain control signal \( \Psi \) toward its maximum \( B_\Psi \) until the inhibitory component \( \Psi H \) kicks in due to the activation of the suppressive signal \( H \), which is defined as follows:

\[
\frac{dH}{dt} = \tau_H \left\{-H + (\varphi - H)\sum_{ij} B_{ij} \right\},
\]

(A47)

where \( \tau_H \) is a time constant. Using the equation (A47), the suppressive signal \( H \) quickly becomes activated and suppresses the gain control activity \( \Psi \) whenever there is an activated output cell at the BHLAW module, which signals the anchoring of blurred “highest luminance” to white. The output of the BHLAW module \( B_{ij} \) is defined as follows:

\[
B_{ij} = f^B(b_{ij}),
\]

(A48)

where the signal function \( f^B(x) \) is a steep sigmoid:

\[
f^B(x) = \frac{x^m}{\sigma^m + x^m},
\]

(A49)
where \( m \) and \( \sigma \) are constants; see Table 1. Function \( b_{ij} \) in (A48) is a blurred version of the anchoring signal \( A_{ij} \):

\[
b_{ij} = \sum_{(p,q) \in N^d} G^A_{pqij} A_{pq},
\]

where the blurring Gaussian anchoring kernel is defined by:

\[
G^A_{pqij} = G^A \exp \left\{ -\frac{(p - i)^2 + (q - j)^2}{\zeta_A^2} \right\} \sum_{(p,q) \in N^d} G^A_{pqij},
\]

where constant

\[
G^A = \frac{W_A}{\sum_{(p,q) \in N^d} \exp \left\{ -\frac{p^2 + q^2}{\zeta_A^2} \right\}},
\]

and \( W_A \) and \( \zeta_A \) are constants. The size of the blurring neighborhood \( N^d \) is determined by parameter \( \varepsilon_A \) in Table 1. When \( m \) in equation (A49) is large, \( H \) approximates a step function

\[
H = \phi \quad \text{ whenever any } \ b_{ij} \geq \sigma
\]

\[
H = 0 \quad \text{ otherwise,}
\]

where \( \phi \) is a constant. In the simulation, equation (A53) was used in place of equations (A47) to (A49).

**APPENDIX B**

To generate the stimuli with different background luminance (Figure 2C), the following formula was used:

\[
I_{ij} = \rho_{ij} E_{ij},
\]

where \( I_{ij} \) is the luminance at point \((i, j)\), \( \rho_{ij} \) is the reflectance at point \((i, j)\), and \( E_{ij} \) is the illumination on point \((i, j)\) (Hurlbert, 1989). For a given stimulus, \( E_{ij} \) was uniform across the image. For practical purposes, \( \rho_{ij} \) in equation (B1) was replaced by the luminance at point \((i, j)\) of the original image. This situation is roughly equivalent to a viewing situation where a picture is exposed to uniform background illumination. The range of \( \rho_{ij} \) was chosen to be \(-4\) to \(5\) in log-scale for a fixed illumination level to examine the full dynamic profile of the shift property. See Figure 2C for the values of illumination \( E_{ij} \) used for the simulations.
<table>
<thead>
<tr>
<th>Names</th>
<th>Symbols</th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>Upper bound of gain control at photoreceptor</td>
<td>$B_z$</td>
<td>500</td>
</tr>
<tr>
<td>Small-time scale input contribution rate for gain control</td>
<td>$C_I$</td>
<td>200</td>
</tr>
<tr>
<td>Large-time scale input contribution rate for gain control</td>
<td>$C_I$</td>
<td>600</td>
</tr>
<tr>
<td></td>
<td>$B_h$</td>
<td>0.04</td>
</tr>
<tr>
<td></td>
<td>$B_s$</td>
<td>$(B_s/C_I)$</td>
</tr>
<tr>
<td></td>
<td>$a_H$</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>$b_H$</td>
<td>0.1</td>
</tr>
<tr>
<td>Shift of permeability of HC gap junction</td>
<td>$\beta_p$</td>
<td>0.08</td>
</tr>
<tr>
<td>Steepness of permeability of HC gap junction</td>
<td>$\lambda_p$</td>
<td>0.01</td>
</tr>
<tr>
<td>Size of connected neighbor for horizontal cell</td>
<td>$\varepsilon_H$</td>
<td>8</td>
</tr>
<tr>
<td>Activation decay</td>
<td>$A$</td>
<td>0.5</td>
</tr>
<tr>
<td>Depolarization constant</td>
<td>$B$</td>
<td>1</td>
</tr>
<tr>
<td>Hyperpolarization constant</td>
<td>$D$</td>
<td>1</td>
</tr>
<tr>
<td>Center spatial scale for the center-surround stage</td>
<td>$\alpha$</td>
<td>0.2</td>
</tr>
<tr>
<td>Surround spatial scales, small, medium</td>
<td>$\beta$</td>
<td>3 (for small scale), 14 (for medium scale)</td>
</tr>
<tr>
<td></td>
<td>$\kappa$</td>
<td>4</td>
</tr>
<tr>
<td>Vertical, horizontal widths of the ON, OFF elliptic simple cell receptive fields</td>
<td>$\gamma_v, \gamma_h$</td>
<td>0.1, 5$\gamma_v$,</td>
</tr>
<tr>
<td>The shift of the centers of the ON, OFF elliptic simple cell receptive fields</td>
<td>$\text{Shift}<em>(L), \text{Shift}</em>(R)$</td>
<td>$-\gamma_v, \gamma_v$</td>
</tr>
<tr>
<td></td>
<td>$a_B$</td>
<td>0.7</td>
</tr>
<tr>
<td></td>
<td>$b_B$</td>
<td>0.15</td>
</tr>
<tr>
<td>Vertical/horizontal width of the complex cell’s gating field</td>
<td>$\gamma_{cv}, \gamma_{ch}$</td>
<td>0.3, 0.7</td>
</tr>
<tr>
<td>Small, Medium, Large Scale Weight</td>
<td>$w_s, w_m, w_l$</td>
<td>0.25, 0.25, 0.5</td>
</tr>
<tr>
<td>Baseline bias of multiple-scale input</td>
<td>$b_M$</td>
<td>0.01</td>
</tr>
<tr>
<td>Spatial constant of the cable of the filling-in unit</td>
<td>$\sigma$</td>
<td>30</td>
</tr>
<tr>
<td></td>
<td>$\varepsilon$</td>
<td>100</td>
</tr>
<tr>
<td>Names</td>
<td>Symbols</td>
<td>Values</td>
</tr>
<tr>
<td>-------------------------------------------</td>
<td>---------</td>
<td>-------------------------------------</td>
</tr>
<tr>
<td>Decay rate for Anchoring</td>
<td>$B_A$</td>
<td>1</td>
</tr>
<tr>
<td>Depolarization constant for Anchoring</td>
<td>$C_A$</td>
<td>10</td>
</tr>
<tr>
<td>Time constant of modulatory unit of anchoring</td>
<td>$\tau_\Psi$</td>
<td>0.01</td>
</tr>
<tr>
<td>Depolarization constant of modulatory unit of anchoring</td>
<td>$B_\Psi$</td>
<td>1.3</td>
</tr>
<tr>
<td>Recharge rate of tonic activity</td>
<td>$T_\Psi$</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>$m$</td>
<td>100</td>
</tr>
<tr>
<td>White</td>
<td>$\omega$</td>
<td>0.5</td>
</tr>
<tr>
<td>Hyperpolarization constant for gain control</td>
<td>$\phi$</td>
<td>8</td>
</tr>
<tr>
<td>Spatial scale for Anchoring</td>
<td>$\zeta_A$</td>
<td>100 (for the area rule), 4 (for the others)</td>
</tr>
<tr>
<td>Size of connection range for the center of center-surround unit</td>
<td>$\varepsilon_C$</td>
<td>6 (for small scale), 28 (for medium scale)</td>
</tr>
<tr>
<td>Size of connection range for the surround of center-surround unit</td>
<td>$\varepsilon_E$</td>
<td>6 (for small scale), 28 (for medium scale)</td>
</tr>
<tr>
<td>Size of connection range for the half kernel of simple cell</td>
<td>$\varepsilon_B$</td>
<td>3</td>
</tr>
<tr>
<td>Size of connection range for the blurring kernel of Anchoring</td>
<td>$\varepsilon_A$</td>
<td>100 (for the area rule), 4 (for the others)</td>
</tr>
<tr>
<td>Size of connection range for the filling-in unit</td>
<td>$\varepsilon_F$</td>
<td>8</td>
</tr>
<tr>
<td>Sizes of various standard kernels</td>
<td>$W_C, W_E, W_B, W_A, W_F$</td>
<td>0.6, 0.6, 4, 1, 1</td>
</tr>
</tbody>
</table>
References


Grossberg S, Hong S (2003), A neural model of surface perception: Lightness, anchoring, and filling-in. Submitted for publication


