



How the venetian blind percept emerges from the laminar cortical dynamics of 3D vision

Yongqiang Cao¹ and Stephen Grossberg^{2*}

¹ HRL Laboratories, LLC, Malibu, CA, USA

² Graduate Program in Cognitive and Neural Systems, Department of Mathematics, Center for Adaptive Systems, Center for Computational Neuroscience and Neural Technology, Boston University, Boston, MA, USA

Edited by:

William Wren Stine, University of New Hampshire, USA

Reviewed by:

William Wren Stine, University of New Hampshire, USA

Birgitta Dresch-Langley, Centre National de la Recherche Scientifique, France

*Correspondence:

Stephen Grossberg, Center for Adaptive Systems, Boston University, Room 205, 677 Beacon Street, Boston, MA 02215, USA
e-mail: steve@bu.edu

The 3D LAMINART model of 3D vision and figure-ground perception is used to explain and simulate a key example of the Venetian blind effect and to show how it is related to other well-known perceptual phenomena such as Panum's limiting case. The model proposes how lateral geniculate nucleus (LGN) and hierarchically organized laminar circuits in cortical areas V1, V2, and V4 interact to control processes of 3D boundary formation and surface filling-in that simulate many properties of 3D vision percepts, notably consciously seen surface percepts, which are predicted to arise when filled-in surface representations are integrated into surface-shroud resonances between visual and parietal cortex. Interactions between layers 4, 3B, and 2/3 in V1 and V2 carry out stereopsis and 3D boundary formation. Both binocular and monocular information combine to form 3D boundary and surface representations. Surface contour surface-to-boundary feedback from V2 thin stripes to V2 pale stripes combines computationally complementary boundary and surface formation properties, leading to a single consistent percept, while also eliminating redundant 3D boundaries, and triggering figure-ground perception. False binocular boundary matches are eliminated by Gestalt grouping properties during boundary formation. In particular, a disparity filter, which helps to solve the Correspondence Problem by eliminating false matches, is predicted to be realized as part of the boundary grouping process in layer 2/3 of cortical area V2. The model has been used to simulate the consciously seen 3D surface percepts in 18 psychophysical experiments. These percepts include the Venetian blind effect, Panum's limiting case, contrast variations of dichoptic masking and the correspondence problem, the effect of interocular contrast differences on stereoacuity, stereopsis with polarity-reversed stereograms, da Vinci stereopsis, and perceptual closure. These model mechanisms have also simulated properties of 3D neon color spreading, binocular rivalry, 3D Necker cube, and many examples of 3D figure-ground separation.

Keywords: Venetian blind effect, visual cortex, stereopsis, binocular vision, perceptual grouping, surface perception, consciousness, LAMINART model

1. INTRODUCTION: DESIGN PRINCIPLES FOR HOW THE BRAIN SEES THE WORLD IN DEPTH

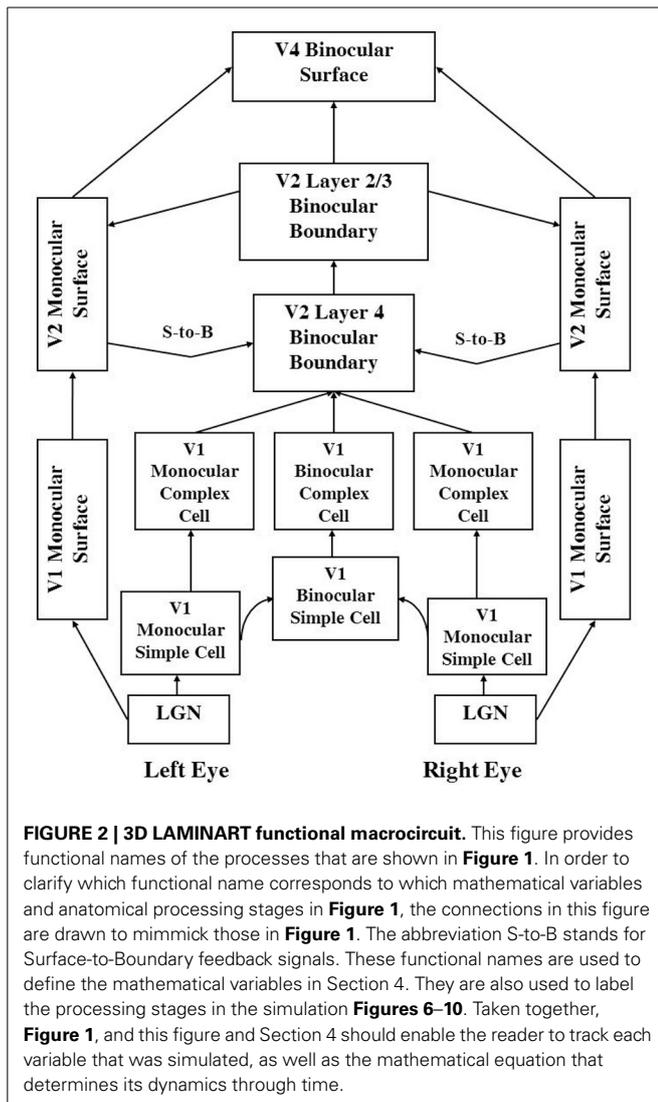
1.1. EXPLAINING 3D PERCEPTS USING LAMINAR CORTICAL NETWORKS

The 3D LAMINART model (Figure 1) predicts how the LGN and cortical areas V1, V2, and V4 computed monocular and binocular visual information to produce three-dimensional (3D) boundary groupings and conscious surface percepts. Each cell type in the model clarifies anatomical and neurophysiological data. Grossberg and Howe (2003), Grossberg and Yazdanbakhsh (2005), and Raizada and Grossberg (2003) provide reviews.

Among the percepts that the model can simulate is the Venetian blind percept that was described in Figure 6.21 of Howard and Rogers (1995). This Venetian blind stimulus consists of two gratings. A low frequency grating is presented to the

left eye, whereas a high frequency one presented to the right eye. During binocular fusion, every second bar of the left grating is in retinal correspondence with every third bar of the right grating. The resulting percept consists of short ramps. Each ramp contains three bars that slope up from left to right. These ramps are interspaced with step returns.

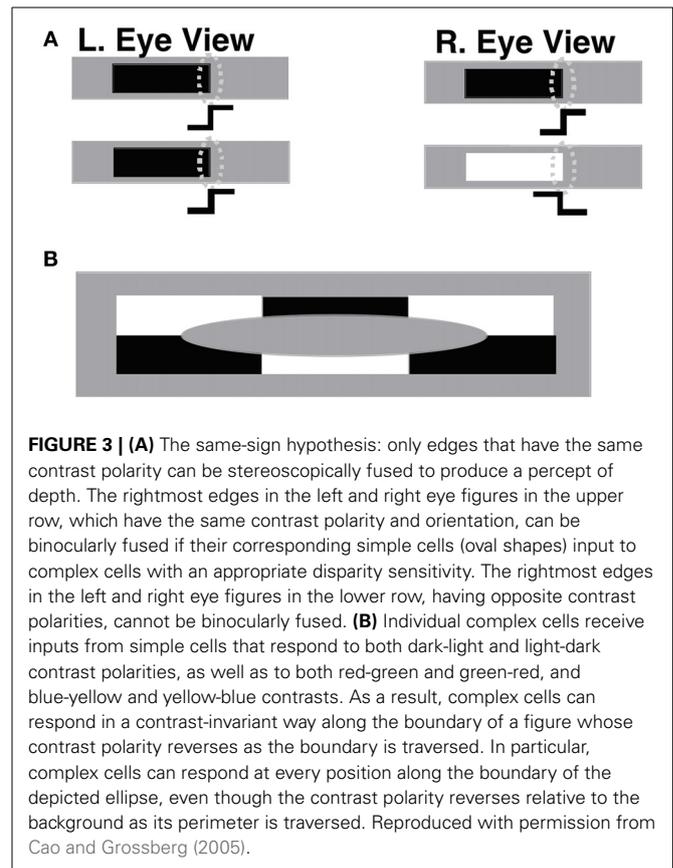
Cao and Grossberg (2005, 2012) showed how the 3D LAMINART model could quantitatively simulate the surface properties that are consciously seen in 18 challenging psychophysical experiments, a feat still not matched by competing models, and did so with a single set of model equations and parameters (see Section 4). The implementation of the model with spiking neurons in Cao and Grossberg (2012) also simulated these surface percepts and demonstrated how analog properties of these percepts emerge from the interactive dynamics of discrete spikes. A simulation of the Venetian blind percept was



limiting case (Panum, 1858; Gillam et al., 1995; McKee et al., 1995). In this percept, a bar presented to one eye is simultaneously matched to two bars presented to the other eye. The unique-matching rule also fails in the Venetian Blind effect. The 3D LAMINART model does not impose unique matches. Rather, correct matches are facilitated by using a *disparity filter* (see DF in **Figure 1**; Grossberg and McLoughlin, 1997; McLoughlin and Grossberg, 1998), whose circuit uses line-of-sight inhibition between active cells that represent different depths at the same position. A parsimonious property of the model, and one that suggests how this constraint may have arisen during evolution, is that these inhibitory interactions are part of the perceptual grouping process that selects and completes 3D boundaries within cortical layer 2/3 (see Section 1.2.5).

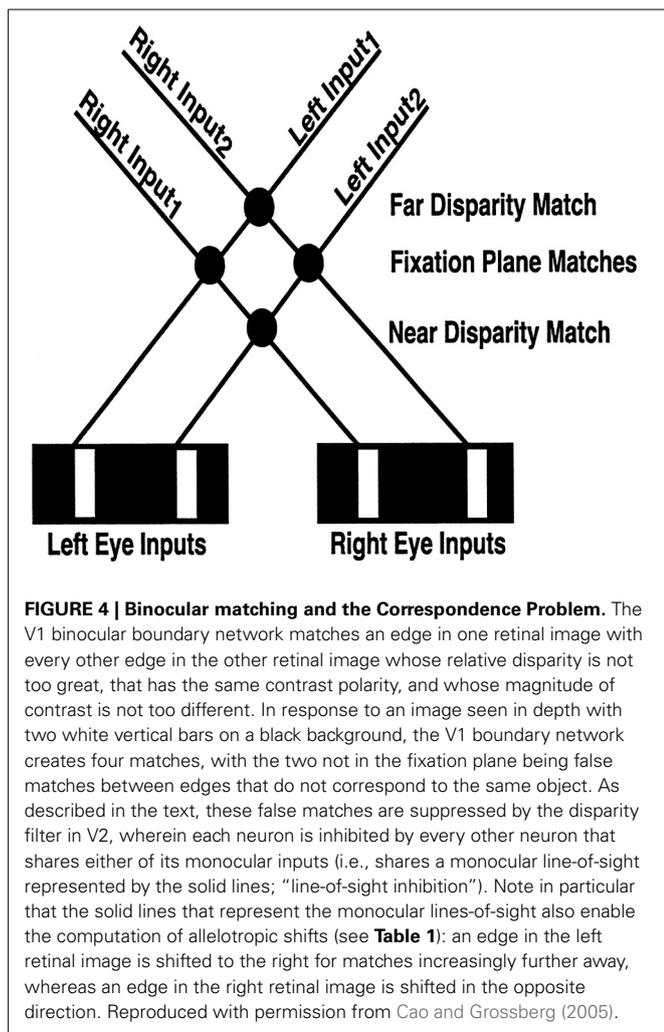
1.2.4. Monocular and binocular information combine in forming depth percepts

Panum's limiting case has homologs in many naturally occurring situations where one edge seen by one eye and two possible edges



with which to match it are seen by the other eye. For example, an object's edge seen by one eye may be occluded by a nearer object when viewed by the other eye. This happens during da Vinci stereopsis (Nakayama and Shimojo, 1990; Gillam et al., 1999). The monocularly viewed region nonetheless has a definite depth that is induced by binocularly viewed scenic features. Monocular information can hereby contribute to forming seamless depthful percepts. These monocular-binocular interactions have been probed by varying the relative contrast of the bars in Panum's limiting case displays and thereby causing alterations in the ensuing percepts of depth (Smallman and McKee, 1995). Another type of display that is useful for the study of monocular-binocular interactions is dichoptic masking, where an object presented to one eye is masked by one presented to the other eye (McKee et al., 1994). All of these variations have been simulated using the 3D LAMINART model.

How do monocular boundaries contribute to depthful percepts? To which depths should a monocular boundary be assigned? This *Monocular-Binocular Interface Problem* was analyzed in Grossberg (1994, 1997) as part of the larger problem of explaining 3D figure-ground percepts. The proposed solution was first implemented in simulations by Grossberg and Howe (2003) of many data about 3D surface perception. This hypothesis predicts that the outputs of monocular boundary cells are added to all depth planes in cortical area V2 pale stripes along their respective lines-of-sight. Layer 4 is a likely site for combining monocular and binocular boundary signals from V1 (**Figure 4**).



Yazdanbakhsh and Watanabe (2004) successfully tested this prediction by doing psychophysical experiments on human subjects. Monocular boundaries that are not at depths where they can form closed boundaries are eliminated by the disparity filter, which is predicted to occur in V2 pale stripes in layer 2/3, where binocular grouping is completed, as explained below. The Monocular-Binocular Interface Problem is hereby resolved at the disparity filter, even as it helps to solve the Correspondence Problem.

1.2.5. 3D perceptual groupings eliminate false matches

Interactions between pyramidal cells in layer 2/3 of the V2 pale stripes were predicted to carry out perceptual grouping using a *bipole property*. This property utilizes both long-range excitatory recurrent axons and shorter-range inhibitory axons. The long-range excitatory recurrent axons connect cells that are (approximately) colinear and coaxial with respect to one another across space. For example, if a Kanizsa square image is presented, each pair of collinear pacmen edges can activate pyramidal cells with like-oriented long-range horizontal connections whose signals summate on target cells between the pacmen. These long-range

horizontal connections also activate the shorter-range inhibitory interneurons. These interneurons inhibit each other and nearby pyramidal cells (**Figure 1**). Due to their mutual inhibition, the total activity in their local inhibitory network is normalized; cf. Grossberg (1973). When the total summing excitation from the long-range excitatory axons on both sides of a target cell, combines with the total normalized inhibition from the shorter-range inhibitory axons, the net excitation can fire the cells cf., von der Heydt and Peterhans (1989) and von der Heydt et al. (1984). This balance of “two-against-one” of excitation and inhibition at target cells implements the bipole property (Grossberg et al., 1997; Grossberg and Raizada, 2000; Raizada and Grossberg, 2001).

When only a single pacman is the inducing stimulus, it induces long-range excitation and disynaptic inhibition from only one side of a recipient cell. In this case, the inhibition is not normalized, so that the excitation and inhibition are approximately balanced, a case of “one-against-one” excitation vs. inhibition, so the target cell is not excited. If the cell receives bottom-up input, it can activate the cell without inputs from any long-range excitatory axons, but its activity can be modulated by such inputs (Bringuier et al., 1999; Crook et al., 2002). Excitatory modulations also help to guide the spread of attention along a boundary grouping (Roelfsema et al., 1998; Ito and Gilbert, 1999; Roelfsema and Spekreijse, 1999; Grossberg and Raizada, 2000), the grouping of 2D and 3D planar percepts (Kapadia et al., 1995; Polat et al., 1998; Bakin et al., 2000), and the grouping of 3D slanted and curved percepts (Grossberg and Swaminathan, 2004).

The 3D LAMINART model predicts that the shorter-range inhibitory interneurons also inhibit the pyramidal cells that correspond to other orientations, notably perpendicular orientations, thereby contributing to figure-ground and binocular rivalry percepts, among others (Grossberg and Swaminathan, 2004; Cao and Grossberg, 2005, 2012; Grossberg and Yazdanbakhsh, 2005; Grossberg et al., 2008). As noted in the previous section, some of these inhibitory interneurons also realize the disparity filter as part of the grouping process. The model hereby predicts that *the selection of a correct 3D grouping includes the suppression of false binocular matches*. The model hereby shows how the solution of the Correspondence Problem also helps to solve the Gestalt grouping problem.

During depth perception, the perceived depths of emergent perceptual groupings, such as illusory contours, often covary with the disparities of their binocularly matched local features. This is not, however, always the case; the perceived depth of perceptual groupings can override local feature disparities (Wilde, 1950; Tausch, 1953; Ramachandran and Nelson, 1976). When this happens, an emergent 3D grouping can suppress “false matches” that are based on the real local disparities of their generative features in the outside world. The model can explain these data because the disparity filter is part of the 3D grouping process, lying as it does within the inhibitory interneurons in layer 2/3 V2 pale stripes, as well as data about how false binocular matches are suppressed that do not correspond to the objects in the outside world. The same perceptual grouping process can also simulate many other challenging data about 3D vision, including data about bistable perception and binocular rivalry (Grossberg and Swaminathan,

2004; Grossberg et al., 2008) and perceptual transparency and Kanizsa stratification (Grossberg and Yazdanbakhsh, 2005).

1.2.6. 3D surface percepts form within Filling-In-Domains

An early model prediction (Grossberg, 1984) that boundary representations on their own do not give rise to visible percepts, indeed that “all boundaries are invisible,” is consistent with many perceptual and neurobiological data. This prediction is expected to hold at least within the interblob cortical stream between V1 and V4 in which perceptual boundaries are formed. In contrast, all visible percepts have been predicted to be a property of surface representations within the blob cortical stream from V1 to V4; see Grossberg (1994, 2014) for reviews. Boundaries are invisible, or amodal, because they pool opposite-polarity contrasts from both achromatic and chromatic receptive fields at the complex cell stage in the interblobs of cortical area V1 in order to build the best possible contrast-invariant boundaries of objects. In particular, they can build complete boundaries from objects that lie in front of textured backgrounds whose relative contrasts reverse along the object’s perimeter (Figure 3B).

Surface representations are completed using a filling-in process that keeps opposite-polarity computations separate to enable distinct lightnesses and colors to fill-in within depth-selective Filling-In Domains, or FIDOs. This filling-in process can reconstruct lightness and color estimates in regions where they have been suppressed by the process of compensating for variable illumination, which is also called “discounting the illuminant” (Grossberg and Todorović, 1988). Boundaries control the depths at which different lightnesses and colors can fill-in via a process called 3D surface capture. The current simulations of the 3D LAMINART model illustrate only the filling-in of achromatic lightnesses in depth in response to psychophysical displays. Grossberg and Hong (2006) and Hong and Grossberg (2004), in their extension of 3D LAMINART to the Anchored Filling-In Lightness Model (aFILM) model, have simulated filling-in of surface lightnesses and colors in response to both psychophysical displays and natural scenes. Other articles have simulated surface percepts in response to artificial sensors such as Synthetic Aperture Radar (SAR), e.g., Grossberg and Williamson (1999) and Mingolla et al. (1999).

How does the brain usually manage to fill-in lightnesses and colors at only the correct depths? The 3D LAMINART surface capture process is reviewed in Grossberg (1994), which also uses it to explain many data about 3D figure-ground perception. The 3D LAMINART model extends FACADE theory to laminar cortical circuitry, while also significantly expanding its explanatory and predictive range. A key surface capture property is that visible surfaces arise in cortical area V4 only if they are enclosed by *connected* boundaries. Figure 5 illustrates how, in response to a da Vinci stereopsis display, a rectangular connected boundary may be composed of one vertical binocular boundary that does encode disparity information, as well as one vertical monocularly viewed boundary and two horizontal boundaries that do not. Such a closed boundary can contain filling-in, and thereby support the formation of a visible surface percept. This surface percept will occur at the depth of the binocular boundary, if all

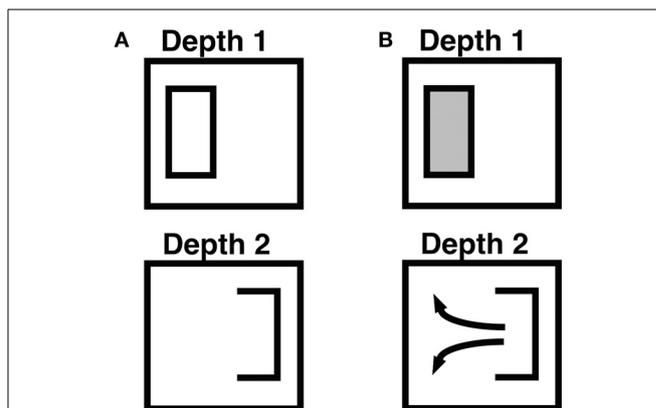


FIGURE 5 | Combining monocular and binocular boundaries, and filling-in within connected boundaries. This figure illustrates how monocular and other depth-unselective boundaries, such as horizontal boundaries, are combined to control surface filling-in. The depicted boundaries, apart from the rectangular boundary frame, represent the boundaries that are induced by a part of an idealized da Vinci stereopsis display in which only the left vertical boundary within the frame is binocularly viewed. See the partially occluded window in Figure 2 of Grossberg (1994) for a complete da Vinci stereopsis display in which this kind of situation can occur. The right vertical boundary within the frame is viewed through only one eye, and is hence a monocular boundary that conveys no disparity information. The horizontal boundaries are processed in the same way as monocular boundaries. **(A)** Open and connected boundaries that are induced by the display after all boundaries are combined. The left vertical boundary is processed selectively at Depth 1 due to the binocular disparity that it induces. The right vertical boundary and the upper and lower horizontal boundaries are added to all boundary representations, across all depths, along the line of sight. This combination of boundaries creates a closed rectangular boundary at Depth 1 and an open boundary at Depth 2. These boundaries are projected topographically to the Filling-In Domain, or FIDO, at the corresponding depth within which boundary-modulated filling-in of surface lightnesses and colors occurs. **(B)** Filling-in of surface lightness is contained or not depending on the connectedness of the boundary. At Depth 1, filling-in is contained within the closed rectangular boundary. At Depth 2, lightness can flow through the boundary gap and become uniform around both sides of the figure’s boundary. As a result, the surface contour representation at Depth 2 generates no output signals via the contrast-sensitive network that outputs to subsequent processing stages. Reproduced with permission from Cao and Grossberg (2005).

other constraints are satisfied. At a different depth plane, where there is no vertical binocular boundary, however, then the total boundary contains a large gaps through which lightness and color signals can flow around the boundaries. Surface contrasts at the boundaries are hereby eliminated and, along with it, any filling-in of those surface features at higher processing stages.

1.2.7. Surface-to-boundary feedback ensures perceptual consistency and initiates figure-ground separation

Boundaries and surfaces form according to computationally *complementary* rules: boundaries are completed *inwardly* using bipole cells, but surface filling-in spreads *outwardly*; boundaries form in an *oriented* way, but surface filling-in spreads *unoriented* in multiple directions; boundaries pool opposite-polarity contrasts at V1 complex cells and hereby become *insensitive to contrast polarity*, whereas visible surfaces are *sensitive to contrast polarity*.

Given that boundaries and surface form using complementary computations, how do they typically work together to generate percepts whose perceived boundaries and surfaces are perceptually consistent? Otherwise expressed, how is complementary consistency achieved? Grossberg (1994) proposed that successfully filled-in regions within the surface representations send Surface-to-Boundary feedback to the boundary system (Figure 2) by sensing whether or not a surface region is filled-in within a connected boundary (see Figure 5B). Contrast-sensitive output circuits accomplish this by detecting where the bounding contours of a successfully filled-in region occur. Such boundary contours occur at positions where contrasts change quickly across space if the boundary can contain filling-in by being connected. If there are large gaps in a boundary that enable filling-in to spread out and around it, then there are no contrast differences around this boundary, so that no output signals are generated there.

These contrast-sensitive output signals are called *surface contours*. They are generated by on-center off-surround networks that operate within disparity and across position in response to filled-in surface activities. The surface contour outputs from the surface stream to the boundary stream strengthen, and thereby confirm, the boundaries that surround successfully filled-in surface regions, at the same time that they inhibit, or prune, redundant boundaries at the same positions and further depths (Grossberg, 1994, 1997; Grossberg and McLoughlin, 1997). The surface contour signals strengthen and prune their target boundaries by activating on-center off-surround networks that operate across disparity and within position within the boundary system. This boundary-enhancing property is predicted to interact with a developmental bias that favors the fixation plane. Taken together, these properties enable the simulation of stereopsis data that have not otherwise been explained; for example, see below and Cao and Grossberg (2005, 2012). Such a fixation plane bias may develop when the relatively high frequency of fixated percepts activates basic activity-dependent mechanisms of self-organizing maps whereby they enhance frequently activated cortical representations, e.g., Buonomano and Merzenich (1998).

Surface contour signals are also predicted to control where the eyes look and how the brain learns invariant object categories. In particular, because surface contour signals are strongest at the distinctive features of an attended object, they can be used to compute eye movement targets. The ARTSCAN, pARTSCAN, and ARTSCAN Search models (Fazl et al., 2009; Grossberg, 2009; Cao et al., 2011; Foley et al., 2012; Chang et al., 2014) predict how surface contour signals may generate predictive eye movement commands, via cortical area V3A, that (1) determine where the eyes will look next and (2) maintain spatial attention in posterior parietal cortex on an object's surface representation in V4, thereby (3) enabling inferotemporal cortex to learn view- and positionally-invariant object categories that represent its boundary and surface representations. Thus, the 3D LAMINART model is part of a larger architecture for active vision wherein 3D boundary and surface representations help to control eye movements for attending, seeing, searching, learning, and recognizing invariant object categories.

2. MODEL DESCRIPTION

The 3D LAMINART model consists of two processing streams: a boundary stream and a surface stream (see Figures 1 and 2). The boundary stream runs from retina/LGN to V1 monocular and binocular boundaries and then to V2 binocular boundaries. The surface stream runs from retina/LGN to V1 and V2 monocular surfaces and then to V4 binocular surfaces. Figure 1 shows a laminar cortical circuit diagram of the 3D LAMINART model, and Figure 2 shows a functional block diagram of the model. A mathematical description of the model is provided in Section 4. In order to facilitate model understanding, the qualitative descriptions of model processes that is provided in this section also list the mathematical equation numbers in which these properties are rigorously defined in Section 4. In order to reduce computational load, the model currently simulates only horizontal and vertical contours and five boundary and surface depths. Five depths were chosen because they are enough to illustrate non-trivial depth separation. The model is extensible to any finite number of depths if finer depth discriminations are desired. Gradual changes of depth using a finite set of depth planes have also been simulated. In particular, Grossberg and Swaminathan (2004, Figure 23d) simulated slanted 3D percepts, including the Necker cube. Such continuous percepts of depth may be achieved by computing a weighted average of filled-in surface activities across multiple depth-selective filling-in domains.

Each model neuron is defined by membrane equation, or shunting, dynamics that have automatic gain control and contrast-normalization properties (Hodgkin, 1964; Grossberg, 1973, 1980). See Equations (1) and (2) in Section 4. Although model neurons and neurons *in vivo* will be clearly distinguished below, model neurons will be referred to by their neurophysiological names because their computational properties closely match those found in the brain. See Grossberg and Howe (2003), Grossberg and Yazdanbakhsh (2005), and Raizada and Grossberg (2003) for comparisons of model cell properties and connections with neurophysiological and neuroanatomical data.

2.1. MONOCULAR BOUNDARIES

2.1.1. LGN cells

The left and right retinal images are processed by LGN cells with circularly symmetric on-center off-surround receptive fields [see Figure 1A; variables $x_{ij}^{L/R}$ in Equations (3)–(5) in Section 4]. These LGN cells discount the illuminant and contrast-normalize the input scene.

2.1.2. V1 monocular simple cells

LGN outputs activate monocular simple cells in layer 4 of the V1 interblobs. The simple cells are oriented filters (Hubel and Wiesel, 1968) that are sensitive to either a similarly oriented dark-light or light-dark contrast in the image, but not both [see Figure 1A; variables $s_{ijk}^{L/R}$ in Equations (6)–(8)]. Outputs from these simple cells project to both binocular simple cells and to monocular simple cells in Layer 3B [for the latter, see Figure 1A; variables $b_{ijk}^{L/R}$ in Equation (9)].

2.2. BINOCULAR BOUNDARIES

2.2.1. V1 binocular simple cells

Outputs from left and right eye monocular simple cells activate binocular simple cells in layer 3B of the interblobs, where stereoscopic fusion begins [see **Figure 1A**; variables b_{ijk}^B in Equation (10)]. The depth selectivity of binocular simple cells is determined by the retinal disparities of the layer 4 monocular cells that project to them. The binocular simple cells in layer 3B are sensitive to just one contrast polarity because only layer 4 simple cells with the same contrast polarity project to them (**Figure 2A**), thereby implementing the *same-sign hypothesis* at layer 3B simple cells that are selective for both binocular disparity and contrast polarity.

The selectivity of stereoscopic fusion at the binocular simple cells in layer 3B also requires the action of inhibitory interneurons [see **Figure 1**; variables $q_{ijkd}^{L/R}$ in Equations (11) and (12)]. The activity of a binocular simple cell is suppressed by these inhibitory interneurons if its left and right eye inputs differ too much in magnitude. The binocular simple cells hereby act like the “obligate cells” of Poggio (1991) by responding only when their left and right eye inputs are approximately equal in magnitude. These layer 3B obligate cells hereby help to solve the Correspondence Problem by responding only to stereoscopically fused stimuli with similar contrast amplitudes from the left and right eye retinal images.

The inhibitory interneurons ensure that binocular simple cells respond only to a narrow range of disparities by inhibiting each other through recurrent inhibitory interneurons, as well as their target binocular simple cell. These recurrent inhibitory interactions normalize total activity within the interneurons. Such normalization enables a “two-against-one” computation that leads to disparity selectivity, much as it leads to the bipole grouping property (see Section 1.2.5) during binocular boundary completion. V1 binocular simple cells and V2 binocular bipole cells thus both seem to use computationally homologous operations in different cortical regions and at different spatial scales. Future experiments should be designed to test this predicted homology and to clarify how it arises during brain evolution and development.

2.2.2. V1 layer 2/3 monocular and binocular complex cells

Layer 2/3 contains complex cells that add inputs from simple cells at the same position that are sensitive to the same orientation but opposite contrast polarities. Both monocular complex cells [see **Figure 1**; variables $c_{ijk}^{L/R}$ in Equations (25) and (26)] and binocular complex cells [see **Figure 1**; variables c_{ijkd} in Equations (13)–(24)] are hereby formed. Because complex cells can respond to both contrast polarities, they can respond all along an object’s boundary, even if contrast polarity with respect to the background reverses as its boundary is traversed (**Figure 3B**). Layer 2/3 complex cells thus implement an early cortical stage of contrast-invariant boundary detection. Complex cells also interact via long-range excitatory recurrent connections [variables H_{ijkdv}^{Ec} in Equations (13) and (15)–(17)] and disynaptic inhibitory interneurons [variables s_{ijkdv}^c in Equations (18)–(21)] that implement non-classical receptive fields via a bipole property

(see Section 2.2.4). This, bipole property in V1 can only modulate a cell, not fire it, if the cell does not also receive a bottom-up input from simple cells.

2.2.3. V2 layer 4 binocular cells

V1 layer 2/3 left and right monocular complex cells and binocular complex cells input to V2 layer 4 cells [see **Figure 1**; variables v_{ijkd} in Equations (27)–(29)]. The monocular cells, which do not compute a binocular disparity, are added to all depth planes in layer 4 along their respective lines-of-sight (**Figure 5**) as part of the brain’s solution of the Monocular-Binocular Interface Problem. The layer 4 cells also receive surface contour feedback signals from left and right V2 monocular surfaces that are formed in the V2 thin stripe region [see Section 2.3.1 and **Figure 1**; variables f_{ijkd}^L , f_{ijkd}^R , and f_{ijkd} in Equations (28) and (29)] to help ensure perceptual consistency, despite the complementarity of boundary and surface computations. These surface contour surface-to-boundary feedback inputs modulate V2 layer 4 cells by enhancing their current level of activity. Surface contour feedback also helps to initiate figure-ground separation (Grossberg, 1994) and plays an indispensable role in explaining various 3D percepts such as da Vinci stereopsis (Cao and Grossberg, 2005).

2.2.4. V2 layer 2/3 bipole grouping cells

Bipole grouping occurs among the binocular cells in V2 layer 2/3 [see **Figure 1A**; variables g_{ijkd} in Equations (30)–(35)]. These cells receive inputs from V2 layer 4 cells. These cells in layer 2/3 possess collinear, coaxial receptive fields that directly excite each other via long-range horizontal axons [variables H_{ijkdv}^{Eg} in Equations (30) and (32)]. They also give rise to short-range, disynaptic inhibitory interneurons [variables s_{ijkdv}^g in Equation (34)] that inhibit themselves as well as their target complex cells [variables H_{ijkd}^{Ig} in Equations (30) and (33)]. This balance of excitation and self-normalizing inhibition achieves the “two-against-one” bipole grouping property (Grossberg et al., 1997; Grossberg, 1999; Grossberg and Raizada, 2000; Grossberg and Williamson, 2001). The boundary grouping process, together with contrast-invariant boundary detection (**Figures 1, 3B**), allows well-localized and connected object boundaries to be completed even in response to disconnected and noisy boundary fragments.

Binocular cells in V1 layer 3B attempt to match every input from a vertical edge from one retinal image with nearby vertical edge input signals from the other retinal image within its disparity range, given that all the inputs code the same contrast polarity and approximately the same contrast magnitude. Even with these restrictions, there remains a Correspondence Problem because false matches may occur in V1.

Figure 4 illustrates the Correspondence Problem that may occur in response to four possible matches if each eye receives inputs from two bars. Only the two matches in the fixation plane are correct, and the other two are false. Such false matches are suppressed in V2 by the disparity filter [see DF in **Figure 1**; variables G_{ijkd}^P are the total DF inhibition in Equations (30) and (35)] as part of the bipole grouping process. The model’s disparity filter (solid lines between **Figure 4** neurons) encourages

unique matching using inhibition between neurons that share a monocular input. These disparity filter inhibitory interactions are, in addition, symmetrical about the fixation plane with the fixation plane inhibiting the near and far disparity planes more than conversely (see **Table 2**). The line-of-sight inhibition and the fixation plane advantage together select two matches in the fixation plane, thereby helping to solve the Correspondence Problem. Because disparity filter interactions occur among the inhibitory interactions that control perceptual grouping in V2 layer 2/3, the model parsimoniously combines suppression of false matches with long-range Gestalt grouping processes.

2.3. MONOCULAR SURFACES

2.3.1. V2 monocular surfaces

Monocular surfaces are computed in the model V2 thin stripes [see **Figure 1**; variables $F_{ijd}^{L/R}$ in Equations (36)–(45)]. The left (right) V2 thin stripe receives boundary signals from V2 layer 2/3 complex cells and illumination-discounted lightness signals from left (right) LGN cells via the left (right) V1 blob region (see **Figure 1**). Earlier FACADE theory and 3D LAMINART modeling has predicted and simulated how surface representations may be generated by a filling-in process, whether through nearest-neighbor diffusive interactions (Cohen and Grossberg, 1984; Grossberg and Todorović, 1988; Grossberg, 1994) or long-range horizontal connections (Grossberg and Hong, 2006). Psychophysical data (e.g., Paradiso and Nakayama, 1991; Pessoa and Neumann, 1998; Pessoa et al., 1998) and neurophysiological data (e.g., Rossi et al., 1996; Lamme et al., 1999) support the existence of filling-in, in contradiction of the critique of Dennett (1991) that such a process does not exist. Surface filling-in has, in fact, been used to explain and simulate many percepts that have not been explained without it, such as surface percepts during figure-ground separation (Kelly and Grossberg, 2000); 2D and 3D neon color spreading and transparency (Grossberg and Mingolla, 1985a; Grossberg, 1994; Grossberg and Yazdanbakhsh, 2005); 3D shape-from-texture (Grossberg et al., 2007); bistable 3D percepts such as the Necker cube illusion (Grossberg and Swaminathan, 2004); and multiple lightness and color percepts (Grossberg and Todorović, 1988; Grossberg and Kelly, 1999; Hong and Grossberg, 2004; Grossberg and Hong, 2006). As illustrated by **Figure 5**, a consciously perceived surface representation can rise from filling-in only if it is enclosed by a connected boundary.

The LGN cells interact through on–center, off–surround circularly symmetric receptive fields whose interactions discount the effects of spatially non-uniform illumination; that is, “discount the illuminant.” These excitatory and inhibitory interactions are balanced and thus attenuate cell activity in response to spatially uniform or slowly varying stimulation. Model LGN cells hereby respond preferentially to luminance borders. At later filling-in stages, these illuminant-discounted surface-border, also called feature contour, signals propagate across surface regions that are enclosed by connected boundaries to complete the lightness representation. Early simulations of filling-in used a boundary-gated nearest-neighbor diffusion equation (Grossberg and Todorović, 1988). Connected boundaries from V2 layer 2/3 create resistive

barriers to limit signal spread (**Figure 5**) of the lightness signals received from the LGN within the V2 thin stripes. Grossberg and Hong (2006) have simulated how filling-in can be carried out by boundary-gated long-range horizontal interactions that operate 1000 times faster than diffusion. An interesting open problem concerns how long-range excitatory interactions may develop, on the one hand, to control the inward and oriented properties of bipole grouping and, on the other hand, to control the complementary outward and unoriented properties of surface filling-in.

The monocular surfaces that are formed in V2 thin stripes are predicted to be invisible, or amodal, and thus do not subserve visible 3D surface percepts. They can, however, directly activate pathways leading to their amodal recognition within the inferotemporal cortex (Grossberg, 1994).

2.4. SURFACE-TO-BOUNDARY FEEDBACK SIGNALS

Successfully filled-in monocular surfaces send surface contour signals from positions at which filled-in contrast changes rapidly enough across space into the boundary representations via V2 layer 4 [see **Figures 1, 2**, Section 1.2.7, and variables $f_{ijkd}^{L/R}$ in Equations (46)–(48)]. These surface-to-boundary feedback signals modulate the activities of V2 boundary cells so that the boundaries that surround the successfully filled-in surfaces are enhanced and redundant boundaries are suppressed, thereby ensuring perceptual consistency and contributing to figure-ground separation

2.5. BINOCULAR SURFACES

2.5.1. V4 surfaces

Consciously visible binocular 3D surface percepts are generated in cortical area V4, where 3D figure-ground separation of object surfaces is also predicted to be completed [see **Figure 1**; variables w_{ijd} in Equations (49) and (50)]. Area V4 receives boundary signals from V2 layer 2/3 and lightness signals from the LGN via V1 blobs and V2 thin stripes. The surface filling-in process is similar to the one described in Section 2.3, except V4 combines monocular lightness signals from both eyes.

The current model simplifies the V4 interactions that have been predicted in Grossberg (1994) to generate 3D surface percepts. See Grossberg and Swaminathan (2004) and Grossberg and Yazdanbakhsh (2005) for simulations of the additional boundary and surface interactions needed to explain percepts that involve the separation of overlapping figures from their backgrounds, such as the Necker cube and 3D neon color spreading.

3. MODEL SIMULATIONS

This section summarizes a simulation of the Venetian blind effect and of Panum’s limiting case to illustrate how monocular and binocular information may interact in the laminar circuits of visual cortex to generate 3D surface percepts. For the Venetian blind simulation, each eye’s stimulus was presented on a grid 30 units high and 115 units wide. For the Panum’s limiting case stimulus, each eye’s stimulus was presented on a grid 30 units high and 60 units wide, which was sufficient to process the smaller width of the stimulus. In all simulations, the background had a

luminance value, in arbitrary units, of 2. In the simulation figures, the light gray bars (if any) had a luminance of 1 and the dark gray bars 0.1. Simulations were performed using the *Matlab* software package.

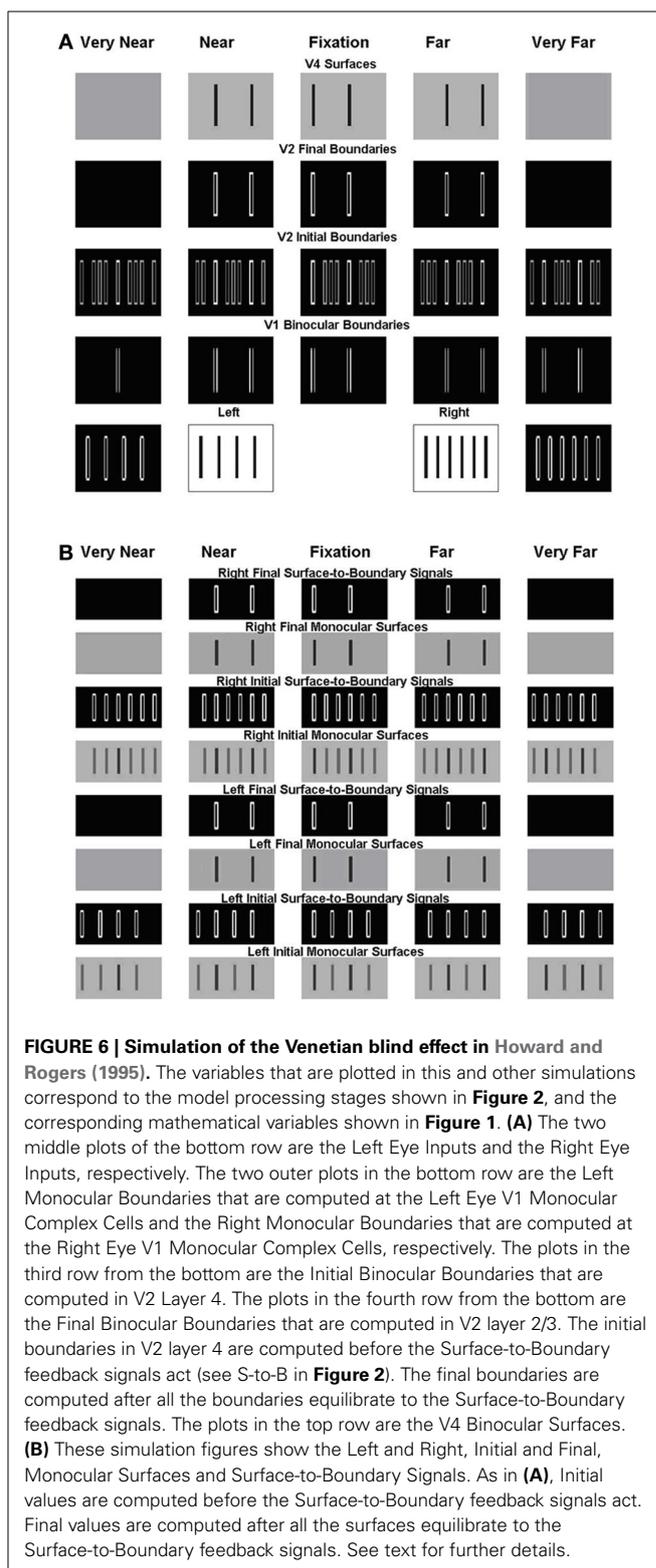
3.1. THE VENETIAN BLIND EFFECT

Section 1.1 described a Venetian blind stereogram from Figure 6.21 of Howard and Rogers (1995). This display consists of a low frequency grating that is presented to the left eye, and a high frequency presented to the right eye. Every second bar of the left grating is in retinal correspondence with every third bar of the right grating; see the middle two plots of the first row of Figure 6A.

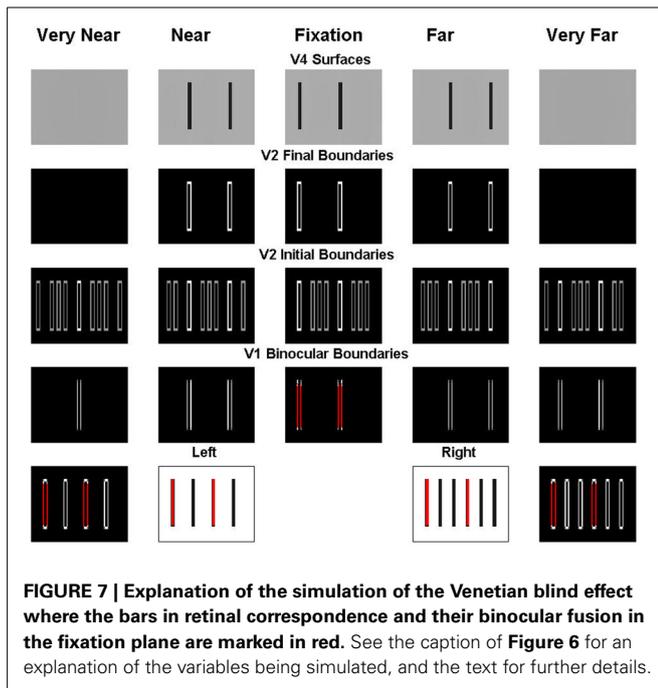
Figure 6 includes two parts: Figure 6A V1 and V2 boundaries and V4 surfaces and Figure 6B V2 monocular surfaces and surface-to-boundary feedback signals. Like the model diagram shown in Figures 1A, 6A should be read from the bottom up, with the bottom two rows representing the input (inner pair of figures with black bars on white background) and the V1 boundary representations (white boundaries on black background), the next two rows representing the V2 boundary representations before (V2 Initial Boundaries) and after (V2 Final Boundaries) surface-to-boundary feedback acts, and the top row representing the V4 surface representations. In Figure 6B, the bottom four rows represent, respectively, in ascending order, the following quantities corresponding to the left eye: the Initial Monocular Surfaces before any feedback interactions occur, the Initial Surface-to-Boundary Signals generated by these surfaces, the Left Final Monocular Surfaces after the Surface-to-Boundary Signals have their effect, and the Final Surface-to-Boundary Signals that are caused by the Final Monocular Surfaces. The same quantities are represented for the right eye in the top four rows. In the top four rows of Figure 6A and all rows of Figure 6B, depth increases from left to right. The middle plot representing the fixation plane. The two leftmost plots represent the two near depth planes. Finally, the two right plots represent the two far depth planes.

This stereogram induces a percept of short ramps, each containing three bars. The bars slope up from left to right and are separated by step returns. The model correctly simulates this surface percept; see the top row of Figure 6A, which shows the simulated depthful surface representations in the model area V4. This row shows that, reading positions from left to right, the first bar of the percept is in the zero disparity plane (fixation plane; second column), the second in the near disparity plane (first column), then there is a step return to the third bar which is located in the far disparity plane (third column), after which the pattern repeats.

Note that every second bar of the left input is in retinal correspondence with every third bar of the right input (bottom row). In Figure 7, these bars are marked in red color (the middle two plots of the bottom row) as are their monocular vertical boundaries (the left most and right most plots of the bottom row). Because of the retinal correspondence, their vertical boundaries are matched in the fixation plane by V1 binocular cells, marked in red color in the middle plot of the second row (reading from the bottom). According to the advantage



of the fixation plane (Section 2.2.4), every vertical boundary located in the same line-of-sight of these binocularly matched vertical boundaries will be killed by the line-of-sight inhibition, whether binocularly matched or not. In particular, vertical

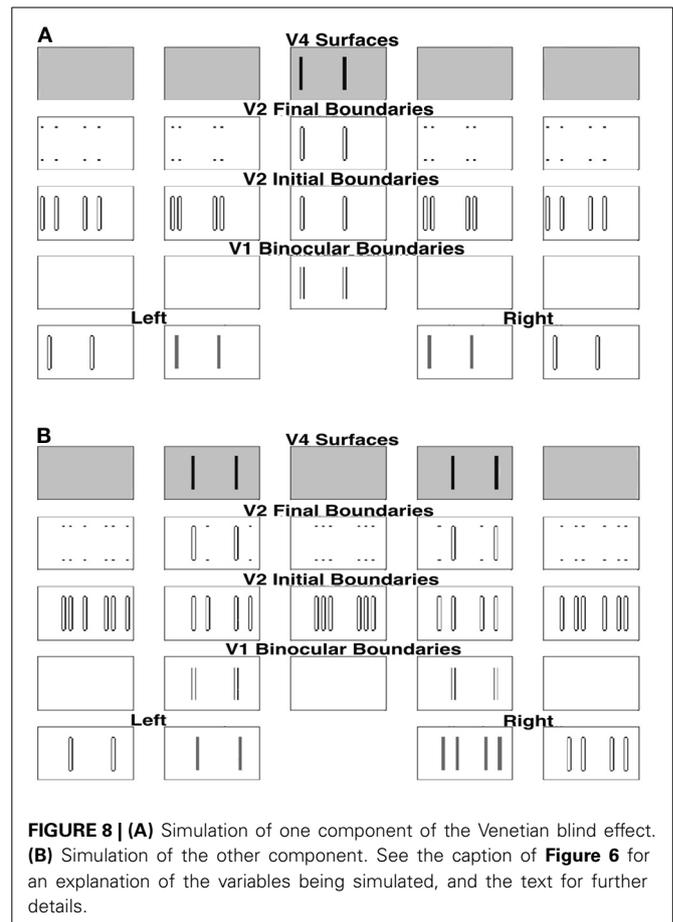


boundaries in the very near and very far depths (the first plot and the last plot of the second row) will all be killed. But the binocularly matched vertical boundaries in the near and far depths (the second plot and the fourth plot of the second row) are not in the same lines-of-sight of the binocular boundaries in the fixation plane, and hence survive the inhibition from them.

In order to simplify the explanation, we can divide the stimulus into two components and consider these separately as follows. First, we extract the bars in retinal correspondence to form the stimulus shown in the middle two plots of the first row of **Figure 8A**. Since the bars in the left and right inputs of this figure are in retinal correspondence, the model correctly predicts that they will appear in the fixation plane, as shown by the middle plot of the fifth row. The specific binocular matches are shown in **Figure 9A**, where the matched bars are marked in the same color.

For the remaining bars, shown in **Figure 8B**, the right eye sees twice the number of bars as the left eye, as in Panum's limiting case, which is explained in the next section. The model simulates fusion of each bar of the left input with two bars of the right input to induce the percept shown in the top row of **Figure 8B**. **Figure 8B** shows how these bars are fused, where the fused bars are marked in the same color and form two side-by-side Panum's limiting cases. The binocular matches are shown in **Figure 9B**.

Adding together the percepts shown in top rows of **Figures 8A,B** yield the surface percept shown in the top row of **Figure 6A**, thereby explaining this Venetian blind effect. **Figure 6B** shows the left and right V2 monocular surfaces and the corresponding surface-to-boundary signals, in both initial and final stages. The surface-to-boundary signals enhance the surviving boundaries with a closed contour that contains



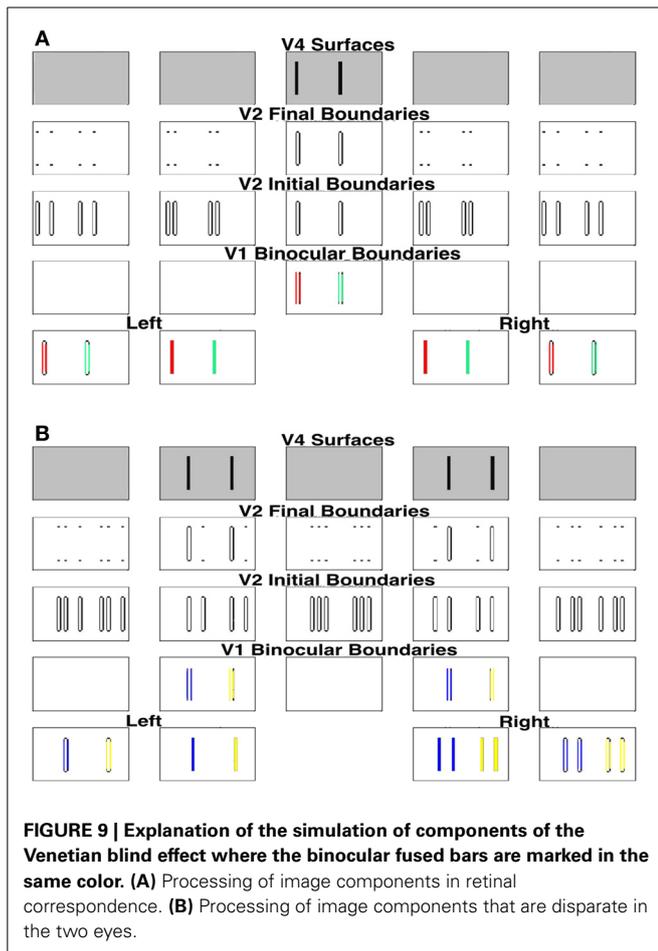
surface filling-in, and therefore helps to generate the correct percept.

The insight that the model provides is that this example of the Venetian blind effect is just a complex version of the Correspondence Problem as it is illustrated by Panum's limiting case, when it is properly understood by combining early stereo matching, later boundary selection by a disparity filter that is part of the boundary grouping process, and surface filling-in of those regions that are completely enclosed by connected boundaries.

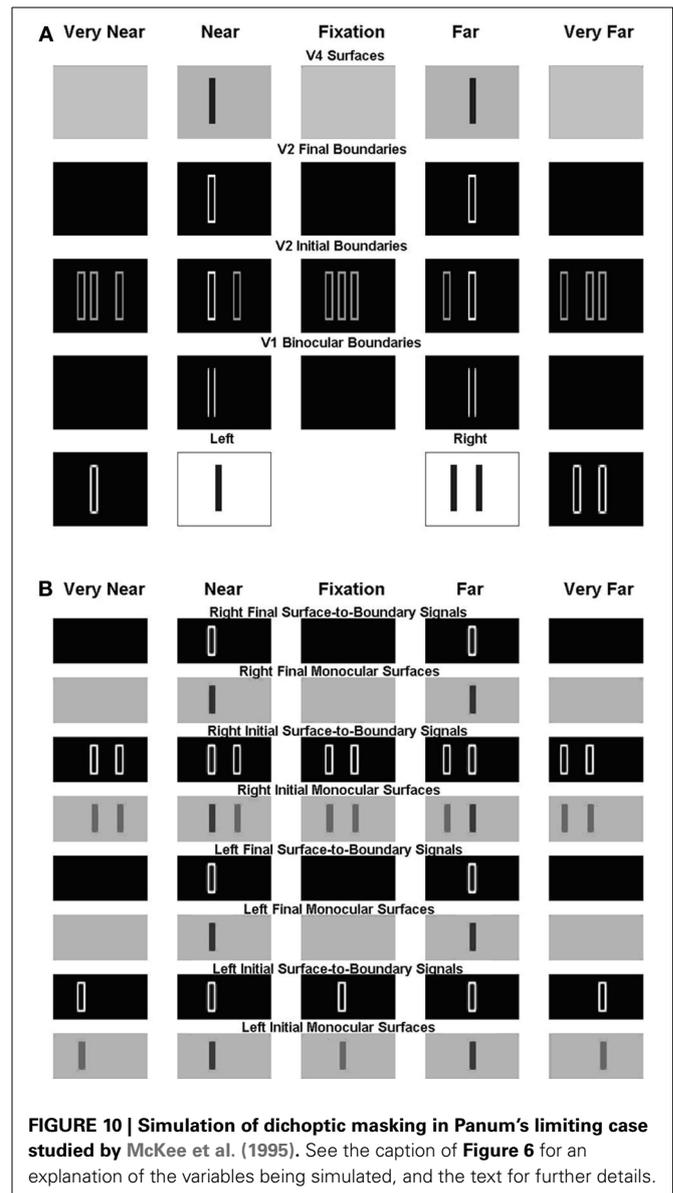
3.2. DICHOTIC MASKING IN PANUM'S LIMITING CASE

As described in the previous section, the present model solves the Correspondence Problem by using a disparity filter that encourages unique matching, via line-of-sight inhibition, but does not enforce it. One advantage of this is that the model can simulate Panum's limiting case, where a bar in one eye is simultaneously fused with two bars in the other eye (Panum, 1858; Gillam et al., 1995; McKee et al., 1995). **Figure 10** shows the model simulation where a bar in one eye masks equally two bars presented to the other eye as reported by McKee et al. (1995).

For this stimulus, the left eye receives a single bar input, whereas the right eye receives an input consisting of two bars; see the middle two plots within the first row of **Figure 10A**. The monocular boundaries that are induced by these stimuli are shown within the outer two plots of the first row. The bar of the



left input fuses with both bars of the right input at V1 binocular boundary cells, thereby forming binocular boundaries in both a near and a far disparity plane; see, the second and fourth plots of the second row. In V2, monocular boundaries are added to all disparity planes along their respective lines-of-sight; see the third row. Left monocular boundaries induce the left bar representation in the first two plots, the middle bar representation in the third plot, and the right bar representation in the fourth and fifth plots of the third row. Right monocular boundaries induce the two right bar representations in the first two plots, the outer two bar representations in the third plot, and the leftmost two bar representations in the fourth and fifth plots of the third row. The V1 monocular and binocular boundaries are added within V2. V2 boundaries that combine binocular input and monocular inputs are stronger than those that do not. Recurrent inhibition of the V2 disparity filter suppresses the V2 vertical boundaries that receive only monocular input and that share one of their lines-of-sight. The binocularly matched boundaries in the near and far depths, represented, respectively, in the third row by the first bar of the second plot and the second bar of third plot, are in the same line-of-sight. But they have the same (or almost the same) strength, and therefore do not kill each other because the disparity filter encourages unique matching, via line-of-sight inhibition, but does not enforce it. The surviving V2 boundary representations



are shown in the fourth row. Those regions in V2 that are enclosed by a connected boundary give rise to surface percepts in V4, as shown in the fifth row.

Figure 10B shows the left and right V2 monocular surfaces and the corresponding surface-to-boundary signals, in both initial and final stages. The surface-to-boundary signals enhance the surviving binocular boundaries in the second and fourth plots of the fourth row of **Figure 9A**, which have closed contours that enclose the surface filling-in process, and therefore generate the correct surface percept. The model correctly predicts that the bar of the left input is matched with both bars of the right input, and so masks them both equally (McKee et al., 1995).

4. MODEL EQUATIONS

This section describes the 3D LAMINART model equations. Since it describes the model used in Cao and Grossberg (2005), the

description of the equations is adapted from the exposition in that article.

Each neuron is a single voltage compartment whose membrane potential, $v(t)$, obeys:

$$C_m \frac{dv(t)}{dt} = (E_{leak} - v(t))g_{leak} + (E_{excit} - v(t))g_{excit}(t) + (E_{inhib} - v(t))g_{inhib}(t). \quad (1)$$

In (1), parameters E denote reversal potentials, g_{leak} is a constant leakage conductance, and the time-varying conductances $g_{excit}(t)$ and $g_{inhib}(t)$ are the total excitatory and inhibitory inputs to the cell (Hodgkin, 1964; Grossberg, 1968, 1973). The products of potential differences times conductances define shunting interactions that enable automatic gain control and normalization of activity to occur among the interacting variables. The following notation is used for the capacitance term $C_m = 1$, the leakage conductance $g_{leak} = A$, and the reversal potentials $E_{excit} = B$, $E_{inhib} = -C$, and $E_{leak} = 0$. When Equation (1) can be rewritten as:

$$\frac{dv}{dt} = -Av + (B - v)g_{excit} - (C + v)g_{inhib}. \quad (2)$$

In (2), A is a constant decay rate, B is the maximum membrane potential, C is the minimum membrane potential, g_{excit} is the total excitatory input, and g_{inhib} is the total inhibitory input. All the simulations use a single set of parameters. For simplicity, the same parameter symbol (e.g., α) may be used in different equations, but with a different value, that is specified in each equation. **Figure 1** labels the model processing stages with the corresponding mathematical variable names to facilitate tracking the network relationships among the various variables. The non-technical summary of model processes in Section 2 parallels the definition of mathematical equations in this section to provide further expository support. Simulations were performed using *Matlab*.

4.1. LGN

The LGN cell membrane potentials, $x_{ij}^{L/R}$, obey the following shunting on-center off-surround equation:

$$\frac{dx_{ij}^{L/R}}{dt} = -\alpha x_{ij}^{L/R} + (\beta - x_{ij}^{L/R}) I_{ij}^{L/R} - x_{ij}^{L/R} \sum_{p \neq i, q \neq j} G_{pqij} I_{pq}^{L/R}, \quad (3)$$

where L/R designates that the cell belongs to the left or right monocular pathway, indices i and j denote the position of the input on the retina, α is a constant (10^{-5}) rate of decay, β is a constant (9.9) maximum membrane potential, $I_{ij}^{L/R}$ is the luminance input of the left or right retinal image to the excitatory on-center, and G_{pqij} is a Gaussian inhibitory off-surround kernel:

$$G_{pqij} = \exp\left(-\frac{(p-i)^2 + (q-j)^2}{2\sigma^2}\right), \quad (4)$$

where σ scales the kernel size (1.5). The steady-state cell membrane potentials are:

$$x_{ij}^{L/R} = \frac{\beta I_{ij}^{L/R}}{\alpha + \sum_{p,q} G_{pqij} I_{pq}^{L/R}}. \quad (5)$$

Equation (5) was used in the simulations.

4.2. V1 LAYER 4 SIMPLE CELLS

At steady-state, the membrane potentials, $s_{ijk}^{L/R,+}$, of simple cells that respond to dark-light contrast polarity are given by:

$$s_{ijk}^{L/R,+} = \sum_{p,q} K_{pqk} [x_{i+p,j+q}^{L/R}]^+. \quad (6)$$

In (6), the superscript L/R, + means L+ or R+. L denotes left monocular, R denotes right monocular, + indicates that the simple cell responds to dark-light contrast polarity, and k denotes orientation. Two orientations, vertical ($k = 1$) and horizontal ($k = 2$), were simulated. The threshold linear signal function $[x]^+ = \max(x, 0)$, and K_{pqk} is an orientationally-tuned Gabor kernel:

$$K_{pqk} = \phi \sin\left(\frac{2\pi(r-0.5)}{\tau}\right) \exp\left[-\frac{1}{2}\left(\frac{(p-0.5)^2}{\sigma_p^2} + \frac{(q-0.5)^2}{\sigma_q^2}\right)\right]. \quad (7)$$

In (7), terms $\phi, \tau, \sigma_p, \sigma_q$ are constants (4.4, 3π , 0.6, 0.6) representing kernel amplitude and dimensions; $r = p$ for cells that respond to vertical boundaries; and $r = q$ for those that respond to horizontal boundaries.

The cell membrane potentials of simple cells with light-dark contrast polarity are the inverse of those in (6):

$$s_{ijk}^{L/R,-} = -s_{ijk}^{L/R,+} = -\sum_{p,q} K_{pqk} [x_{i+p,j+q}^{L/R}]^+. \quad (8)$$

4.3. V1 LAYER 3B MONOCULAR SIMPLE CELLS

The steady-state membrane potentials, $b_{ijk}^{L/R,+/-}$ of the layer 3B monocular simple cells obey:

$$b_{ijk}^{L/R,+/-} = 2 [s_{ijk}^{L/R,+/-}]^+. \quad (9)$$

The factor of 2 compensates for monocular simple cells receiving an input from one eye whereas binocular simple cells receive inputs from both eyes.

4.4. V1 LAYER 3B BINOCULAR SIMPLE CELLS

The layer 3B binocular simple cells implement the obligate property by receiving excitatory inputs from layer 4 and balanced inhibitory input from layer 3B inhibitory interneurons at the same position and disparity. Their membrane potentials,

$b_{ijkd}^{B,+/-}$ obey:

$$\begin{aligned} \frac{d}{dt} b_{ijkd}^{B,+/-} = & -\gamma_1 b_{ijkd}^{B,+/-} + \left(1 - b_{ijkd}^{B,+/-}\right) \\ & \left(\left[s_{(i+s)jk}^{L,+/-} - \theta \right]^+ + \left[s_{(i-s)jk}^{R,+/-} - \theta \right]^+ \right) \\ & - \alpha \left(\left[q_{ijkd}^{L,+/-} \right]^+ + \left[q_{ijkd}^{L,-/+} \right]^+ \right) \\ & + \left[q_{ijkd}^{R,+/-} \right]^+ + \left[q_{ijkd}^{R,-/+} \right]^+, \end{aligned} \quad (10)$$

where the parameters γ_1 , α , and θ (0.1, 7.2, 0.4) represent the decay rate of the membrane potential, the inhibitory gain, and the signal threshold. The variables $q_{ijkd}^{L/R,+/-}$ are membrane potentials of inhibitory interneurons in layer 3B, k is the orientation, d is the disparity to which the model neuron selectively fires, and index s is the positional shift between left and right eye inputs that depends on the disparity; see **Table 1**.

Layer 3B inhibitory interneuronal cell membrane potentials, $q_{ijkd}^{L/R,+/-}$, receive excitatory input from layer 4 and inhibitory input from all other inhibitory interneurons that code the same position and disparity. Their left (L) and right (R) subpopulations obey:

$$\begin{aligned} \frac{d}{dt} q_{ijkd}^{L,+/-} = & -\gamma_2 q_{ijkd}^{L,+/-} + \left[s_{(i+s)jk}^{L,+/-} - \theta \right]^+ - \beta \left(\left[q_{ijkd}^{R,+/-} \right]^+ \right. \\ & \left. + \left[q_{ijkd}^{R,-/+} \right]^+ + \left[q_{ijkd}^{L,-/+} \right]^+ \right), \end{aligned} \quad (11)$$

and

$$\begin{aligned} \frac{d}{dt} q_{ijkd}^{R,+/-} = & -\gamma_2 q_{ijkd}^{R,+/-} + \left[s_{(i-s)jk}^{R,+/-} - \theta \right]^+ - \beta \left(\left[q_{ijkd}^{L,+/-} \right]^+ \right. \\ & \left. + \left[q_{ijkd}^{L,-/+} \right]^+ + \left[q_{ijkd}^{R,-/+} \right]^+ \right). \end{aligned} \quad (12)$$

In (11) and (12), the parameters γ_2 , β , and θ (4.5, 4, 0.4) are the membrane potential decay rate, the inhibitory gain, and the signal threshold, respectively.

Table 1 | The allelotropic shift (s) is the amount that the left and right monocular contours must be displaced to form a single fused binocular contour.

Disparity (d)	V. Near disparity	Near disparity	Zero disparity	Far disparity	V. Far disparity
Allelotropic shift (s)	-8	-4	0	+4	+8

It depends on the disparity. It is zero for matches in the fixation plane because these matches are between contours at retinal correspondence. **Figure 3** illustrates the allelotropic shift and shows that a left monocular contour needs to be shifted more to the right for matches that are further from the observer, whereas a right monocular contour needs to be shifted in the opposite direction.

Mild constraints on these parameter values are sufficient to ensure that the binocular simple cells act like “obligate cells” (Poggio, 1991) that respond vigorously only when their left and right inputs are approximately equal in size. Equation (10) was solved at equilibrium, using the equations described in the Obligate Theorem (see Section 4) to speed up the simulations.

4.5. V1 LAYER 2/3 MONOCULAR AND BINOCULAR COMPLEX CELLS

V1 layer 2/3 monocular and binocular complex cells pool the cell membrane potentials of monocular and binocular layer 3B simple cells of both contrast polarities at each position and orientation. These complex pyramidal cells also emit long-range, collinear, coaxial connections within layer 2/3 whereby they excite each other. The long-range excitatory connections also diverge to excite short-range, disynaptic interneurons that inhibit target complex cells and nearby inhibitory interneurons. The recurrent inhibition normalizes the total activity of the inhibitory interneuronal network. The “two-against-one” balance of excitation and inhibition that converges on the pyramidal cells implements the bipole property that controls boundary grouping (Grossberg and Mingolla, 1985a,b; Grossberg and Raizada, 2000). Monocular and binocular bipole cells obey the same laws but have different inputs. The membrane potential, c_{ijkd} , of binocular collinear bipole cell in V1 layer 2/3 obeys:

$$\begin{aligned} \frac{d}{dt} c_{ijkd} = & -\alpha c_{ijkd} + (\beta - c_{ijkd}) \\ & \left(I_{ijkd}^c \left(\gamma_1 + \gamma_2 \left[\sum_v H_{ijkdv}^{Ec} - H_{ijkd}^{Lc} \right]^+ \right) + \gamma_3 [c_{ijkd} - \beta_c]^+ \right) \\ & - (1 + c_{ijkd}) (C_{ijkd}^P + C_{ijkd}^S), \end{aligned} \quad (13)$$

with parameters α , β , γ ($\alpha = 20$, $\beta = 7$, $\gamma_1 = 1$, $\gamma_2 = 1$, $\gamma_3 = 0.5$). The input I_{ijkd}^c from V1 layer 3B binocular simple cells obeys:

$$I_{ijkd}^c = \mu \left(\left[b_{ijkd}^{B,+} - \theta \right]^+ + \left[b_{ijkd}^{B,-} - \theta \right]^+ \right). \quad (14)$$

with parameters μ and θ (20, 0.1). The excitatory input H_{ijkdv}^{Ec} derives from long-range recurrent connections in V1 layer 2/3 to a complex cell at position (i, j) , orientation k , disparity d , and side v of the bipole cell (Grossberg and Swaminathan, 2004). Term $\sum_v H_{ijkdv}^{Ec}$ sums inputs from both sides $v = 1, 2$ of the bipole cell:

$$H_{ijkdv}^{Ec} = \sum_{pq} W_{pqijkv}^c [c_{pqkd} - \zeta_c]^+. \quad (15)$$

In (15), ζ_c is a threshold (0), and W_{pqijkv}^c is the long-range connection weight from the bipole cell at position (p, q) to the bipole cell at position (i, j) , orientation k , and side v . The connection weights for the horizontal orientation ($k = 2$) obey ($v = 1$ for left branch and $v = 2$ for right branch):

$$W_{pqij21}^c = \left[\text{sign}(i - p) \exp \left(- \left(\frac{(i - p)^2}{\sigma_p^2} + \frac{(j - q)^2}{\sigma_q^2} \right) \right) \right]^+, \quad (16)$$

and

$$W_{pqij22}^c = \left[\text{sign}(p-i) \exp \left(- \left(\frac{(i-p)^2}{\sigma_p^2} + \frac{(j-q)^2}{\sigma_q^2} \right) \right) \right]^+, \quad (17)$$

where $\text{sign}(x) = 1$ if $x > 0$, -1 if $x < 0$, and 0 otherwise. Parameters $\sigma_p = 8$, $\sigma_q = 0.3$, and the spatial connection range (diameter) is 3 . Connection weights for the vertical orientation are obtained by rotation.

Term H_{ijkd}^{Ic} is the total inhibitory input from the inhibitory interneurons:

$$H_{ijkd}^{Ic} = \sum_v \left[s_{ijkdv}^c \right]^+. \quad (18)$$

In (18), s_{ijkdv}^c is the activity of the interneuron that inhibits the bipole cell at position (i, j) , orientation k , disparity d , from side v , where:

$$\frac{d}{dt} s_{ijkdv}^c = \delta_I \left(-s_{ijkdv}^c + H_{ijkdv}^{Ec} - \eta s_{ijkdv}^c \left[s_{ijkdv}^c \right]^+ \right). \quad (19)$$

In (19), u and v are the two branches of orientation k , δ_I is a large gain that ensures rapid response of the inhibitory interneuron, and $\eta = 1$. Term H_{ijkdv}^{Ec} from the long-range connection excites the interneuron, whereas term $-\eta s_{ijkdv}^c \left[s_{ijkdv}^c \right]^+$ defines the recurrent inhibition that normalizes the total inhibitory activity, thereby enabling the 2-to-1 bipole interaction. Solving (19) at equilibrium yields the steady-state inhibitory activities that are used in the simulations:

$$s_{ijkdv}^c = \left(-B_v + \sqrt{B_v^2 + 4\eta H_{ijkdv}^{Ec}} \right) / 2\eta, \quad (20)$$

and

$$s_{ijkdu}^c = \left(-B_u + \sqrt{B_u^2 + 4\eta H_{ijkdu}^{Ec}} \right) / 2\eta. \quad (21)$$

In (20) and (21), respectively, $B_v = 1 + \eta \left(H_{ijkdu}^{Ec} - H_{ijkdv}^{Ec} \right)$ and $B_u = 1 + \eta \left(H_{ijkdv}^{Ec} - H_{ijkdu}^{Ec} \right)$. A bipole cell will not fire when it receives excitatory input from only one side of its long-range connection, but it can fire when it receives excitatory inputs from both sides. In (19), for example, when H_{ijkdu}^{Ec} equals zero, then s_{ijkdu}^c equals zero. As a result, s_{ijkdv}^c equals H_{ijkdv}^{Ec} . The total excitatory input from the long-range connections then equals the total inhibitory input from the inhibitory interneurons, and cancel. However, when the collinear excitatory long-range recurrent inputs H_{ijkdu}^{Ec} and H_{ijkdv}^{Ec} from both sides of the target cell are far from zero, the sum of these excitatory inputs is larger than the total self-normalizing inhibitory input, and thus the cell can fire if the cell is in V2 see Equation (30). If the cell is in V1, as in Equation (13), then the excitatory inputs from its long-range connections are modulatory, thereby enhancing bipole cell firing only if it also receives excitatory input from layer 4.

Term $\gamma_3 [c_{ijkd} - \beta_c]^+$ in (13) is self-excitatory feedback, with threshold $\beta_c = 0.03$. Term C_{ijkd}^P is the inhibitory input at the same position and disparity from other bipole cells that code different orientations:

$$C_{ijkd}^P = \gamma_4 \left(\sum_{r \neq k} [c_{ijrd} - \beta_c]^+ \right), \quad (22)$$

where $\gamma_4 = 5$. Term C_{ijkd}^S is the inhibitory input from spatial competition across position and orientation, but within disparity. Competition across position sharpens and localizes the spatial positions of boundaries. Competition across orientation prevents abutting lines in the image of different orientations from generating illusory contours that can penetrate the interior of a differently oriented boundary. This latter property is called *spatial impenetrability* (Grossberg and Mingolla, 1987; Grossberg and Williamson, 2001). Term C_{ijkd}^S obeys:

$$C_{ijkd}^S = \gamma_5 \sum_{p \neq i, q \neq j, r} W_{pqijk} [c_{pqrd} - \beta_c]^+. \quad (23)$$

In (23), $\gamma_5 = 1$, and W_{pqijk} is an elliptic Gaussian kernel elongated at the orientation perpendicular to k . For the W_{pqijk} vertical orientation ($k = 1$):

$$W_{pqij1} = \exp \left(- \left(\frac{(i-p)^2}{\sigma_p^2} + \frac{(j-q)^2}{\sigma_q^2} \right) \right), \quad (24)$$

where $\sigma_p > \sigma_q$ ($\sigma_p = 8$, $\sigma_q = 0.3$). The horizontally oriented kernel is obtained by rotation.

The membrane potential, $c_{ijk}^{L/R}$, of a monocular complex cell obeys the same equation as that of a binocular complex cell c_{ijkd} :

$$\begin{aligned} \frac{d}{dt} c_{ijk}^{L/R} = & -\alpha c_{ijk}^{L/R} + \left(\beta - c_{ijk}^{L/R} \right) \\ & \left(I_{ijk}^{c, L/R} \left(\gamma_1 + \gamma_2 \left[\sum_v H_{ijkv}^{Ec} - H_{ijk}^{Ic} \right]^+ \right) + \gamma_3 \left[c_{ijk}^{L/R} - \beta_c \right]^+ \right) \\ & - \left(1 + \frac{L/R}{c_{ijk}} \right) \left(C_{ijk}^P + C_{ijk}^S \right). \end{aligned} \quad (25)$$

The input $I_{ijk}^{c, L/R}$ from V1 layer 3B monocular simple cells obeys:

$$I_{ijk}^{c, L/R} = \left[b_{ijk}^{L/R, +} - \theta \right]^+ + \left[b_{ijk}^{L/R, -} - \theta \right]^+. \quad (26)$$

The parameters $\alpha = 20$, $\beta = 8$, and $\theta = 0.4$. The other parameters are the same as the binocular parameters.

4.6. V2 LAYER 4

Most V2 cells are binocular (Hubel and Livingstone, 1987), consistent with the model's combination of V1 left and right monocular inputs with binocular inputs in layer 4 of V2, such that monocular inputs are added to all depth planes along their

respective lines-of-sight, yielding. Initially, the following initial membrane potential, v_{ijkd} , of a V2 layer 4 cell:

$$v_{ijkd} = \alpha h \left(\left[c_{ijkd}^- \theta \right]^+ \right) + \beta \left(h \left(\left[c_{(i+s)jk}^L - \theta_m \right]^+ \right) + h \left(\left[c_{(i-s)jk}^R - \theta_m \right]^+ \right) \right). \quad (27)$$

In (27), s is the positional shift between left and right eye input (see **Table 1**); h is a signal function with $h(x) = 1$ if $x > 0$, 0 otherwise; Θ and θ_m are signal thresholds (0.06, 0.3); and α and β (2.6, 0.8) are the gains of the excitatory binocular and monocular connections, respectively.

These layer 4 boundary cells in the V2 pale stripes also receive feedback signals from left and right monocular surfaces in the V2 thin stripes. These are the surface-to-boundary surface contour feedback signals that were discussed in constraint (7) of Section 1. When these surface contour signals are active, (27) is replaced, at steady-state, by:

$$v_{ijkd} = \left(\alpha h \left(\left[c_{ijkd} - \theta \right]^+ \right) + \beta \left(h \left(\left[c_{(i+s)jk}^L - \theta_m \right]^+ \right) + h \left(\left[c_{(i-s)jk}^R - \theta_m \right]^+ \right) \right) \right) (1 + \alpha_f f_{ijkd}) (\delta + (1 - \delta) h(f_{ijkd})). \quad (28)$$

In (28), f_{ijkd} is the total surface contour feedback signal, α_f is its excitatory gain (1.1), and α , β , h , θ and θ_m are the same as in (27). Parameter δ (0.2) scales the activities of layer 4 cells. If $\delta < 1$, then the activities of layer 4 cells that do not receive surface contours signals are suppressed to some degree. The total surface contour feedback signal f_{ijkd} obeys:

$$f_{ijkd} = \left[f_{ijkd}^L - \theta_f \right]^+ + \left[f_{ijkd}^R - \theta_f \right]^+. \quad (29)$$

In (29), f_{ijkd}^L and f_{ijkd}^R are the surface contour signals derived from left and right monocular surfaces, respectively, and θ_f is a threshold (0.03).

4.7. V2 LAYER 2/3 COMPLEX CELLS

The bipole cells in V2 layer 2/3 implement perceptual grouping by long-range horizontal connections, and include the disparity filter as part of the inhibitory interactions that control perceptual grouping. See constraint (5) of Section 1. The membrane potential, g_{ijkd} , of the bipole cell at position (i,j) that codes orientation k and disparity d obeys:

$$\frac{d}{dt} g_{ijkd} = -\alpha g_{ijkd} + (\beta - g_{ijkd}) \left(\gamma_1 I_{ijkd}^g + \gamma_2 \left[\sum_v H_{ijkdv}^{Eg} - H_{ijkd}^{Ig} \right]^+ \right) - (1 + g_{ijkd}) G_{ijkd}^P, \quad (30)$$

where α , β , γ_1 , and γ_2 are constants ($\alpha = 30$, $\beta = 10$, $\gamma_1 = 1.4$, $\gamma_2 = 1$). Term I_{ijkd}^g in (30) is the input from V2 layer 4:

$$I_{ijkd}^g = [v_{ijkd}]^+. \quad (31)$$

As for V1 layer 2/3 cells, V2 layer 2/3 bipole cells receive long-range recurrent excitatory signals from other (almost) collinear and coaxial bipole cells at nearby positions with the same disparity preference. Term H_{ijkdv}^{Eg} is the total excitatory input from branch v of the bipole cell at position (i,j) , orientation k and disparity d :

$$H_{ijkdv}^{Eg} = \sum_{pq} W_{pqijkv}^g [g_{pqkd} - \zeta_g]^+. \quad (32)$$

The long-range connection weight W_{pqijkv}^g is the same as W_{pqijkv}^c in (16) and (17), but with a larger spatial range (diameter = 7) and $\sigma_p = 15$, $\sigma_q = 0.1$, and $\zeta_g = 0.03$. Term H_{ijkd}^{Ig} is the total input from the inhibitory interneurons:

$$H_{ijkd}^{Ig} = \sum_v \left[s_{ijkdv}^g \right]^+. \quad (33)$$

The activity, s_{ijkdv}^g , of the inhibitory interneuron for branch v has the same form as the inhibitory interneurons that are defined in Equation (19) for V1 layer 2/3:

$$\frac{d}{dt} s_{ijkdv}^g = \delta_I \left(-s_{ijkdv}^g + H_{ijkdv}^{Eg} - \eta s_{ijkdv}^g \left[s_{ijkdu}^g \right]^+ \right). \quad (34)$$

The parameters in (34) are also the same as those in (19).

Term G_{ijkd}^P in (30) is the disparity filter (DF in **Figure 1A**) whose inhibitory signals cross disparities along the lines-of-sight:

$$G_{ijkd}^P = \gamma_3 \sum_{d' \neq d} M_{dd'} \left(\left[g_{(i+s'-s)jkd'} - \beta_g \right]^+ + \left[g_{(i+s-s')jkd'} - \beta_g \right]^+ \right). \quad (35)$$

In (35), γ_3 is the gain (5) of the total disparity filter inhibition; $M_{dd'}$ is the connection strength of line-of-sight inhibition from all cells that share a monocular input between disparities d and d' (see **Table 2**); and $\left[g_{(i+s'-s)jkd'} - \beta_g \right]^+$ and $\left[g_{(i+s-s')jkd'} - \beta_g \right]^+$ are V2 layer 2/3 bipole cell inhibitory inputs along the left and right lines-of-sight with the corresponding disparity-induced positional shifts s and s' (see **Table 1**) and threshold β_g (0.03).

The disparity filter (G^P) works across a range of parameter values. As illustrated in **Table 2**, it needs to be symmetrical about the fixation plane (i.e., the near and far disparity planes equally inhibit and are equally inhibited by the zero disparity plane), and the zero disparity plane inhibits the near and far disparity planes more than conversely.

Table 2 | The inhibition coefficients $M_{dd'}$ that define line-of-sight inhibition.

	V. Near	Near	Zero	Far	V. Far
V. Near	–	3	5	3	2
Near	0.4	–	2.5	2	0.4
Zero	0.3	1.5	–	1.5	0.3
Far	0.4	2	2.5	–	0.4
V. Far	2	3	5	3	–

Each neuron is inhibited by every other neuron that shares either of its inputs by an amount that depends on the disparities of the inhibited and inhibiting neurons (cf. Figure 3).

4.8. V2 THIN STRIPE MONOCULAR SURFACE FILLING-IN

Monocular surface filling-in occurs within V2 thin stripes in response to lightness signals from LGN via V1 blobs, and binocular boundary signals from V2 layer 2/3 bipole cells. The boundary signals create resistive barriers that contain the spread of lightness during the filling-in process. The activity $F_{ijd}^{L/R}$ in each Filling-In Domain, or FIDO, is the membrane potential of a left (L) or right (R) monocular surface cell at position (i, j) and disparity d . These activities obey a nearest-neighbor diffusion equation (Grossberg and Todorović, 1988):

$$\varepsilon \frac{d}{dt} F_{ijd}^{L/R} = -\alpha F_{ijd}^{L/R} + \sum_{(p,q) \in N_{ij}} (F_{pqd}^{L/R} - F_{ijd}^{L/R}) \Phi_{pqijd} + x_{ijd}^{L/R}. \quad (36)$$

The rate parameter $\varepsilon \ll 1$ ensures that this surface filling-in process within the thin stripes is faster than the boundary grouping process within the pale stripes. Parameter α is the decay rate (1.0); and N_{ij} is the set of the nearest-neighbor positions of (i, j) :

$$N_{ij} = \{(i, j-1), (i-1, j), (i+1, j), (i, j+1)\}; \quad (37)$$

and $x_{ijd}^{L/R}$ is the lightness input from the left (L) or right (R) LGN:

$$x_{ijd}^L = [x_{(i+s)j}^L]^+ \quad (38)$$

and

$$x_{ijd}^R = [x_{(i-s)j}^R]^+, \quad (39)$$

where s is the disparity-sensitive positional shift (see Table 1). The boundary-gated diffusion coefficients, Φ_{pqijd} , in (36) are suppressed at positions where the boundary signals are large:

$$\Phi_{pqijd} = \frac{\delta}{1 + \rho (g_{(i-0.5)(j-0.5)d} + g_{(i-0.5)(j+0.5)d})} \quad \text{if } p = i-1 \text{ and } q = j, \quad (40)$$

$$\Phi_{pqijd} = \frac{\delta}{1 + \rho (g_{(i+0.5)(j-0.5)d} + g_{(i+0.5)(j+0.5)d})}, \quad \text{if } p = i+1 \text{ and } q = j, \quad (41)$$

$$\Phi_{pqijd} = \frac{\delta}{1 + \rho (g_{(i-0.5)(j-0.5)d} + g_{(i+0.5)(j-0.5)d})}, \quad \text{if } p = i \text{ and } q = j-1, \quad (42)$$

$$\Phi_{pqijd} = \frac{\delta}{1 + \rho (g_{(i-0.5)(j+0.5)d} + g_{(i+0.5)(j+0.5)d})}, \quad \text{if } p = i \text{ and } q = j+1. \quad (43)$$

The diffusion parameter $\delta = 2000$, and the boundary-gating parameter $\rho = 200$. The boundary terms in these equations sum over all orientations of bipole cell activations at the corresponding position and disparity:

$$g_{ijd} = \gamma \sum_k [g_{ijkd} - \theta_g]^+, \quad (44)$$

with gain $\gamma = 10$ and threshold $\theta_g = 0.03$.

In Equations (40)–(43), the lattice of boundary-gating signals is offset by $[0.5, 0.5]$ relative to the lattice of discounted lightness inputs. This enables the boundary-gating signals to contain filling-in without trapping the lightness inputs within the boundary itself. These two processing streams are also spatially displaced in the cortical map. Spurious edge effects were avoided by using wrap-around whereby the last element of a row/column is adjacent to the first element of the same row/column (Grossberg and Howe, 2003).

The steady-state of (36) is used in the simulations; namely:

$$F_{ijd}^{L/R} = \frac{x_{ijd}^{L/R} + \sum_{(p,q) \in N_{ij}} F_{pqd}^{L/R} \Phi_{pqijd}}{\alpha + \sum_{(p,q) \in N_{ij}} \Phi_{pqijd}}. \quad (45)$$

This approximation is justified by the assumption of filling-in that is fast relative to the rate of boundary formation. During each time step of boundary grouping, monocular surfaces are filled-in using (45) and generate surface-to-boundary surface contour signals before the process is reiterated.

4.9. SURFACE CONTOUR FEEDBACK SIGNALS

The surface contour signals project from V2 monocular surfaces in the thin stripes to V2 layer 4 cells in the pale stripes to modulate binocular boundaries in layer 2/3; see Equation (28). Output signals from the left (L) and right (R) monocular surface activities (45) are derived from oriented filters:

$$f_{ijkd}^{L/R,+} = \sum_{p,q} K_{pqk} [F_{i+p,j+q,d}^{L/R}]^+, \quad (46)$$

$$f_{ijkd}^{L/R,-} = -f_{ijkd}^{L/R,+} = -\sum_{p,q} K_{pqk} [F_{i+p,j+q,d}^{L/R}]^+. \quad (47)$$

where the Gabor kernel K_{pqk} is defined in Equation (7). Their sum defines the final surface contour signals:

$$f_{ijkd}^{L/R} = \left[f_{ijkd}^{L/R,+} \right]^+ + \left[f_{ijkd}^{L/R,-} \right]^+. \quad (48)$$

V4 binocular surface filling-in: Visible 3D percepts. V4 is predicted in the model to support visible percepts of 3D surfaces. To accomplish this, V4 receives lightness signals from the LGN via V1 blobs and V2 thin stripes, and boundary signals from V2 layer 2/3. It combines the monocular lightness signals from the two eyes that correspond to the same 3D location. Its binocular lightness input, z_{ijd} , sums the rectified monocular lightness signals from the left (L) and right (R) eyes that correspond to the same 3D position:

$$z_{ijd} = \left[x_{(i+s)j}^L \right]^+ + \left[x_{(i-s)j}^R \right]^+. \quad (49)$$

In (49), i, j are positional indices, d disparity and s the positional shift defined in **Table 1**. V4 cell membrane potentials, $w_{ij d}$, undergo binocular surface filling-in using a steady-state diffusion equation similar to (36):

$$w_{ij d} = \frac{z_{ij d} + \sum_{(p,q) \in N_{ij}} w_{pq d} \Phi_{pqij d}}{\alpha + \sum_{(p,q) \in N_{ij}} \Phi_{pqij d}}. \quad (50)$$

In (50), parameter $\alpha = 1$; the set of nearest-neighbors N_{ij} is defined in (A37); and $\Phi_{pqij d}$ is defined in (40)–(43) with the diffusion parameter $\delta = 1000$ and the boundary-gating parameter $\rho = 400$.

4.10. OBLIGATE THEOREM

The following theorem shows that the obligate property holds at the binocular simple cells in layer 3B. See Grossberg and Howe (2003) for a proof.

Obligate Theorem. Consider the system:

$$\begin{aligned} \frac{db_{ijkd}^{B,+}}{dt} = & -\gamma_1 b_{ijkd}^{B,+} + \left(1 - b_{ijkd}^{B,+}\right) \left(S_{(i+s)jk}^{L,+} + S_{(i-s)jk}^{R,+}\right) \\ & - \alpha \left(\left[q_{ijkd}^{L,+}\right]^+ + \left[q_{ijkd}^{L,-}\right]^+ + \left[q_{ijkd}^{R,+}\right]^+ + \left[q_{ijkd}^{R,-}\right]^+\right), \end{aligned} \quad (51)$$

$$\begin{aligned} \frac{dq_{ijkd}^{L,+}}{dt} = & -\gamma_2 q_{ijkd}^{L,+} + S_{(i+s)jk}^{L,+} \\ & - \beta \left(\left[q_{ijkd}^{R,+}\right]^+ + \left[q_{ijkd}^{R,-}\right]^+ + \left[q_{ijkd}^{L,-}\right]^+\right), \end{aligned} \quad (52)$$

$$\begin{aligned} \frac{dq_{ijkd}^{R,+}}{dt} = & -\gamma_2 q_{ijkd}^{R,+} + S_{(i-s)jk}^{R,+} \\ & - \beta \left(\left[q_{ijkd}^{L,+}\right]^+ + \left[q_{ijkd}^{L,-}\right]^+ + \left[q_{ijkd}^{R,-}\right]^+\right), \end{aligned} \quad (53)$$

$$\begin{aligned} \frac{dq_{ijkd}^{L,-}}{dt} = & -\gamma_2 q_{ijkd}^{L,-} + S_{(i+s)jk}^{L,-} \\ & - \beta \left(\left[q_{ijkd}^{R,-}\right]^+ + \left[q_{ijkd}^{R,+}\right]^+ + \left[q_{ijkd}^{L,+}\right]^+\right), \end{aligned} \quad (54)$$

and

$$\begin{aligned} \frac{dq_{ijkd}^{R,-}}{dt} = & -\gamma_2 q_{ijkd}^{R,-} + S_{(i-s)jk}^{R,-} \\ & - \beta \left(\left[q_{ijkd}^{L,-}\right]^+ + \left[q_{ijkd}^{L,+}\right]^+ + \left[q_{ijkd}^{R,+}\right]^+\right), \end{aligned} \quad (55)$$

where

$$S_{(i+s)jk}^{L,+/-} = \left[s_{(i+s)jk}^{L,+/-} - \theta \right]^+, \quad (56)$$

and

$$S_{(i-s)jk}^{R,+/-} = \left[s_{(i-s)jk}^{R,+/-} - \theta \right]^+. \quad (57)$$

In (6) and (7), $s_{(i+s)jk}^{L,+/-}$ and $s_{(i-s)jk}^{R,+/-}$ are the monocular simple cell activities defined by (6) and (8). The parameters $\theta \geq 0$, $\gamma_1 > 0$ and

$$0 < \beta < \gamma_2 < \alpha < \gamma_2 + \beta \quad (58)$$

Then, in response to constant inputs, the system converges exponentially to the following unique equilibrium. Let $\Gamma = \gamma_1 + S_{(i+s)jk}^{L,+} + S_{(i-s)jk}^{R,+}$.

$$(1) \text{ if } 0 < S_{(i+s)jk}^{L,+}, S_{(i-s)jk}^{R,+}; \text{ and } \frac{\beta}{\gamma_2} \leq \frac{S_{(i+s)jk}^{L,+}}{S_{(i-s)jk}^{R,+}} \leq \frac{\gamma_2}{\beta},$$

then at equilibrium

$$b_{ijkd}^{B,+} = \frac{1}{\Gamma} \left(1 - \frac{\alpha}{\gamma_2 + \beta}\right) \left(S_{(i+s)jk}^{L,+} + S_{(i-s)jk}^{R,+}\right); \quad (59)$$

$$(2) \text{ if } 0 < S_{(i+s)jk}^{V,L,+}, S_{(i-s)jk}^{V,R,+}; \text{ and } \frac{S_{(i+s)jk}^{L,+}}{S_{(i-s)jk}^{R,+}} > \frac{\gamma_2}{\beta_2},$$

then at equilibrium

$$b_{ijkd}^{B,+} = \frac{1}{\Gamma} \left(S_{(i-s)jk}^{R,+} + \left(1 - \frac{\alpha}{\gamma_2}\right) S_{(i+s)jk}^{L,+}\right); \quad (60)$$

$$(3) \text{ if } 0 < S_{(i+s)jk}^{V,L,+}, S_{(i-s)jk}^{V,R,+}; \text{ and } \frac{S_{(i+s)jk}^{L,+}}{S_{(i-s)jk}^{R,+}} < \frac{\beta}{\gamma_2},$$

then at equilibrium

$$b_{ijkl}^{B,+} = \frac{1}{\Gamma} \left(S_{(i+s)jk}^{L,+} + \left(1 - \frac{\alpha}{\gamma_2} \right) S_{(i-s)jk}^{R,+} \right); \quad (61)$$

$$(4) \text{ for all other values of } S_{(i+s)jk}^{L,+}, S_{(i-s)jk}^{R,+},$$

$$\text{at equilibrium } b_{ijkl}^{B,+} \leq 0. \quad (62)$$

5. DISCUSSION: CONSCIOUS 3D SURFACE PERCEPTS ARE PART OF SURFACE-SHROUD RESONANCES

The current simulations illustrate how the interactions among identified neurons in laminar cortical circuits can give rise to consciously seen surface percepts such as the Venetian blind and Panum's limiting case percepts. These simulations, along with those in articles such as Cao and Grossberg (2005, 2012); Fang and Grossberg (2009), Grossberg and Howe (2003), Grossberg et al. (2007), Grossberg and McLoughlin (1997), Grossberg and Raizada (2000), Grossberg and Swaminathan (2004), Grossberg and Yazdanbakhsh (2005), Kelly and Grossberg (2000), and Raizada and Grossberg (2001), show how the model cortical interactions of FACADE theory and its laminar cortical extension in the 3D LAMINART model provide unified explanations, and testable predictions, about a wide variety of psychophysical and neurobiological data about 2D and 3D perceptual grouping, surface perception, and figure-ground separation.

The filled-in surface representations in all of these articles have parametric properties that closely match visually perceived and reported surface percepts by human subjects. It is for this reason that the liberty is taken of calling them model representations of conscious percepts. However, the model as stated in the current article is insufficient to represent the dynamics that may subserve a conscious percept in the brain. This insufficiency may be better understood from the vantage point of other theoretical results which clarify how the current model may be consistently embedded into a larger theory wherein more sophisticated correlates of conscious events may be represented.

This insufficiency may be summarized in the light of two theoretical predictions. The first prediction is part of FACADE theory (Grossberg, 1994) and the 3D LAMINART model. It claims that "conscious visual percepts are surface percepts" that are represented within the surface cortical stream through V1 blobs, V2 thin stripes, and their projections to V4 and beyond. This prediction coexists with the companion prediction that "all boundaries are invisible" within the boundary cortical stream through V1 interblobs, V2 pale stripes, and their projections to V4 and beyond.

The second prediction is part of Adaptive Resonance Theory, or ART (Grossberg, 1976, 2012, 2013; Carpenter and Grossberg, 1991, 1993). It claims that "all conscious states are resonant states."

Putting together these two predictions raises the question: what resonance enables filled-in surface representations to be consciously seen? Recent research on how the brain coordinates attention, perception, eye movement search, learning, and recognition of invariant object categories, embodied in a class of models whose variations are called ARTSCAN, distributed

ARTSCAN (dARTSCAN), positional ARTSCAN (pARTSCAN), and ARTSCAN Search (Fazl et al., 2009; Grossberg, 2009; Cao et al., 2011; Foley et al., 2012; Chang et al., 2014), has shed new light on this question. In particular, these models have explained and predicted how surface representations, say in cortical area V4, bid for spatial attention, say in posterior parietal cortex (PPC), and how a winning surface, to which focused spatial attention is paid, enables spatial attention in PPC to fit itself to the shape of the surface. Such form-fitting spatial attention is sometimes called an *attentional shroud* (Tyler and Kontsevich, 1995) and the feedback interaction that maintains spatial attention in PPC upon the surface of interest in V4 is said to form a *surface-shroud resonance*.

Grossberg (2012, 2013) has predicted that every conscious visual surface percept is part of a surface-shroud resonance. Such a resonance has been proposed to propagate both top-down to lower cortical levels, such as V1, where finer features of seen representations may be represented, as well as bottom-up to higher cortical areas. For purposes of the present article, it suffices to observe that the current 3D LAMINART model can be consistently embedded into a larger system for focusing spatial attention from the Where cortical stream upon consciously seen visual surface representations using surface-shroud resonances, while learning to recognize them in the What cortical stream using feature-category resonances, such as those modeled by ART (Grossberg, 1980, 2012; Carpenter and Grossberg, 1991, 1993). Said more simply, a mechanistic account is now available for documenting the differences between seeing and knowing, and how they are coordinated into seamless moments of conscious awareness using interactions between the What and Where cortical streams.

REFERENCES

- Bakin, J. S., Nakayama, K., and Gilbert, C. D. (2000). Visual responses in monkey area V1 and V2 to three-dimensional surface configurations. *J. Neurosci.* 20, 8188–8198.
- Bringuiet, V., Chavane, F., Glaeser, L., and Frégnac, Y. (1999). Horizontal propagation of visual activity in the synaptic integration field of area 17 neurons. *Science* 283, 695–699. doi: 10.1126/science.283.5402.695
- Buonomano, D. V., and Merzenich, M. M. (1998). Cortical plasticity: from synapses to maps. *Annu. Rev. Neurosci.* 21, 149–198. doi: 10.1146/annurev.neuro.21.1.149
- Cao, Y., and Grossberg, S. (2005). A laminar cortical model of stereopsis and 3D surface perception: closure and da Vinci stereopsis. *Spat. Vis.* 18, 515–578. doi: 10.1163/156856805774406756
- Cao, Y., and Grossberg, S. (2012). Stereopsis and 3D surface perception by spiking neurons in laminar cortical circuits: a method of converting neural rate models into spiking models. *Neural Netw.* 26, 75–98. doi: 10.1016/j.neunet.2011.10.010
- Cao, Y., Grossberg, S., and Markowitz, J. (2011). How does the brain rapidly learn and reorganize view- and positionally-invariant object representations in inferior temporal cortex? *Neural Netw.* 24, 1050–1061. doi: 10.1016/j.neunet.2011.04.004
- Carpenter, G., and Grossberg, S. (1993). Normal and amnesic learning, recognition, and memory by a neural model of cortico-hippocampal interactions. *Trends Neurosci.* 16, 131–137. doi: 10.1016/0166-2236(93)90118-6
- Carpenter, G. A., and Grossberg, S. (eds.). (1991). *Pattern Recognition by Self-organizing Neural Networks*. Cambridge, MA: MIT Press.
- Chang, H.-C., Grossberg, S., and Cao, Y. (2014). Where's Waldo? How perceptual, cognitive, and emotional brain processes cooperate during learning to categorize and find desired objects in a cluttered scene. *Front. Integr. Neurosci.* 8:43. doi: 10.3389/fnint.2014.00043
- Cohen, M. A., and Grossberg, S. (1984). Neural dynamics of brightness perception: features, boundaries, diffusion, and resonance. *Percept. Psychophys.* 36, 428–456. doi: 10.3758/BF03207497

- Crook, J. M., Engelmann, R., and Lowel, S. (2002). Caba-inactivation attenuates collinear facilitation in cat primary visual cortex. *Exp. Brain Res.* 143, 295–302. doi: 10.1007/s00221-002-1007-y
- Dennett, D. C. (1991). *Consciousness Explained*. Boston, MA: Little, Brown, and Company.
- Fang, L., and Grossberg, S. (2009). From stereogram to surface: how the brain sees the world in depth. *Spat. Vis.* 22, 45–82. doi: 10.1163/156856809786618484
- Fazl, A., Grossberg, S., and Mingolla, E. (2009). View-invariant object category learning, recognition, and search: how spatial and object attention are coordinated using surface-based attentional shrouds. *Cogn. Psychol.* 58, 1–48. doi: 10.1016/j.cogpsych.2008.05.001
- Foley, N. C., Grossberg, S., and Mingolla, E. (2012). Neural dynamics of object-based multifocal visual spatial attention and priming: object cueing, useful-field-of-view, and crowding. *Cogn. Psychol.* 65, 77–117. doi: 10.1016/j.cogpsych.2012.02.001
- Gillam, B., Blackburn, S., and Cook, M. (1995). Panum's limiting case: double fusion, convergence error, or "da Vinci stereopsis." *Perception* 24, 333–346. doi: 10.1068/p240333
- Gillam, B., Blackburn, S., and Nakayama, K. (1999). Stereopsis based on monocular gaps: metrical encoding of depth and slant without matching contours. *Vision Res.* 39, 493–502. doi: 10.1016/S0042-6989(98)00131-X
- Grimson, W. E. (1981). A computer implementation of a theory of human stereo vision. *Philos. Trans. R. Soc. B* 292, 217–253. doi: 10.1098/rstb.1981.0031
- Grossberg, S. (1968). Some physiological and biochemical consequences of psychological postulates. *Proc. Natl. Acad. Sci. U.S.A.* 60, 758–765. doi: 10.1073/pnas.60.3.758
- Grossberg, S. (1973). Contour enhancement, short-term memory and constancies in reverberating neural networks. *Stud. Appl. Math.* 52, 217–257. doi: 10.1007/978-94-009-7758-7_8
- Grossberg, S. (1976). Adaptive pattern classification and universal recoding, II: feedback, expectation, olfaction, and illusions. *Biol. Cybern.* 23, 187–202.
- Grossberg, S. (1980). How does a brain build a cognitive code? *Psychol. Rev.* 87, 1–51. doi: 10.1037/0033-295X.87.1.1
- Grossberg, S. (1984). "Outline of a theory of brightness, color, and form perception," in *Trends in Mathematical Psychology*, eds E. Degreef and J. van Buggenhaut (Amsterdam: North-Holland), 5559–5586.
- Grossberg, S. (1994). 3D vision and figure-ground separation by visual cortex. *Percept. Psychophys.* 55, 48–120. doi: 10.3758/BF03206880
- Grossberg, S. (1997). Cortical dynamics of three-dimensional figure-ground perception of two-dimensional figures. *Psychol. Rev.* 104, 618–658. doi: 10.1037/0033-295X.104.3.618
- Grossberg, S. (1999). How does the cerebral cortex work? Learning, attention and grouping by the laminar circuits of visual cortex. *Spat. Vis.* 12, 163–186. doi: 10.1163/156856899X00102
- Grossberg, S. (2009). Cortical and subcortical predictive dynamics and learning during perception, cognition, emotion, and action. *Philos. Trans. R. Soc. Lond.* 364, 1223–1234. doi: 10.1098/rstb.2008.0307
- Grossberg, S. (2012). Adaptive resonance theory: how a brain learns to consciously attend, learn, and recognize a changing world. *Neural Netw.* 37, 1–47. doi: 10.1016/j.neunet.2012.09.017
- Grossberg, S. (2013). *Adaptive Resonance Theory*. Scholarpedia. Available online at: http://www.scholarpedia.org/article/Adaptive_resonance_theory
- Grossberg, S. (2014). "The visual world as illusion: the ones we know and the ones we don't," in *Oxford Compendium of Visual Illusions*, eds A. Shapiro and D. Todorovic (Oxford: Oxford University press), (in press).
- Grossberg, S., and Hong, S. (2006). A neural model of surface perception: lightness, anchoring, and filling-in. *Spat. Vis.* 19, 263–321. doi: 10.1163/156856806776923399
- Grossberg, S., and Howe, P. D. L. (2003). A laminar cortical model of stereopsis and three-dimensional surface perception. *Vision Res.* 43, 801–829. doi: 10.1016/S0042-6989(03)00011-7
- Grossberg, S., and Kelly, F. (1999). Neural dynamics of binocular brightness perception. *Vision Res.* 39, 3796–3816. doi: 10.1016/S0042-6989(99)00095-4
- Grossberg, S., Kuhlmann, L., and Mingolla, E. (2007). A neural model of 3D shape-from-texture: multiple-scale filtering, boundary grouping, and surface filling-in. *Vision Res.* 47, 634–672. doi: 10.1016/j.visres.2006.10.024
- Grossberg, S., and McLoughlin, N. P. (1997). Cortical dynamics of 3-D surface perception: binocular and half-occluded scenic images. *Neural Netw.* 10, 1583–1605. doi: 10.1016/S0893-6080(97)00065-8
- Grossberg, S., and Mingolla, E. (1985a). Neural dynamics of perceptual grouping: textures, boundaries, and emergent segmentations. *Percept. Psychophys.* 38, 141–147. doi: 10.3758/BF03198851
- Grossberg, S., and Mingolla, E. (1985b). Neural dynamics of form perception: boundary completion, illusory figures, and neon color spreading. *Psychol. Rev.* 92, 173–211. doi: 10.1037/0033-295X.92.2.173
- Grossberg, S., and Mingolla, E. (1987). Neural dynamics of surface perception: boundary webs, illuminants, and shape-from-shading. *Comput. Vis. Graph. Image Process.* 37, 116–165. doi: 10.1016/S0734-189X(87)80015-4
- Grossberg, S., Mingolla, E., and Ross, W. D. (1997). Visual brain and visual perception: how does the cortex do perceptual grouping? *Trends Neurosci.* 20, 106–111. doi: 10.1016/S0166-2236(96)01002-8
- Grossberg, S., and Raizada, R. D. (2000). Contrast-sensitive perceptual grouping and object-based attention in the laminar circuits of primary visual cortex. *Vision Res.* 40, 1413–1432. doi: 10.1016/S0042-6989(99)00229-1
- Grossberg, S., and Swaminathan, G. (2004). A laminar cortical model for 3D perception of slanted and curved surfaces and of 2D images: development, attention and bistability. *Vision Res.* 44, 1147–1187. doi: 10.1016/j.visres.2003.12.009
- Grossberg, S., and Todorović, D. (1988). Neural dynamics of 1-D and 2-D brightness perception: a unified model of classical and recent phenomena. *Percept. Psychophys.* 43, 241–277. doi: 10.3758/BF03207869
- Grossberg, S., and Williamson, J. R. (1999). A self-organizing neural system for learning to recognize textured scenes. *Vision Res.* 39, 1385–1406. doi: 10.1016/S0042-6989(98)00250-8
- Grossberg, S., and Williamson, J. R. (2001). A neural model of how horizontal and interlaminar connections of visual cortex develop into adult circuits that carry out perceptual groupings and learning. *Cereb. Cortex* 11, 37–58. doi: 10.1093/cercor/11.1.37
- Grossberg, S., and Yazdanbakhsh, A. (2005). Laminar cortical dynamics of 3D surface perception: stratification, transparency, and neon color spreading. *Vision Res.* 45, 1725–1743. doi: 10.1016/j.visres.2005.01.006
- Grossberg, S., Yazdanbakhsh, A., Cao, Y., and Swaminathan, G. (2008). How does binocular rivalry emerge from cortical mechanisms of 3-D vision? *Vision Res.* 48, 2232–2250. doi: 10.1016/j.visres.2008.06.024
- Hodgkin, A. L. (1964). *The Conduction of the Nervous Impulse*. Springfield, IL: Charles C. Thomas.
- Hong, S., and Grossberg, S. (2004). A neuromorphic model for achromatic and chromatic surface representation of natural images. *Neural Netw.* 17, 787–808. doi: 10.1016/j.neunet.2004.02.007
- Howard, I. P., and Rogers, B. J. (1995). *Binocular Vision and Stereopsis*. New York, NY: Oxford University Press.
- Hubel, D. H., and Livingstone, M. S. (1987). Segregation of form, color, and stereopsis in primate area 18. *J. Neurosci.* 7, 3378–3415.
- Hubel, D. H., and Wiesel, T. N. (1968). Receptive fields and functional architecture of monkey striate cortex. *J. Physiol.* 195, 215–243.
- Ito, M., and Gilbert, C. D. (1999). Attention modulates contextual influences in the primary visual cortex of alert monkeys. *Neuron* 22, 593–604. doi: 10.1016/S0896-6273(00)80713-8
- Julesz, B. (1971). *Foundations of Cyclopean Perception*. Chicago: The University of Chicago Press.
- Kapadia, M. K., Ito, M., Gilbert, C. D., and Westheimer, G. (1995). Improvement in visual sensitivity by changes in local context: parallel studies in human observers and in V1 of alert monkeys. *Neuron* 15, 843–856. doi: 10.1016/0896-6273(95)90175-2
- Kelly, F. J., and Grossberg, S. (2000). Neural dynamics of 3-D surface perception: figure-ground separation and lightness perception. *Percept. Psychophys.* 62, 1596–1619. doi: 10.3758/BF03212158
- Lamme, V. A. F., Rodriguez-Rodriguez, V., and Spekreijse, H. (1999). Separate processing dynamics for texture elements, boundaries and surfaces in primary visual cortex of the Macaque monkey. *Cereb. Cortex* 9, 406–413. doi: 10.1093/cercor/9.4.406
- Marr, D., and Poggio, T. (1976). Cooperative computation of stereo disparity. *Science* 194, 283–287. doi: 10.1126/science.968482
- McKee, S. P., Bravo, M. J., Smallman, H. S., and Legge, G. E. (1995). The "uniqueness constraint" and binocular masking. *Perception* 24, 49–65. doi: 10.1068/p240049
- McKee, S. P., Bravo, M. J., Taylor, D. G., and Legge, G. E. (1994). Stereo matching precedes dichoptic masking. *Vision Res.* 34, 1047–1060. doi: 10.1016/0042-6989(94)90009-4

- McLoughlin, N. P., and Grossberg, S. (1998). Cortical computation of stereo disparity. *Vision Res.* 38, 91–99. doi: 10.1016/S0042-6989(97)00122-3
- Mingolla, E., Ross, W., and Grossberg, S. (1999). A neural network for enhancing boundaries and surfaces in synthetic aperture radar images. *Neural Netw.* 12, 499–511. doi: 10.1016/S0893-6080(98)00144-0
- Nakayama, K., and Shimojo, S. (1990). da Vinci stereopsis: depth and subjective occluding contours from unpaired image points. *Vision Res.* 30, 1811–1825. doi: 10.1016/0042-6989(90)90161-D
- Panum, P. L. (1858). *Physiologische Untersuchungen ueber das Sehen mit zwei Augen*. Kiel: Schwerssche Buchhandlung (translated by C. Hubscher, 1940, Hanover, NH: Dartmouth Eye Institute).
- Paradiso, M. A., and Nakayama, K. (1991). Brightness perception and filling-in. *Vision Res.* 31, 1221–1236. doi: 10.1016/0042-6989(91)90047-9
- Pessoa, L., and Neumann, H. (1998). Why does the brain fill-in? *Trends Cogn. Sci.* 2, 422–424. doi: 10.1016/S1364-6613(98)01237-6
- Pessoa, L., Thompson, E., and Noë, A. (1998). Finding out about filling-in: a guide to perceptual completion for visual science and the philosophy of perception. *Behav. Brain Sci.* 21, 723–802. doi: 10.1017/S0140525X98001757
- Poggio, G. F. (1991). “Physiological basis of stereoscopic vision,” in *Vision and Visual Dysfunction. Binocular Vision*, ed D. Regan (London: MacMillan), 224–238.
- Polat, U., Mizobe, K., Pettet, M. W., Kasamatsu, T., and Norcia, A. M. (1998). Collinear stimuli regulate visual responses depending on cell’s contrast threshold. *Nature* 391, 580–584. doi: 10.1038/35372
- Raizada, R., and Grossberg, S. (2001). Context-sensitive bindings by the laminar circuits of V1 and V2: a unified model of perceptual grouping, attention, and orientation contrast. *Vis. Cogn.* 8, 431–466. doi: 10.1080/13506280143000070
- Raizada, R., and Grossberg, S. (2003). Towards a theory of the laminar architecture of cerebral cortex: computational clues from the visual system. *Cereb. Cortex* 13, 100–113. doi: 10.1093/cercor/13.1.100
- Ramachandran, V. S., and Nelson, J. I. (1976). Global grouping overrides point-to-point disparities. *Perception* 5, 125–128. doi: 10.1068/p050125
- Roelfsema, P. R., Lamme, V. A. F., and Spekreijse, H. (1998). Object-based attention in the primary visual cortex of the Macaque monkey. *Nature* 395, 376–381. doi: 10.1038/26475
- Roelfsema, P. R., and Spekreijse, H. (1999). Correlates of gradual spread of attention over a traced curve in Macaque area V1. *Soc. Neurosci. Abstr.* 7.2.
- Rossi, A. F., Rittenhouse, C. D., and Paradiso, M. A. (1996). The representation of brightness in primary visual cortex. *Science* 273, 1104–1107. doi: 10.1126/science.273.5278.1104
- Smallman, H. S., and McKee, S. P. (1995). A contrast ratio constraint on stereo matching. *Proc. R. Soc. Lond. B*, 260, 265–271. doi: 10.1098/rspb.1995.0090
- Tausch, R. (1953). Die beidäugige Raumwahrnehmung—ein Prozess auf Grund der Korrespondenz und Disparation von Gestalten anstelle der Korrespondenz oder Disparation einzelner Netzhautelemente. *Z. Exp. Angew. Psychol.* 1, 394–421.
- Tyler, C. W., and Kontsevich, L. L. (1995). Mechanisms of stereoscopic processing: stereoattention and surface perception in depth reconstruction. *Perception* 24, 127–153. doi: 10.1068/p240127
- von der Heydt, R., and Peterhans, E. (1989). Mechanisms of contour perception in monkey visual cortex. I. lines of pattern discontinuity. *J. Neurosci.* 9, 1731–1748.
- von der Heydt, R., Peterhans, E., and Baumgartner, G. (1984). Illusory contours and cortical neuron responses. *Science* 224, 1260–1262. doi: 10.1126/science.6539501
- Wilde, K. (1950). Der Punktreiheneffekt und die Rolle der binocularen Querdisparation beim Tienfenshen. *Psychol. Forsch.* 23, 223–262. doi: 10.1007/BF00416941
- Yazdanbakhsh, A., and Watanabe, T. (2004). Asymmetry between horizontal and vertical illusory lines in determining the depth of their embedded surface. *Vision Res.* 44, 2621–2627. doi: 10.1016/j.visres.2004.05.021

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 14 November 2013; accepted: 16 June 2014; published online: 05 August 2014.

Citation: Cao Y and Grossberg S (2014) How the venetian blind percept emerges from the laminar cortical dynamics of 3D vision. *Front. Psychol.* 5:694. doi: 10.3389/fpsyg.2014.00694

This article was submitted to *Perception Science*, a section of the journal *Frontiers in Psychology*.

Copyright © 2014 Cao and Grossberg. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.