A LAMINAR CORTICAL MODEL FOR 3D PERCEPTION OF SLANTED AND CURVED SURFACES AND OF 2D IMAGES: DEVELOPMENT, ATTENTION, AND BISTABILITY

Stephen Grossberg[1] and Gurumurthy Swaminathan[2]

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

Running Title: A cortical model of curved 3D surface perception

Technical Report CAS/CNS TR-2003-002

January 2003
Revised: October, 2003

All correspondence should be addressed to:

Prof. Stephen Grossberg
Department of Cognitive and Neural Systems
Boston University
677 Beacon Street, Boston, MA 02215
TEL: (617) 353-7858, FAX: (617) 353-7755
Email: steve@bu.edu
http://www.cns.bu.edu/Profiles/Grossberg

ABSTRACT

A model of laminar visual cortical dynamics proposes how 3D boundary and surface representations arise from viewing slanted and curved 3D objects and 2D images. The 3D boundary representations emerge from non-classical receptive field interactions within intracortical and intercortical feedback circuits. Such non-classical interactions within cortical areas V1 and V2 contextually disambiguate classical receptive field responses to ambiguous visual cues using cells that are sensitive to colinear contours, angles, and disparity gradients. Remarkably, these cell types can all be explained as variants of a unified perceptual grouping circuit whose most familiar example is a 2D colinear bipole cell. Model simulations show how this circuit can develop cell selectivity to colinear contours and angles, how slanted surfaces can activate 3D boundary representations that are sensitive to angles and disparity gradients, how 3D filling-in occurs across slanted surfaces, how a 2D Necker cube image can be represented in 3D, and how bistable 3D Necker cube percepts occur. The model also explains data about slant aftereffects and 3D neon color spreading. It shows how chemical transmitters that habituate, or depress, in an activity-dependent way can help to control development and also to trigger bistable 3D percepts and slant aftereffects. Attention can influence which of these percepts is perceived by propagating selectively along object boundaries.

Keywords: 3D perceptual grouping, 3D surface perception, cortical development, attention, angle cells, disparity-gradient cells, Necker cube, bistable percept, slant aftereffect, filling-in, V1, V2, V4, LAMINART model, FACADE model

1 Introduction

A central problem for visual neuroscience concerns how 3D objects are represented by the human visual system. Computational models that deal with 3D inputs typically concentrate only on planar objects. However, most of the objects in the world are slanted, tilted, or curved and span multiple depths with respect to an observer. In this article, the term *planar* refers to frontoparallel planar objects while the terms *slanted* and *tilted* refer to slanted and tilted planar objects, where slant is defined as deviation around the horizontal axis and tilt is defined as deviation around the vertical axis. Both binocular cues, such as disparity, and monocular cues, such as perspective, shading, and junctions, provide information about slant and tilt of an object. This article proposes how the brain combines monocular and binocular cues in a context-sensitive way to represent and perceive the 3D structure of slanted, tilted, and curved objects.

Monocular cues taken by themselves can be ambiguous. Consider Figure 1a where the two objects are made up of same set of surfaces. Depending on how the individual surfaces are combined, we perceive two different 3D objects. The same parallelogram can signal a near-to-far or a far-to-near slanted surface, depending upon the context. Contextual cues thus play a key role in disambiguating ambiguous local cues. In response to some 2D images, such as Necker cube images, the percept changes over time and depends on various factors such as attention and internal receptive field biases (Kawabata, 1986).
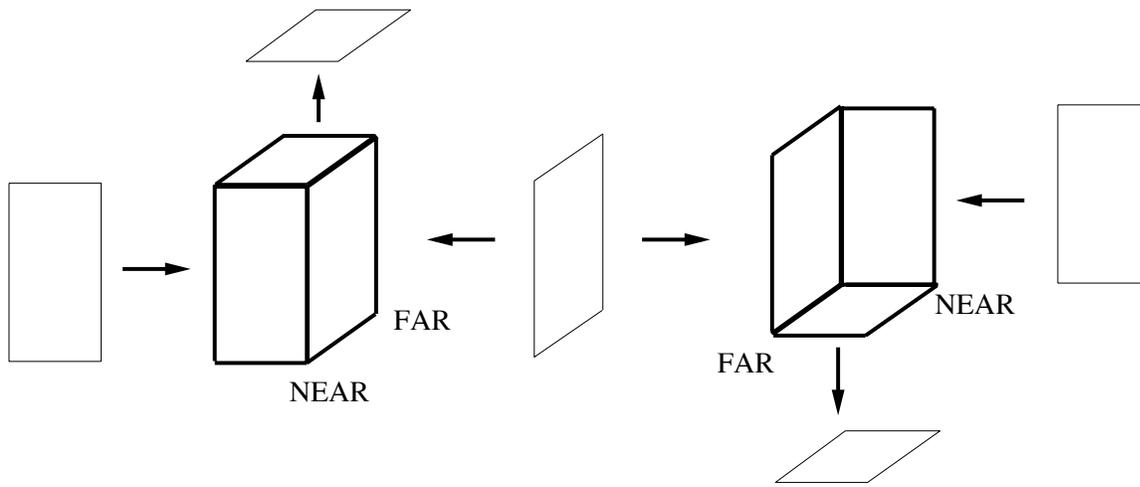
(a)
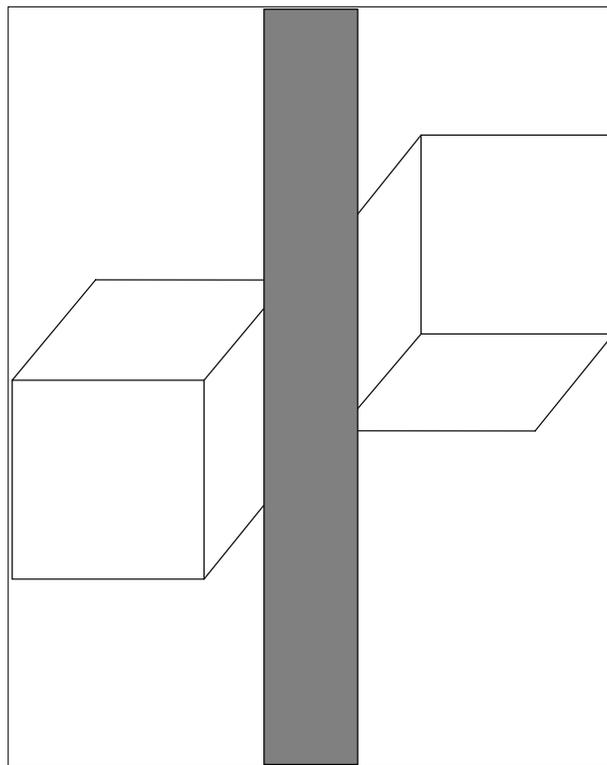
(b)

Figure 1. (a) Same angles and shapes, but different surface tilts: The two figures in bold lines are made of same set of surfaces. But due to the different arrangement of surfaces they give rise to completely different percept. The left bold figure has a positive tilt (near to far) while the right bold figure has a negative tilt (far to near). (b) Even though the sides of the cube are colinear in 2D, they are not colinear in their 3D interpretation (Tse, 1999).

Binocular disparity is a common cue for generating 3D planar percepts (Grossberg and Howe, 2002; Grossberg and McLoughlin, 1997; Julesz, 1971; Marr and Poggio, 1976; Ozhawa, 1998). Disparity information can also be used to determine the slant of an object. A slanted object is registered at multiple disparities and these representations need to be grouped across depth for it to be perceived as a single object. Information about tilt and curvature of an object can also be gleaned from disparity cues.

Neurophysiolgical studies have found cells in extrastriate cortex to be tuned to features important in 3D perception. In Macaque cortical area V2, cells are tuned to relative disparity (Thomas et al., 2002), disparity edges (von der Heydt et al., 2000), angles (Pasupathy and Connor, 1999), border ownership (Zhou et al., 2000) and figure-ground relations (Bakin et al., 2000). There is evidence for cells tuned to slanted 3D boundaries in V4 (Hinkle and Connor, 2001). Curvature tuning is found in V4 (Pasupathy and Connor, 2001), IT (Janssen et al., 2000), and parietal cortex (Taira et al., 2000).

Psychophysical studies have shown the importance of relative disparity, or disparity gradients, in human visual perception. Targets specified by a different stereoscopic slant than the distracters are detected pre-attentively (Holliday and Braddick, 1991) and so are targets presented on a surface of different slant than that of the distracters (He and Nakayama, 1995; Nakayama and Silverman, 1986). Also, multi-element tracking results do not differ if the elements are on a planar or a slanted surface (Viswanathan and Mingolla, 1999). Ryan and Gillam (1993) provided evidence that three-dimensional aftereffects can result from disparity gradient adaptation by showing that the size of the aftereffect varied with the disparity gradient of the adapting lines. Lee (1999) showed that the size of aftereffects is also dependent on the difference in disparity between the adapting and test surfaces. Many models of perceptual grouping (Grossberg and Mingolla, 1985a; Guy and Medioni, 1996; Williams and Jacobs, 1995) deal with grouping of 2D percepts. Grouping of objects, however, typically takes place in three dimensions. Illusory surface experiments (Nakayama and Shimojo, 1992) illustrate that depth needs to be taken into account during grouping. Although in some cases, 2D grouping principles work well on the 2D projection of 3D images, in other cases, 2D grouping principles gives rise to a different result than the 3D percept. For example, in Figure 1b, even though the two lines of the cube are colinear in the 2D plane, they are not colinear in the 3D interpretation (Tse, 1999) and hence are not grouped.

Grossberg and colleagues (Grossberg, 1984, 1994; Grossberg and Mingolla, 1985a; Grossberg and Todorović, 1988) have proposed that the grouping of boundaries and the filling-in of surfaces are distinct, indeed complementary (Grossberg, 2000), processes. Whereas boundaries complete inwardly in an oriented fashion, surfaces fill-in outwardly in an unoriented fashion until a boundary is reached. The outward filling-in process needs to be controlled across multiple depth planes when it fills-in 3D curved surfaces. A potential problem is that a multiple-depth boundary may have gaps at some depths, but not others, which could allow spreading colors and brightnesses to spill out during filling-in. A related problem involved in filling-in of 3D curved surfaces is clearly seen in 3D illusory displays (Carman and Welch, 1992; Liinasuo et al., 2000). Here the filling-in signal needs to spread in a controlled way across depths where there are no boundaries or filling-in inducers in the original images.

This article develops a neural model of 3D curved object representation wherein object fragments at multiple depth planes can be grouped together by disparity-gradient cells that are sensitive to an object's slant and tilt. These disparity-gradient cells can also form illusory contours in curved 3D neon color displays. The model also includes cells that are tuned to angles and explains how disparity-gradient and angle cells can be self-organized by principles that have been previously been used to self-organize 2D colinear bipole grouping cells (Grossberg and Williamson, 2001). The model hereby proposes that the statistics of the visual environment help to determine the distribution of colinear bipole cells within one depth, colinear bipole cells across depths (disparity-gradient cells), and non-colinear bipole cells (angle cells) as variations of a single design theme of how horizontal connections form in cortical layer 2/3A. The model clarifies how monocular cues in an image, notably combinations of angles, can bias the activation of some disparity-gradient cells more than others to form a 3D percept in response to 2D images, such as Necker cube images. Activity-dependent habituative mechanisms also occur in the model. Habituation is essential for the development of disparity-gradient and angle cells as well as of other properties of cortical cells (Grossberg and Seitz, 2003; Grunewald and Grossberg, 1998; Olson and Grossberg, 1998). These habituative mechanisms can lead to multi-stable percepts when two or more 3D interpretations of a 2D image are approximately equally salient, as in Necker cube percepts. The model also explains how filling-in can be carried out across multiple depths.

The present model is called the 3D LAMINART model because it generalizes to the explanation of 3D data a previously described LAMINART model which proposes how 2D perceptual grouping, attention, development, and learning are carried out by the laminar circuits of cortical areas V1 and V2 (Grossberg, 1999; Grossberg et al., 1997; Grossberg and Raizada, 2000; Grossberg and Williamson, 2001; Raizada and Grossberg, 2001). The LAMINART model was extended to explain data about stereopsis and 3D *planar* surfaces in Grossberg and Howe (2003). This extension showed how earlier modeling concepts from the FACADE model of 3D vision and figure-ground perception (Grossberg, 1994; Grossberg and McLoughlin, 1997; McLoughlin and Grossberg, 1998) could be embedded consistently within the LAMINART model circuits, and further developed to explain psychophysical, neurophysiological, and anatomical data about stereopsis and 3D planar surface perception. Grouping mechanisms were not needed to simulate the targeted data in Grossberg and Howe (2003), although it was proposed how this 3D LAMINART model could also explain 3D *planar* grouping data. The present article shows how this 3D LAMINART model can be further extended, again in a self-consistent way, to explain psychophysical, neurophysiological, and anatomical data about the perception of slanted and curved 3D surfaces. Here, 3D grouping mechanisms are essential to explain targeted data.

This paper is organized as follows. Section 2 describes the laminar architecture of the model. Section 3 provides an overview of how the model interprets 2D images in 3D. Section 4 discusses the model simulations that show how the long-range horizontal connections in V1 can develop into colinear and angle cells and how various Necker cube 2D images are interpreted in 3D. Section 5 discusses how the model can explain various data related to 3D grouping and slant aftereffects. It also explains how the monocular and binocular cues can interact in the model to give rise to a stable representation, and compares the present model with alternative models. The mathematical description of the model is described in the Appendix.

2   Model description

The model carries out this extension by adding three key cell types that are needed to control 3D grouping, as shown in model block diagram in Figure 2. The three key additions in the model are (1) colinear bipole cells, (2) non-colinear bipole cells (angle cells) and (3) disparity-gradient cells (Figure 2a). Colinear bipole cells played an important role in the original 2D LAMINART model

to carry out perceptual grouping and boundary completion. They were not needed to simulate the data about planar 3D surface perception considered by Grossberg and Howe (2003) because boundary completion was not required to explain these data. In the present analysis, colinear and non-colinear bipole cells get activated by line segments and angles in the images, respectively. They activate the disparity-gradient cells that group boundaries across depth. This multiple-depth boundary representation by disparity-gradient cells is used to control filling-in of slanted and curved surfaces. The mathematical description of the model is described in the Appendix. We give an overview of the model and describe each of its novel features in detail in the following sections.

(a)   **(b)**
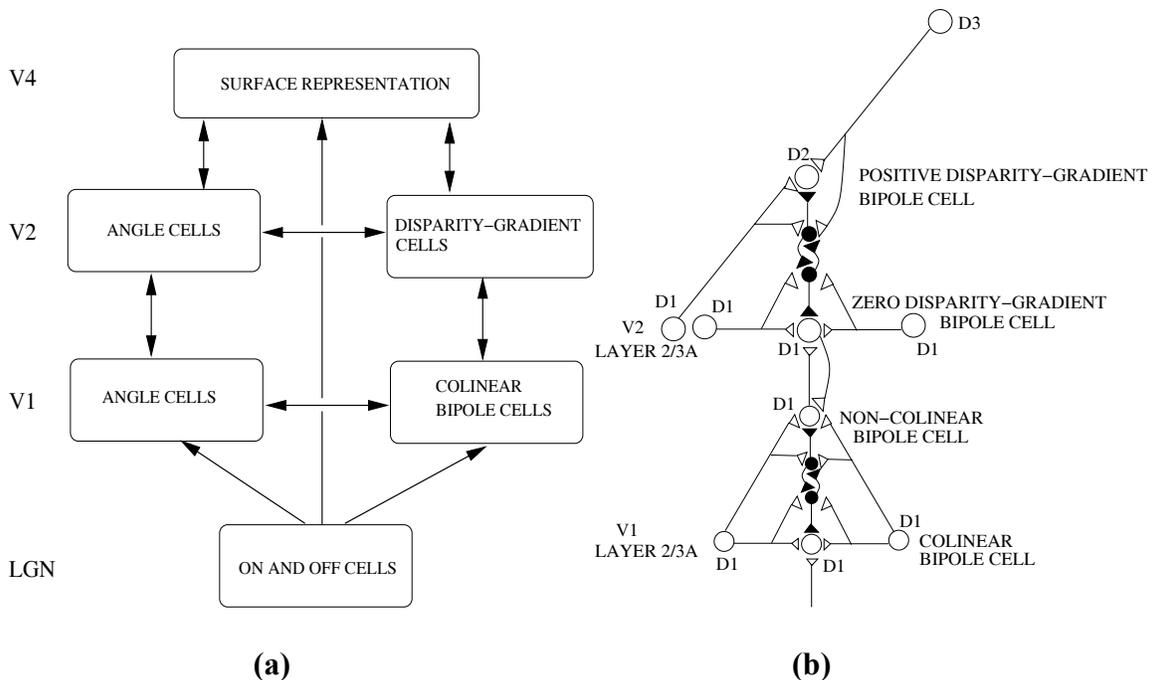
Figure 2. (a) Block diagram of the model: The input image undergoes on-center, off-surround processing in the LGN. In layer 2/3A of V1, angle cells and colinear bipole cells get activated by angles and line segments in the images, respectively. Angle cells and colinear bipole cells interact with each other via long-range horizontal connections in layer 2/3A of V1. Colinear bipole cells activate disparity-gradient cells, while V1 angle cells activate V2 angle cells. V2 angle cells and disparity-gradient cells interact via long-range horizontal connections in layer 2/3A of V2. Disparity-gradient cells group across position and disparity to form closed boundary segments, which is used as a barrier for filling-in of surfaces in V4 which receives filling-in signal from the LGN. (b) Laminar circuit for 3D boundary grouping: V1 angle cells and colinear bipole cells are in layer 2/3A of V1. Layer 2/3A cells in V1 activate layer 2/3A cells in V2. Layer 2/3A of V2 contains V2 angle cells and disparity-gradient cells. Layer 2/3A of V2 feedback to layer 2/3A of V1. In the full 3D LAMINART model this feedback is mediated via layer 1 (Figure 3). Disparity-gradient cells group across disparity gradient and disparities. D1, D2, and D3 represent various depths. Open (black) circles (triangles) represent excitatory (inhibitory) cells (connections).

6
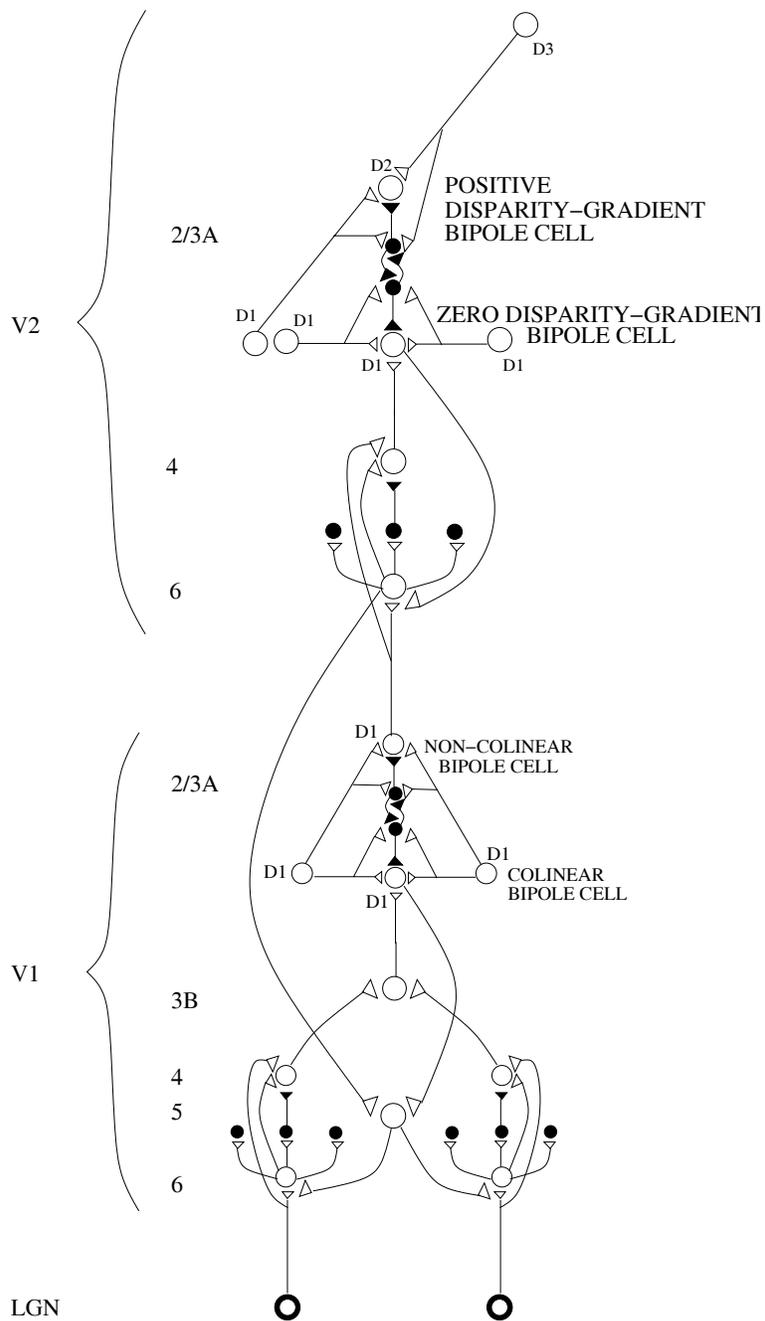
Figure 3. 3D LAMINART MODEL: The 2D LAMINART model (Grossberg and Raizada, 2000) is extended to 3D. The LGN provides bottom-up activation to layer 4 directly and via layer 6 → 4 on-center off-surround pathway, which provides divisive contrast normalization (Grossberg,1973, 1980; Heeger, 1992). Monocular simple cells in layer 4 activate binocular simple cells in layer 3B. Layer 2/3A complex cells combine the output of contrast sensitive simple cells to get contrast insensitive output. Layer 2/3A consists of angle cells and colinear bipole cells. Layer 2/3A activates layer 4 of V2 directly and via layer 6 as was the case for V1. Layer 2/3A of V2 consists of disparity-gradient cells and V2 angle cells. Open (black) circles (triangles) show excitatory (inhibitory) neurons (connections). V2 layer 2/3A cells feedback onto V1 layer 2/3A cells via layer 6 of V2, layer 5 and 6→4 of V1.

2.1 Laminar Architecture

The laminar architecture of the model is show in Figure 2b. Model circuits are consistent with all the anatomical and neurophysiological constraints that were used to explain many other types of data using earlier versions of the 3D LAMINART and FACADE models. Definitive anatomical and neurophysiological experiments on 3D slanted and curved surface perception have not yet been performed. We nonetheless interpret all model cell types using the best information available in order to make testable predictions that can guide future experiments. It is conceivable that a cell type which we localize in V2 may occur in a different cortical area in some species. What cannot change however, without altering key functional properties is the *order* in which various model operations occur.

In order to keep the simulations tractable, the model omits interactions in layers 1, 4, and 6 that are not required to explain its targeted 3D grouping data. Figure 3 shows how these interactions can be consistently embedded into a more complete 3D LAMINART model.

Layer 2/3A of V1 contains complex cells. These cells combine the outputs from simple cells that are sensitive to the same orientation but opposite contrast polarities (Callaway, 1998; Poggio, 1972). How the inputs from the two eyes are combined in by circuits in layers 3B and 2/3A has been quantitatively modeled in related work (Grossberg and Howe, 2002). Here we assume that these inputs to layer 2/3A have already been computed.

The complex cells in layer 2/3A are assumed to be of two kinds: They are (1) colinear bipole cells that link colinear line segments, or other oriented contrast gradients, over short distances and (2) non-colinear bipole cells (angle cells) that get activated by various angles in an image of the scene. Section 4.1 shows how model cells in layer 2/3A can self organize into colinear and non-colinear bipole cells by developing layer 2/3A horizontal connections. Layer 2/3A of V1 connects to layer 2/3A of V2 (Tootell and Hamilton, 1989). Layer 2/3A of V2 has two kinds of model cells: disparity gradient cells and angle cells (Pasupathy and Connor, 1999). The angle cells in V2 are similar to the ones in layer 2/3A of V1 and receive input from V1 angle cells. The disparity-gradient cells of V2 receive inputs from V1 colinear bipole cells and link cells of different disparities to form straight or curved segments in 3D.

The formation of curved boundaries in 3D using disparity-gradient cells naturally generalizes how curved boundaries are formed in 2D using colinear bipole cells. The receptive field of a colinear bipole cell *prefers* to group cells that are colinear across space with respect to

the cell's position and preferred orientation, and also have the same preferred orientation. Bipole cells can, however, also group cell activations that deviate from colinearity and the preferred orientational preference to form curved groupings (Gove et al., 1995; Grossberg and Mingolla, 1985b). Later psychophysical experiments have supported this predicted bipole cell receptive field; e.g., Field et al. (1993) and Kellman and Shipley (1991). Kellman and Shipley (1991) have called this ability to group curves *relatability conditions*. Recent data have shown that these conditions generalize to 3D (Kellman, 2003). Disparity-gradient cells can generate such curved boundaries in 3D by naturally generalizing the 2D colinear bipole cell receptive field; see Appendix Section B.4.4.

Appropriate combinations of angle cells help to select the correct disparity-gradient cells (flat, near-to-far, far-to-near) through contextual interactions. The angle cells and the disparity-gradient cells are proposed to interact with each other via horizontal connections in layer 2/3A of V2 (Amir et al., 1993). Disparity-gradient and angle cells in layer 2/3A of V2 feed back into V1 layer 2/3A colinear and angle bipole cells, respectively. In the model in Figure 3, this feedback is mediated via layer 6 of V2 and layer 6→4 interactions of V1. The feedback enhances V1 layer 2/3A cells that are supported by V2 groupings, while suppressing non-supported cells. In the model shown in Figure 3, the top-down V2-to-V1 feedback has an on-center off-surround form (via 6→4 interactions in V1), which is consistent with data of Hupé et. al. (1998) and was modeled in Grossberg (1999) and Grossberg and Raizada (2000). This property is also consistent with results of Lee and Blake (2002), who showed that V1 activity is reduced during binocular rivalry. In the model, off-surround interactions suppress the non-dominant eye signals.

The present article focuses on straight groupings in 3D—that is, slanted groupings—but these results directly generalize to curved surfaces in 3D in much the same way that they do in 2D.

2.2   V1 colinear bipole cells

It is known that layer 2/3A of V1 has long-range horizontal connections (Callaway, 1998). These intra-laminar connections primarily connect to cells of similar orientation (Bosking et al., 1997; Schmidt et al., 1997). Such connections have been used to explain psychophysical and neural data about attention and perceptual grouping (Grossberg and Raizada, 2000). In the present model, the long-range horizontal connections of colinear bipole cells link line segments over

short distances among cells that are sensitive to the same disparity. Grossberg and Williamson (2001) showed how these connections can develop within the laminar circuits of the visual cortex.

Colinear bipole cell excite each other via long-range horizontal connections that also give rise to short-range disynaptic inhibition via pools of interneurons; see Figure 2b. This balance of excitation and inhibition at target cells helps to implement the *bipole property*. When the inducing stimulus (e.g., a pacman) is only on one side, it excites the corresponding oriented receptive fields of layer 2/3A cells, which send out long-range horizontal excitation onto the target cell. However, this excitation also activates a commensurate amount of disynaptic inhibition (as in Figure 2b). This creates a case of "one-against-one", and the target cell is not excited above-threshold. However, the cell activity can be modulated by input from even one side if the cell receives bottom-up input (Bringuier et al., 1999; Crook et al., 2002). The modulation is achieved by combining the bottom-up input with input from long-range horizontal connections (see Appendix, equation (3)). The combined bottom-up and horizontal input from one side can overcome the disynaptic inhibition from the inhibitory interneurons and thus can activate the cell. These modulations play an important role in the spreading of attention (Grossberg and Raizada, 2000; Ito and Gilbert, 1999; Roelfsema et al., 1998; Roelfsema and Spekreijse, 1999), the grouping of 2D and 3D planar percepts (Bakin et al., 2000; Kapadia et al., 1995; Polat et al., 1998), and the grouping of 3D slanted and curved percepts, as discussed below. When two colinearly aligned inducing stimuli are present, one on each side, a boundary grouping can form: Long-range excitatory inputs converge onto the cell from both sides and summate. These excitatory inputs also activate a shared pool of inhibitory interneurons, which as well as inhibiting the target cell, also inhibit each other, thus normalizing the total amount of inhibition emanating from the interneuron pool. This summating excitation and normalizing inhibition create a case of "two-against- one" and the target cell is excited above-threshold (von der Heydt and Peterhans, 1989; von der Heydt et al., 1984). When there is direct bottom-up input, it can activate the cell without horizontal interactions.

2.3 V1 and V2 angle cells

There is direct neurophysiological evidence of cells tuned to angles in area 17 of the cat, which is homologous to V1 in the Macaque monkey (Shevelev, 1998), and in V2 (Hegde and Van Essen,

2000) and V4 (Pasupathy and Connor, 1999) of the Macaque. Cells are tuned to both angles and to the orientation of the angles. Some cells are tuned to a particular angle with a particular orientation. Some are tuned to various angles of a particular orientation; that is, they get activated strongly by different angles that have a common orientation, but weakly to colinear line segments; and some are tuned to an angle at any orientation.

The previous section discussed how long-range horizontal connections in layer 2/3A of V1 are used by colinear bipole cells to link line segments over short distances. These long-range connections can also get tuned to angles in the images (Grossberg and Mingolla, 1987; Neumann and Stiehl, 1990). In the model, layer 2/3A of V1 contains bipole cells, called non-colinear bipole cells that are tuned to angles by means of long-range horizontal connections that connect to different orientations. These non-colinear bipole cells have similar properties to layer 2/3A colinear bipole cells. They get input from other cells from two sides, or even three sides, depending on the angle that the cell represents. In the absence of direct bottom-up input, such cells get activated only if they receive sufficient excitation from all their sides. How the long-range horizontal connections in layer 2/3A can develop into angle cells and colinear bipole cells is shown in Section 4.1.

Angle cells are also present in layer 2/3A of V2 (Hegde and Van Essen, 2000). The model includes angle cells, named V2 angle cells, in layer 2/3A of V2. V2 layer 2/3A also has horizontal connections, but these are longer than those in V1 layer 2/3A (Amir et al., 1993; Grosof et al, 1993; Ramsden et al., 2001; von der Heydt et al., 1984). The V2 angle cells are similar to V1 angle cells and receive bottom-up input from V1 angle cells and horizontal input from disparity-gradient cells in the model.

2.4 V2 disparity-gradient cells

Many psychophysical data describe how the visual system handles slanted and curved surfaces. Humans are often more sensitive to relative disparities—that is, disparity differences between center and surround stimuli—than absolute disparities (DeAngelis, 2000). The aftereffect experiments done in Ryan and Gillam (1993) provided evidence that three-dimensional aftereffects can result from disparity-gradient adaptation by showing that the size of the aftereffect varied with the disparity gradient of the adapting lines. Lee (1999) showed that the size of the aftereffect is also dependent on the disparity difference between the adapting and

testing surface. He also showed that the slant of the aftereffect produced is always opposite to the slant of the adapting surface indicating that the aftereffects are mediated by cells coding slant—that is, disparity gradients. Seyama et al. (2000) found that aftereffects can also depend on the tilt difference between the testing and the adapting surface. The above data show the importance of disparity gradients in the representation of slanted surfaces.

On the neurophysiological side, there is evidence for cells tuned to relative disparity—that is, cells tuned to a constant disparity difference between center and surround independent of the disparity of the center or surround—in V2 of the macaque monkey (Thomas et al., 2002). There are also cells that are tuned to disparity gradients—that is, cells that respond to slanted or tilted bar stimuli—in V4 of the macaque (Hinkle and Connor, 2001). In MT, there are cells tuned to slant and tilt of a surface defined by motion (Nguyenkim and DeAngelis, 2001). In the parietal cortex (Sakata et al., 1999) and IT (Janssen et al., 2000), some cells are tuned to slanted or tilted bar stimuli, just like the cells in V4 mentioned before, and some cells were tuned to slanted or tilted surfaces, like the cells in MT mentioned before. These data support the existence of cells tuned to disparity gradients in the visual system.

In the model, V2 layer 2/3A contains cells that are tuned to disparity gradients. These disparity-gradient cells are sensitive to disparities, disparity gradients, and are orientationally tuned and receive bottom-up input from the colinear bipole cells in layer 2/3A of V1 and horizontal input from V2 angle cells and other disparity-gradient cells.
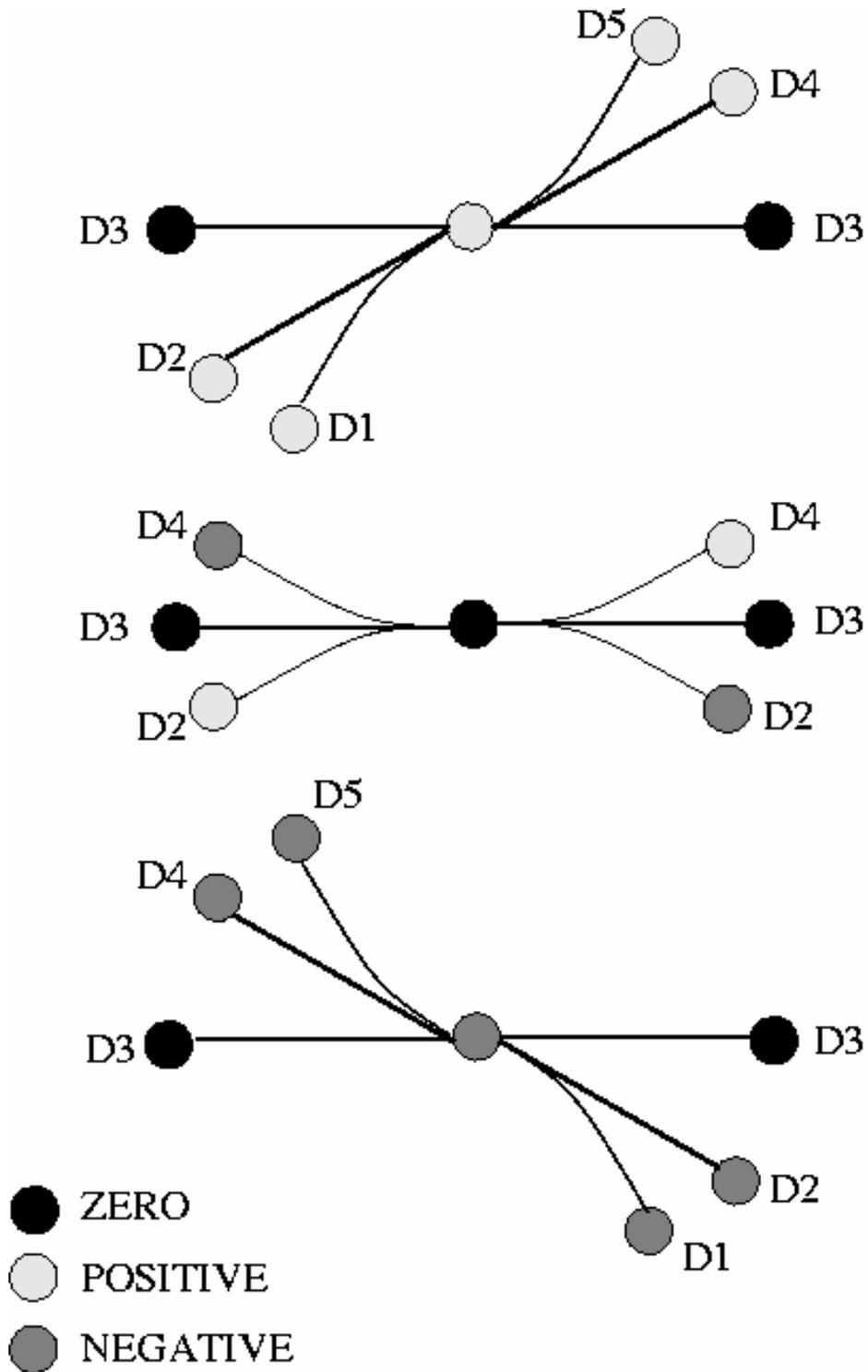
Figure 4. Disparity-gradient cells: Positive disparity-gradient cells link from near-to-far; negative disparity-gradient cells link from far-to-near; and zero disparity-gradient cells link within disparity. Black circles indicate zero disparity-gradient cells, gray circles indicate negative disparity-gradient cells, and light gray circles indicate positive disparity-gradient cells. D1- D5 indicates different depths from near (D1) to far (D5).

Figure 4 illustrates how model disparity-gradient cells connect with each other. Three cells corresponding to positive, zero and negative disparity gradients are shown. Positive disparity-gradient cells connect with other disparity-gradient cells from near depth to far depth, whereas negative disparity-gradient cells connect from far depth to near depth. Zero disparity-gradient cells connect within depth. The Appendix mathematically describes the connections between cells of different disparity-gradients and orientations that enable the cells to smoothly represent curved surfaces in both 2D and 3D space.

2.5 V4 surface representation

Once the boundaries are registered at corresponding disparities, then filling-in between these boundaries is proposed to generate visible 3D surface percepts in cortical area V4 (Grossberg, 1994; Grossberg and Todorović, 1988). The existence of a filling-in process has been supported by psychophysical (Paradiso and Nakayama, 1991; Pessoa and Neumann, 1998; Pessoa, Thompson and Noë, 1998) and neurophysiological experiments (Lamme et al., 1999; Rossi et al., 1996). A filling-in process has been used to explain many percepts, such as da Vinci stereopsis (Grossberg and Howe, 2002; Grossberg and Mcloughlin, 1997), figure-ground perception (Kelly and Grossberg, 1999), 2D and 3D neon color spreading (Grossberg, 1994; Grossberg and Mingolla, 1985a), and both monocular and binocular brightness percepts (Grossberg and Kelly, 1999; Grossberg and Todorović, 1988). In previous models (Grossberg, 1997; Kelly and Grossberg, 2000), the boundary representation at a particular depth acts as a barrier to filling-in signals only at that depth. One problem that must be solved to fill-in curved surfaces is that the boundary representation for a slanted or curved surface may have gaps at some depths even if it has no gaps at other depths. What prevents surface lightness and color from dissipating through these gaps? We call this problem the *lightness dissipation problem*.

(a)

(b)

D1

D2

D1

D2

(c)
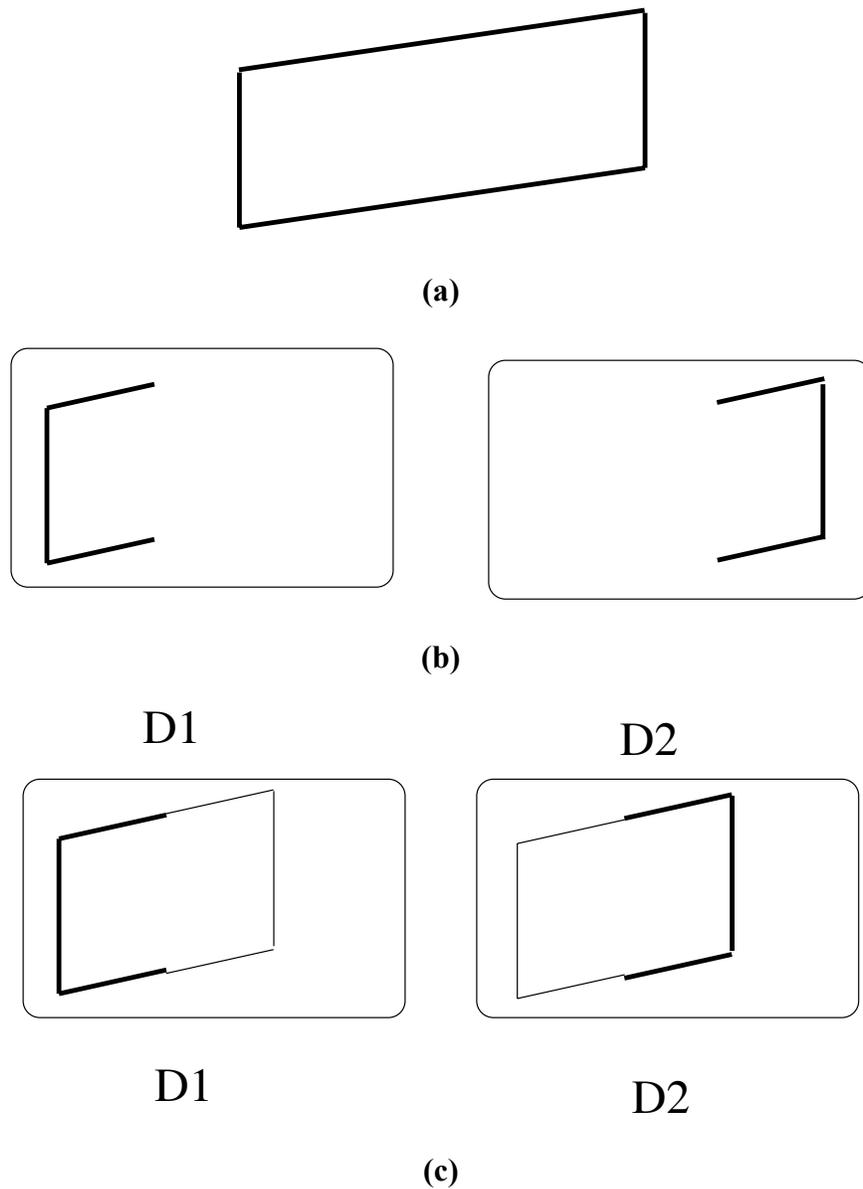
Figure 5. Filling-in of slanted surfaces. (a) The input is a tilted rectangle. (b) Multiple depth representation of the tilted rectangle. (c) Filling-in barriers: The boundary representation act as a strong filling-in barriers at the corresponding depth while acting as a weak barrier at the nearby depths thus creating closed boundary compartments within each depth. D1 (near) and D2 (far) represents different depths.

This problem is overcome in the present model as follows: A boundary signal that acts as a strong barrier to filling-in at its preferred depth also weakly acts as a barrier to filling-in at other depths. For example, consider a tilted rectangle in depth, as in Figure 5a. Each boundary

representation is activated at its preferred depths, as in Figure 5b, and this boundary representation has gaps at each depth. If no other boundaries existed, filling-in signals would flow out of the boundary gaps at each depth. The model proposes that the boundary at a particular depth is also represented, albeit weakly, at nearby other depths. This hypothesis has earlier been made to explain how a finite pool of depth-selective boundaries can control a continuous change in perceived depth (Grossberg, 1994, 1997). Here it is predicted to also contribute to percepts of slanted and curved surfaces in depth. In particular, the total boundary signal that acts as a barrier to filling-in at each depth is shown in Figure 5c. Now, a closed boundary exists at each depth, and the filling-in signal is at least partially contained at each depth. Because of differences in boundary strength, however, the filled-in activity is not uniformly strong at each position. It is stronger wherever there is a strong boundary, since lightness and color can dissipate more through a weaker boundary than a stronger one. It is shown in Section 4.6.2 how a tilted surface representation can be generated by such differential filling-in across different depths.

3   Overview of model interpretation of 2D images in 3D

An important mechanism in the model for 3D interpretation of 2D images is that angle cells contextually bias the activation of disparity-gradient cells, and grouping among disparity-gradient cells disambiguates the 3D interpretation of the 2D image. Figure 6a illustrates how different angles can bias the activation of disparity-gradient cells to favor different depth relationships (near-to-far, flat, far-to-near). Figure 6b suggests one way in which angle cells might be organized in cortical hypercolumns. This organization clusters together angle cells that bias disparity-gradient cells with similar depth preferences. Such a clustering of both angle cells and disparity-gradient cells would be a plausible outcome of cortical development, but direct evidence for it is lacking. To see how these interactions work, consider a right triangle and a parallelogram. The parallelogram can be seen as a flat 2D surface or as a tilted rectangle in 3D (Figure 1a), while the right triangle is seen only as a flat 2D surface. These two interpretations arise in the model due to the selective activation of disparity-gradient cells by V2 angle cells (Figure 6a) and the subsequent grouping among these disparity-gradient cells (Figure 4).

As discussed before, non-colinear bipole cells (angle cells) are activated by image corners while colinear bipole cells are activated by line segments. So, for the triangle and the

parallelogram, V1 angle cells get activated at corners and colinear bipole cells get activated along straight edges. This segregation of activation occurs due to the bipole property of these cells, since V1 angle cells get maximum input at the corners, whereas the colinear bipole cells get maximum input along straight edges. Colinear bipole cells directly activate the disparity-gradient cells. There are three (populations of) disparity-gradient cells, corresponding to positive, negative and zero disparity gradients, at each position, orientation, and disparity. The colinear bipole cells corresponding to the same position and orientation activate all three—zero, positive, and negative—disparity-gradient cells equally. V1 angle cells activate corresponding V2 angle cells, which in turn activate the disparity-gradient cells. An important difference is that the V2 angle cells activate the disparity-gradient cells selectively, unlike colinear bipole cells that activate all disparity-gradient cells equally. This selective activation of disparity-gradient cells by V2 angle cells is assumed to be learned from general image statistics. Another difference is that V2 angle cells activate the disparity-gradient cells at nearby positions, while the colinear bipole cells activate disparity-gradient cells at their corresponding positions.

(a)   (b)
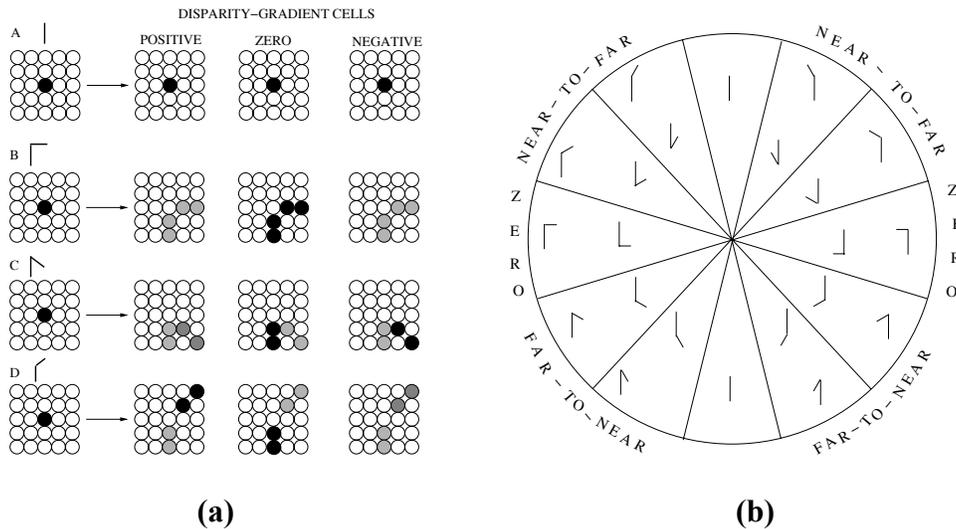
Figure 6. (a) Activation of disparity-gradient cells by angle cells and colinear bipole cells. Colinear bipole cells activate all disparity-gradient cells equally at the corresponding positions. Angle cells selectively activate disparity-gradient cells at nearby positions. Circles represent cells at various spatial positions. Shaded circle represents active cell, unshaded circle represents inactive cell. The amount of activation depends on the darkness of the shade. (b) Hypercolumn representation of angle cells: The top part of the circle represents angles that activate positive disparity-gradient cells preferentially, while the bottom part activates negative disparity-gradient cells preferentially. The middle part of the circle at the left and right of the circle represents angles that activate zero disparity-gradient cells preferentially. The middle part of the circle at the top and bottom represents colinear line segments.

The activation of disparity-gradient cells by V2 angle cells and colinear bipole cells is shown in Figure 6a. Each circle in the figure represents a cell at a particular spatial position and the color inside the circle indicates the strength of the activation. In particular, black circles indicate strong activation, white circles indicate zero activation and gray circles indicate intermediate activation. A colinear bipole cell that codes the vertical orientation activates all disparity-gradient cells—positive, zero, and negative—equally, at the same position, orientation, and disparity. The activation of disparity-gradient cells by V2 angle cells depends on the angle to which the cell is tuned. For example, the V2 angle cell tuned to angle B in the figure mostly activates zero disparity-gradient cells at nearby positions along the horizontal and vertical orientations. Similarly, the V2 angle cell that codes angle C mostly activates zero disparity-gradient cells along the vertical orientation and positive disparity-gradient cells along the oblique orientation. The V2 angle cell that codes angle D mostly activates zero disparity-gradient cells along the vertical orientation and negative disparity-gradient cells along the oblique orientation. An important point to note is that although there is a preference for an angle cell to activate a particular disparity-gradient cell strongly, it can also activate other nearby disparity-gradient cells weakly.

Figure 6b shows the arrangement of angle cell preferences in a hypercolumn structure. The angles are obtained by combining a vertical line with obliques lines of different orientation. The angles that are in the top part of the hypercolumn activate positive disparity-gradient cells more than negative disparity-gradient cells, while the angles in the bottom part activate negative disparity-gradient cells more than positive disparity-gradient cells. This selectivity can be learned from general image statistics. In particular, the angles in the top part of the hypercolumn are usually part of a tilted surface going from near to far in the real world. Similarly, the angles in the bottom part are usually part of a tilted surface that goes from far to near. The hypercolumn representation shows how the preferential activation of disparity-gradient cells by angle cells can change smoothly as the angle changes smoothly.
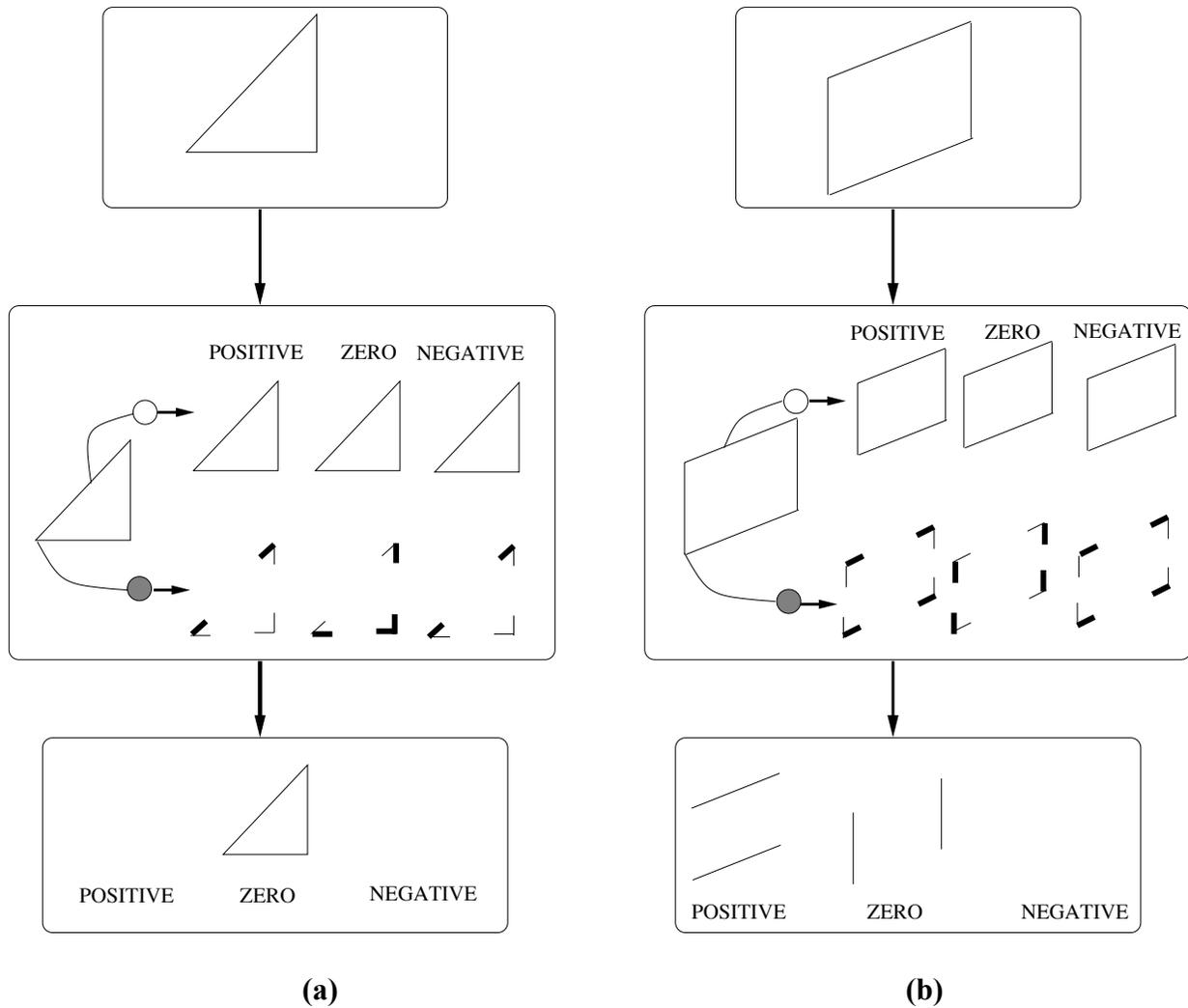
Figure 7. Activation of disparity-gradient cells by angle cells and colinear bipole cells. Top part shows the input, middle part shows the activation of disparity-gradient cells by angle cells and colinear bipole cells and the bottom part shows the _nal representation by disparity-gradient cells. Shaded circle represents angle cell and unshaded circle represents colinear bipole cell. The thickness of the lines indicates activation strength. (a) For the triangle input, the cross disparity-gradient cells are activated along oblique lines and zero disparity-gradient cells along horizontal lines. The triangle is represented by zero disparity-gradient cells within disparity. See text for details. (b) For the parallelogram input, the cross disparity-gradient cells are activated more along oblique lines and zero disparity-gradient cells are activated more along vertical lines. The parallelogram is represented by positive disparity-gradient cells across depth and zero disparity-gradient cells. See text for details.

For the triangle, zero disparity-gradient cells are activated strongly along the horizontal and vertical edges, while cross disparity-gradient cells (both positive and negative) are activated along the oblique edge. Since the zero disparity-gradient cells group more strongly within depth

19

than across depths, all the vertical and horizontal edges of the triangle are represented within depth. Thus the corner that is shared by the horizontal and vertical edge is also represented at the same depth as the edges. This interaction binds the horizontal and vertical boundaries within depth and causes the other angle cells to be activated at that depth. Thus, for the triangle, the three corners are all represented within the same depth. This enables the weakly activated zero disparity-gradient cells along the oblique edge (cf., angle C in Figure 6a) to group strongly and to inhibit the cross disparity-gradient cells. This is because zero disparity-gradient cells group preferentially within depth, while cross disparity-gradient cells group across depth. Hence, all the edges of the triangle are represented within depth, as shown in Figure 7a.

For the parallelogram, the zero disparity-gradient cells are activated strongly along the vertical edges and the cross disparity-gradient cells are activated along oblique edges. Let us assume that one of the vertical edges of the parallelogram is represented at a particular depth. Thus the corners shared by the vertical edge are represented at the same depth. The cross disparity-gradient cells along oblique edges group strongly across depths than within depth, as in angles C and D of Figure 6a. This causes the second vertical line to be pushed into a different depth than the first vertical line, as shown in Figure 7b. The difference between the triangle and the parallelogram is that the horizontal and vertical edges of the triangle share a corner which forces them to be represented at the same depth, while for the parallelogram the vertical edges do not share a corner and hence can be represented at different depths. If one of the cross disparity-gradient cells, either positive or negative, groups more strongly, the parallelogram is either seen as going from near-to-far or far-to-near, respectively. If they balance out, then the parallelogram is seen in a single depth plane.

In summary, the 3D interpretation of a 2D image starts by the activation of disparity-gradient cells by V2 angle cells in the model, and is completed by the grouping of disparity-gradient cells. The grouping uses the local preferences initiated by angle cells to enforce a globally consistent interpretation. Thus the disparity-gradient cells which developed in response to 3D image statistics for 3D grouping also help to disambiguate 3D percepts of 2D images.

4   Model simulations

This section summarizes model simulations that show how layer 2/3A horizontal connections in V1 develop into colinear bipole and angle cells with the properties described in the previous

sections. The model developmental equations are given in Appendix A. Then it is shown how the laminar model circuit can respond to 2D images containing monocular cues, such as angles, with a 3D boundary representation. This is demonstrated by simulating a Necker cube 2D image in 3D, including its bistability. These model equations are given in Appendix B. It is also shown how the model can represent 3D slanted and curved boundaries using disparity cues alone, without any monocular cues. Finally, simulations of 3D surface filling-in are carried out using these 3D boundary representations.

The simulations are done using the laminar circuit shown in Figure 2b, since the horizontal connections in V1 and V2 are rate-limiting in generating the targeted properties. The self-organization simulations show how connections in layer 2/3A of V1 develop into angle cells and colinear bipole cells within depth. A similar scheme can be used to self-organize V2 layer 2/3A cells into angle cells and disparity-gradient cells across depth. The layer 2/3A cells of V1 in both the developmental and the Necker cube simulations receive inputs that were generated by hand to be consistent with previous model simulation outputs (Grossberg and Howe, 2002; Grossberg and Raizada, 2000; Grossberg and Williamson, 2001). The input generation procedure is described for each simulation in later sections.

4.1 Development of colinear and non-colinear bipole cells

The cells in layer 2/3A of V1 have long-range horizontal connections with other cells in layer 2/3A. These horizontal connections have been shown to develop in response to endogenous and visual cues to link colinear line segments in such a way as to satisfy the bipole property (Grossberg and Williamson, 2001). Here it is shown how such cells can also get tuned to angles in the images as well as to their colinear statistics. The challenging aspect in the present simulation is that the various angles share some features. Hence, if one cell emerges a winner for a particular angle, it has the tendency to emerge as the winner for other angles that have common features with the angle that it has learned before. Habituative transmitter gates are used to overcome this tendency. The winning cell habituates and does not fire for a while in response to subsequent input presentations, thus allowing other cells to code subsequent inputs; see Section 4.1.3.

4.1.1 Simulation set-up

There are sixteen excitatory cells at each spatial position in layer 2/3A of the model. These sixteen cells at each position will code the various angle and colinear bipole cells that develop there. Half of them receive inputs from vertically oriented layer 3B cells and half from horizontally oriented layer 3B cells (Callaway, 1998; Callaway and Wiser, 1996). Cells also receive horizontal inputs from cells at different spatial positions within a defined neighborhood. The horizontal input from each spatial position was computed as the total activity, within orientation, of all the cells at that spatial position. Figure 8a shows a schematic of the simulation setup.

Each excitatory cell at a spatial position has four inhibitory interneurons associated with it whose learned interactions with the excitatory cell will give rise to the bipole property; see Figure 8b. The inhibitory interneurons have inhibitory connections with the excitatory cells and with other inhibitory interneurons. This setup models the long-range excitation and short-range inhibition found in layer 2/3A of V1 (Callaway, 1998). The inhibitory connection from the interneurons to the excitatory cell balances the excitation from the horizontal connections to ensure that the cell can fire only if it receives direct bottom-up input, or sufficient input from both sides. At the same time, the recurrent inhibitory connections among the interneurons normalize the total inhibition so that the cell can fire when its inputs satisfy the bipole property. Both the excitatory and the inhibitory adaptive weights between these cells are learned, starting from zero initial values.

As in the developmental model of Grossberg and Williamson (2001), the excitatory adaptive weights are learned using the *instar* learning law, which has become the standard law for learning self-organizing maps (Grossberg, 1976a, 1980; Kohonen, 1989). During instar learning, the activity in the postsynaptic target cell turns on learning, and the adaptive weight learns the expected value of the total signal from its presynaptic source cells during the interval when the target cell is active. An *outstar* learning law (Grossberg, 1968, 1980) was used to learn the weights between an inhibitory interneuron and its excitatory neuron, and the weights among the inhibitory interneurons. Outstar learning accomplishes the balance between inhibition and excitation (see Section 2.2) by causing the inhibitory synaptic weight to track the expected activation of the target excitatory cells at times when its source inhibitory interneuron has positive activity. Likewise, the inhibitory interneuronal weights track the positive activity of the

target inhibitory interneuron. This property enables the network to normalize the total inhibitory input from the interneurons, which enables layer 2/3A excitatory cells to fire if there is excitatory input on both sides sufficient to overcome this normalized inhibition.
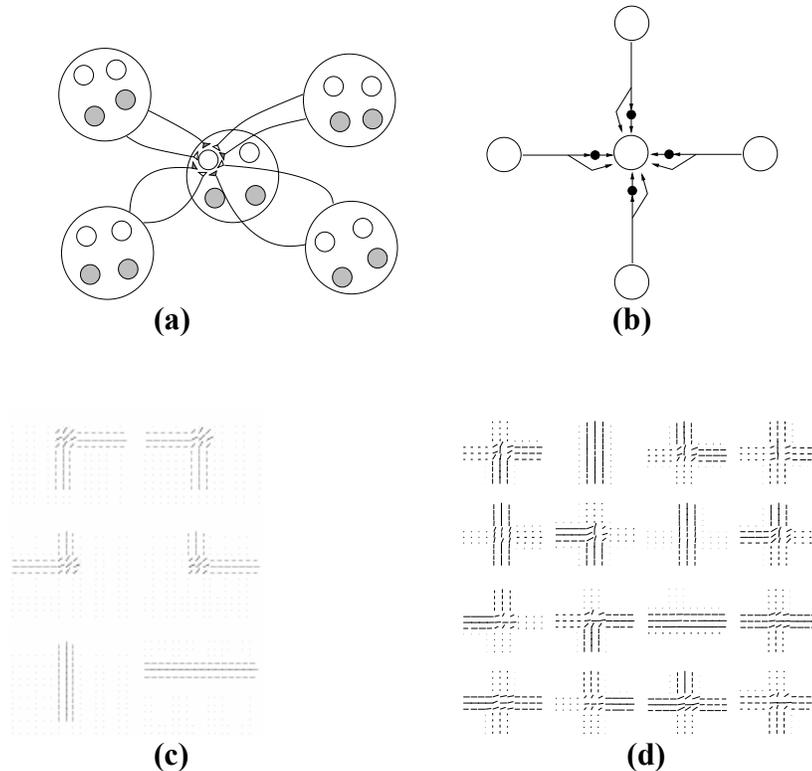
Figure 8. Developmental simulation of layer 2/3A cells into angle cells and colinear bipole cells: (a) Simulation setup. Five positions are shown in the figure. At each position there are two cells that receive bottom-up input from horizontally oriented cells, shown as shaded circles, and two cells that receive bottom-up input from vertically oriented cells, shown as unshaded circles. Each cell receives two types of connections from cells at other positions: connections from horizontally oriented cells, indicated by shaded triangles, and connections from vertically oriented cells, indicated by unshaded triangles. The weights of these connections are learned in the simulations. For simplicity, inhibitory interneurons are not shown. (b) Each excitatory cell is associated with four inhibitory interneurons (black disks). The four inhibitory interneurons receive part of the horizontal input, depending on their position (left, right, top or bottom) received by the excitatory cell at their position. (c) Inputs used in the simulation. There are six different inputs corresponding to four different right angles, and vertical and horizontal line segments. Horizontal and vertical lines represent input from horizontally and vertically oriented cells, respectively. Oblique lines indicate the presence of inputs from both horizontal and vertical cells. The length of the lines represents the strength of inputs. (d) Result of the developmental simulation showing long-range layer 2/3A connection weights for the sixteen cells at a single spatial position. Each cell receives two types of connections from other positions: connections from horizontally oriented cells, and connections from vertically oriented cells. Horizontal and vertical lines represent weights from horizontally and vertically oriented cells, respectively. Oblique lines indicate weight from both horizontally and vertically oriented cells. See text for details.

4.1.2 Input presentation

The various inputs used in the simulation are shown in Figure 8c. In order to make the simulation more tractable, the input is presented at or around a constant spatial position and the weights that develop at that spatial position were used at all other spatial positions. This simplification saves a great deal of computational time and is justified by the hypothesis that image statistics are the same across position. Since the weights are learned at a single position the inputs were presented to the network such that the intersection (in case of angles) or the center of the input (in case of colinear line segments) was centered on that position. During each iteration, a random input was chosen and then presented to the network.

4.1.3 Activity-dependent habituative transmitter gates

The simulation shows that, at each position, each cell can get tuned to one of the input features, in our case to a set of angles and colinear lines. As noted above, the challenging aspect of the simulation is that the various inputs share similar features. Hence, if one cell emerges a winner in the first iteration, then it could become the winner for any subsequent iteration. This is because the weights are zero initially and updated at each iteration. Since the inputs share features, the winning cell would have a bias to win again over other cells. Chemical transmitters that habituate in an activity-dependent way overcome this tendency (Grossberg, 1969, 1976b, 1980). These transmitters gate, or multiply, the combined bottom-up and horizontal input before the gated signal can activate the cell. Since the transmitter multiples the input, after it habituates, the gated input to the cell decreases. This enables other cells to emerge as winners during subsequent input presentations and to get tuned to other input features. Recent neurophysiological experiments have confirmed the predicted existence of such habituative gates, or depressing synapses, at cortical cells (Abbott et al., 1997; Markram and Tsodyks, 1996; Tsodyks et al., 1998). Other properties of cortical development have also been shown to depend on habituative thransmitter gates, notably properties of cortical maps, such as ocular dominance and orientation columns that include singularities, fractures, and linear zones; opponent simple cells that are sensitive to opposite contrast polarities; complex cell disparity-tuning properties; and coordinated development of receptive field profiles across the cortical layers (Grossberg and Seitz, 2003; Grunewald and Grossberg, 1998; Olson and Grossberg, 1998). Later sections show how the same habituative gates can also lead to bistable 3D percepts of Necker cube 2D image. This analysis

hereby suggests that bistable percepts may arise from mechanisms that are needed to control cortical development.

4.1.4   Simulation results

The simulation was run until the excitatory and inhibitory weights converged which took approximately 6000 input presentations. Since the horizontal weights are zero initially, random selection of the bottom-up inputs to each neuron ensures that a single neuron will have more bottom-up input than others and hence emerge as a winner through competitive interactions. Once the neuron emerges as a winner, it learns the input by self-organizing the excitatory horizontal connection weights, and inhibitory connection weights, with other layer 2/3A cells. Once the horizontal weights become sufficiently large through learning, they influence the activation of the cell such that the cell that codes the input emerges as a winner. The neuron also habituates once it emerges as a winner so that, in subsequent input presentations, other neurons can win the competition and learn to map other inputs. Since there are at least two neurons for each input, even if one gets habituated, the other can win the competition if the same input is presented for two consecutive input presentations. The results of the simulation are shown in Figure 8d, which shows the horizontal connection weights, from cells at other positions and of different orientations, for each of the sixteen cells. The results show that some neurons get tuned to different angles and some others are tuned to colinear lines.

4.2   Necker cube simulation

In order to clarify how a 2D image can give rise to one or more 3D percepts, a Necker cube 2D image was simulated. The simulation shows how a 2D Necker cube image can be interpreted in 3D and how bistable percepts occur. The simulation was done using the laminar model shown in Figure 2b. There are four different cell types in layer 2/3A of V1 and V2: They are angle cells and colinear bipole cells in V1, and angle cells and disparity-gradient cells in V2. The layer 2/3A cells in V1 input to layer 2/3A cells in V2 and they, in turn, send feedback signals to layer 2/3A of V1. This simplification from the full 3D LAMINART model of Figure 3 was done to ease the computational load. The simulation results should not change if the full 3D LAMINART model is used, since the rate-limiting interactions for the simulated data are captured by the simplified model in Figure 2b.

Each V1 colinear bipole cell is determined by its position, orientation, and disparity. Each V1 angle cell is determined by its position, angle type, and disparity. V2 angle cells are similar to V1 angle cells and are also determined by their position, angle type and disparity. Each V2 disparity-gradient cell is determined by its position, orientation, disparity, and disparity-gradient; see Figure 2b. Four orientations (horizontal, vertical and two obliques), three disparities (D1, D2, and D3), eight different angles corresponding to the eight corners of the Necker cube, and three different disparity-gradients (zero, positive and negative) were used in the simulation. Parameter D1 represents the set of all disparities that correspond to the fixation depth of a planar image, whereas D2 represents a slightly further depth and D3 a still further depth. The model network is similar to the network used in the developmental simulation in Section 4.1. In that simulation, layer 2/3A cells developed into V1 angle cells and colinear bipole cells. Similar rules can be used to develop V2 angle cells and disparity-gradient cells. All these cells are variants of bipole cells. The architectural similarity of different kinds of cells enables the model to be simple but at the same time able to simulate a wide variety of data.

(a)   (b)
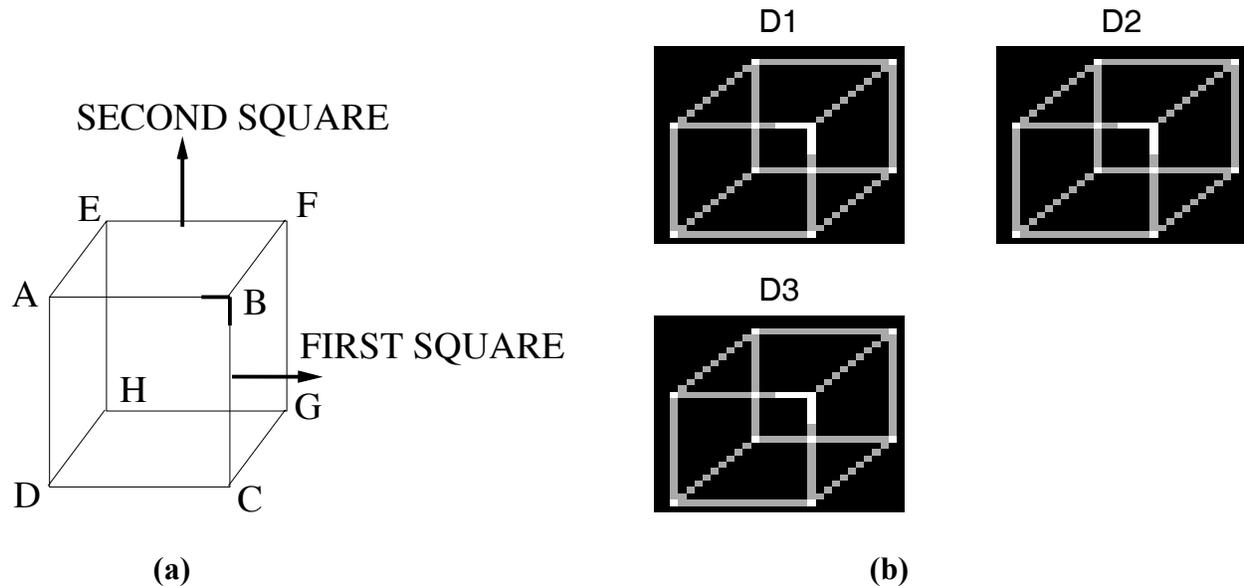
Figure 9. (a) Schematic of input to Necker cube simulation. Three types of input, with highlights at vertex A, B, and H, were used. The terms "first square" and "second square" are used in the text to refer to the two squares. (b) Model input with highlight at vertex B to layer 2/3A cells at multiple depths. This input is generated by hand to be consistent with previous model simulations (Kelly and Grossberg, 2000). D1, D2, and D3 represent different depths. See text for details.

4.2.1   Input generation

As was done for the developmental simulations, the inputs to layer 2/3A of V1 were generated by hand to be consistent with previous model simulation outputs (Grossberg and Mcloughlin, 1997; Kelly and Grossberg, 1999). For the developmental simulations, the inputs were presented at a single disparity, since the simulation focused on interactions within disparity. For the Necker cube simulation, the inputs to layer 2/3A cells were presented at multiple depths (D1, D2, and D3). In all the Necker cube simulations, the input to a single vertex was strengthened by increasing the activation of the corresponding horizontal and vertical line segments within a small neighborhood near the vertex, as illustrated in Figure 9a. This was done to simulate an attentional focus at that particular vertex. Kawabata (1986) showed that the interpretation of the Necker cube depends on which angle is attended. In particular, the Necker cube percept for which the square on which the highlighted angle is present is seen in front is more probable than the other interpretation. The analysis below indicates how attention to any edge fragment, or indeed any momentary enhancement of the boundary corresponding to that edge due to internal changes of state, can yield similar results. Grossberg (1994) described why and how a 2D image can initially activate a population of complex cells that code different disparities. Correspondingly, the Necker cube stimulus generated an input to layer 2/3A complex cells at multiple depths, as shown in Figure 9b. In the present simulations, an attentional bias is represented by a larger input to the attended features.

4.2.2   Activation of layer 2/3 cells in V1

In the model V1, angle cells get activated at the corners and the colinear bipole cells get activated by the straight edges. This segregation of activation occurs by two mechanisms. First, due to the bipole property of these cells, angle cells get maximum input at corners, while the colinear bipole cells get maximum input along straight edges. Second, competition between different angle and colinear bipole cells at the same position and disparity sharpens the responses of the cells to the input. As a result, angle cells are activated at corners while colinear bipole cells are activated along straight edges. The simulation output for the segregation of activation for V1 angle cells and colinear bipole cells is shown in Figure 10.

D1 D2 D1 D2

D3 D3

(a)   **(b)**
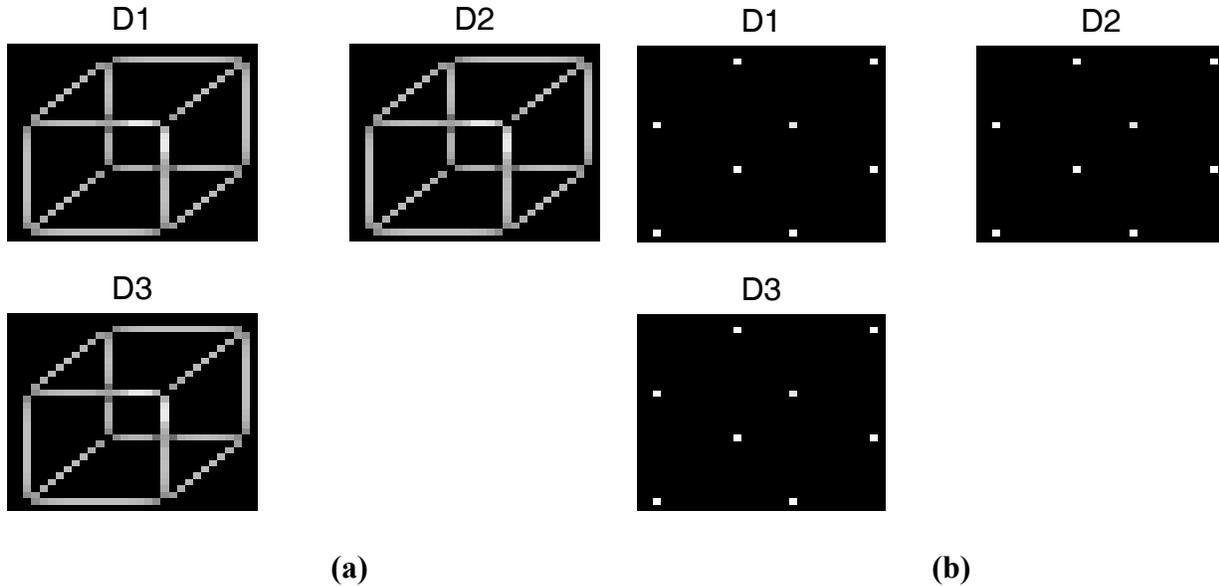
Figure 10. Model output for the activation of V1 colinear bipole and angle cells in layer 2/3A for the Necker cube input. (a) V1 colinear bipole cells, (b) V1 angle cells. There are eleven colinear bipole and angle cells along each edge of the Necker cube. The colinear bipole cells (a) get activated along straight edges while angle cells (b) get activated at corners. D1, D2, and D3 represent different depths.

4.2.3 Activation of layer 2/3A cells in V2
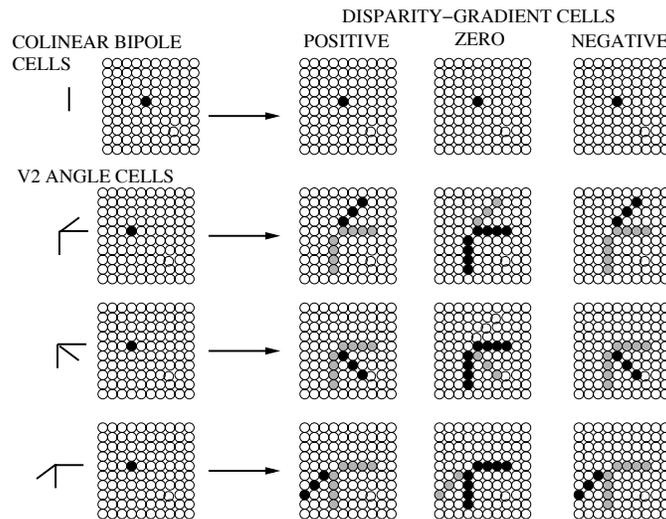
Figure 11. Activation of disparity-gradient cells by colinear bipole cells and V2 angle cells. The colinear bipole cells activate all disparity-gradient cells equally at the corresponding position. Angle cells activate disparity-gradient cells selectively at nearby positions. Shaded circle represent active cell, unshaded circle represents inactive cell. The strength of the activation is indicated by the darkness of the shade. See text for details.

Figure 11 shows how V1 colinear bipole cells and V2 angle cells activate the disparity-gradient cells for some of the corners of the Necker cube. A colinear bipole cell in V1 activates all disparity-gradient cells in V2 equally at their corresponding position, orientation and depth. V2 angle cells strongly activate zero disparity-gradient cells along the horizontal and vertical orientations, and cross disparity-gradient cells along the oblique orientations, at nearby positions. As discussed in Section 3, the selectivity of the angle cells in activating disparity-gradient cells can be learned from 3D image statistics, and is described mathematically in the Appendix for various angles corresponding to the corners of the Necker cube. The activation of disparity-gradient cells by colinear bipole cells and V2 angle cells for the complete Necker cube input is shown in Figure 12a and the combined activation of disparity-gradient cells is shown in Figure 12b.
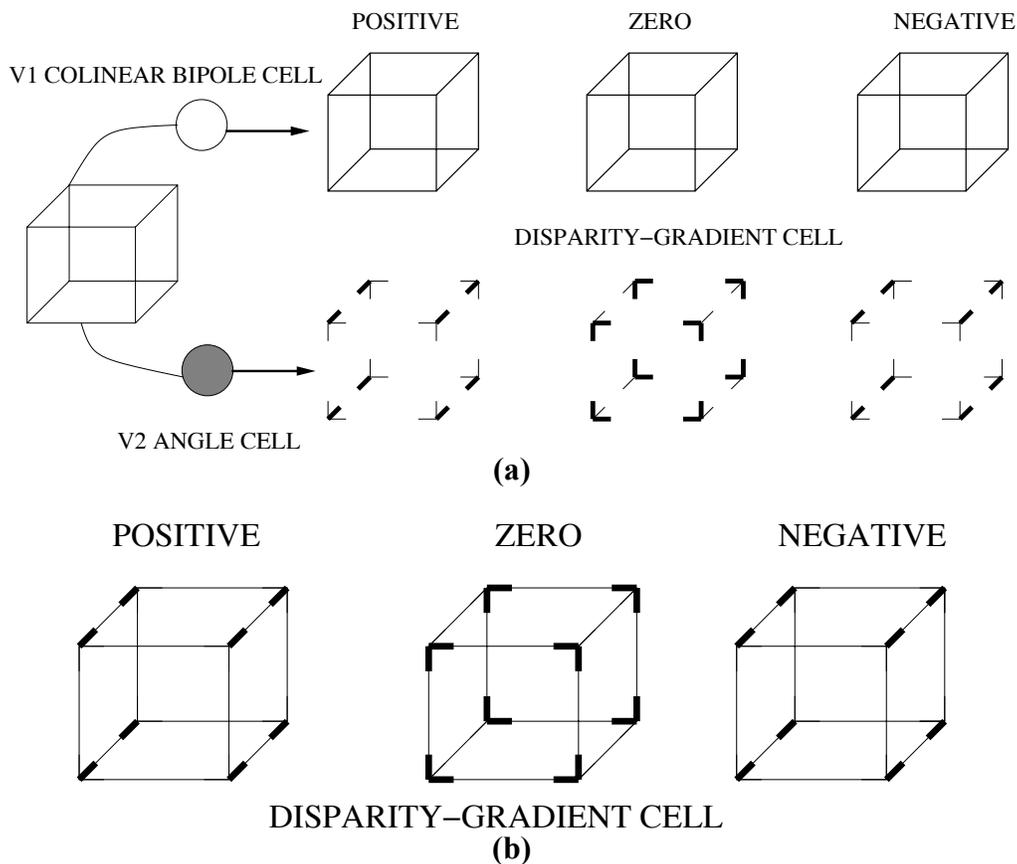
Figure 12. (a) Activation of disparity-gradient cells by angle and colinear bipole cells for the Necker cube input. Colinear bipole cells activate all disparity-gradient cells equally along straight edges, while angle cells selectively activate disparity-gradient cells near corners. Thickness of the lines indicates strength of activation. (b) Combined activation, by angle and colinear bipole cells, of disparity-gradient cells.

V2 angle cells and disparity-gradient cells at the same position, but different disparities, angles, and disparity-gradients, compete with each other. This competition sharpens the response of the cells to the input and helps to disambiguate ambiguous activations.

4.2.4 Grouping and attentional propagation by disparity-gradient cells

Disparity-gradient cells interact with disparity-gradient cells at other positions, disparities, and disparity-gradients, through long-range horizontal connections in layer 2/3A of V2. This anatomical interpretation is consistent with neurophysiological data showing that bipole grouping capable of completing boundaries, including illusory countours, occurs in V2 (Peterhans and von der Heydt, 1999; von der Heydt et al., 1984), and that boundaries in V2 reflect 3D figure-ground properties (von der Heydt et al., 2000; Zhou et al., 2000). These horizontal interactions play a key role in explaining the type of 3D grouping percepts that have been reported; e.g., Tse (1999) and Liinasuo et al. (2000). The grouping principles are a natural extension to the explanation of 3D slanted and curved surfaces of grouping principles that have been used before to explain 2D and 3D planar surface percepts (Gove et al., 1995; Grossberg, 1994, 1997; Grossberg and Howe, 2003; Grossberg and McLoughlin, 1997; Grossberg and Mingolla, 1985a, 1985b; Kelly and Grossberg, 2000).

For the Necker cube input, the zero disparity-gradient cells group along horizontal and vertical arms of the cube while the cross disparity-gradient cells group along the oblique arms. At the same time, attentional highlighting of the angle at vertex B increases the activation of horizontal and vertical arms of the first square (Figure 9a). The increased activation of the horizontal and vertical arms of the first square, as depicted in Figure 13a, is caused by the spread of attention along the boundary of the first square. Roelfsema et al. (1998) showed how attention to one position on a curve can enhance activation of cortical cells at distal positions on the curve. These data are consistent with the hypothesis that attention spreads along the cortical cells that represent the curve. Grossberg and Raizada (2000) simulated the Roelfsema et al. (2000) neurophysiological data, using the 2D LAMINART model, by showing how attention can amplify boundary groupings that form via long-range horizontal connections in layer 2/3A. A similar mechanism is used here to propagate the attentional highlight along the horizontal and vertical arms of the first square. In particular, the activity of the horizontal and vertical disparity-

gradient cells in the model can be modulated by their horizontal and vertical inputs from either side, if they also receive bottom-up input from V1 colinear bipole cells. Thus, the disparity-gradient cells on the highlight translate their increased activation to other disparity-gradient cells along the vertical and horizontal arms, as schematized in Figure 13a and simulated in Figure 13b.

4.2.5   Breaking of T-junctions

The FACADE model proposed how boundaries corresponding to the stem of a T-junction in an image or scene can be split from boundaries corresponding to the top during figure-ground separation (Grossberg, 1994). In particular, long-range grouping combined with short-range competition, across orientations and positions, can break the stem of the T from its top (Figure 13). In the present example, the long-range grouping by disparity-gradient cells increases the activation of boundaries that correspond to the attended square. At the T junctions of the Necker cube, depicted by shaded circles in Figure 13a, the activity of the horizontal and vertical arms of the first square is greater than the corresponding activity of the horizontal and vertical arms of the second square. Spatial and orientational competition among vertical and horizontal disparity-gradient cells enable the first square to inhibit the activity of second square near the T junctions, thereby detaching the stem of the T from its top (Kelly and Grossberg, 1999). The simulation output is shown in Figure 13b.
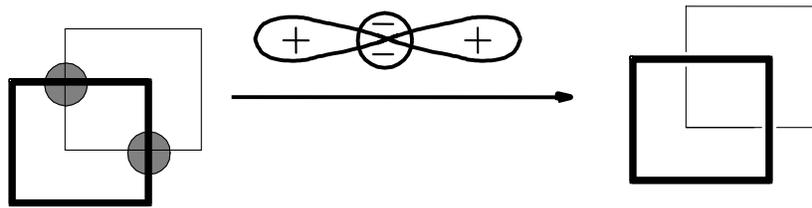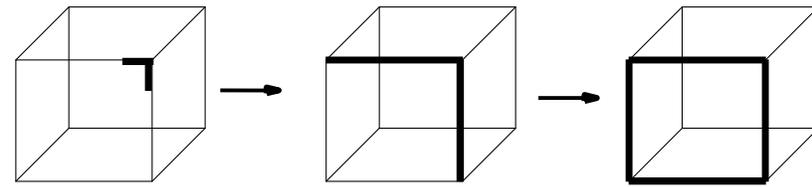
(a)
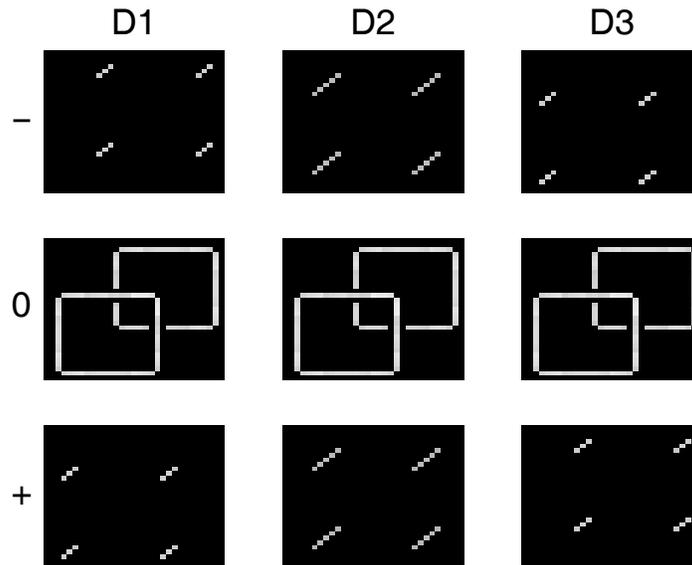
(b)

Figure 13. (a) T junction gap formation: The increased activation of vertical and horizontal arms near vertex B spreads through long-range horizontal connections among disparity-gradient cells along the arms, and thorough V2 angle cells at corners. The increased activation of the first square (Figure 9a) inhibits the activation of vertical and horizontal arms around T junction by spatial and orientational competition across positions and orientations. (b) Corresponding model output. Note that due to T junction break-up, there is a closed boundary that supports filling-in for the first square, but not for the second square.

4.2.6 Filling-in and near-to-far inhibition

As discussed in Section 2.5, filling-in a boundary representation can lead to visible surface percepts. In the FACADE model, boundary signals at multiple depths are used to capture surface signals within depth-selective filling-in domains, or FIDOs. Grossberg (1994) showed that too many boundary and surface fragments are initially formed because each complex cell is sensitive to a range of disparities. In particular, for the Necker cube input, the first square is represented at multiple depths (Figure 9b), which leads to redundant boundary representations. Elimination of the redundant boundaries is realized by the process whereby the boundary and surface properties are bound into a consistent boundary-surface percept. In particular, if a region within a FIDO is surrounded by a closed boundary, then it can contain its filling-in signals. A contour-sensitive output process can sense the bounding contour of this region. These output signals strengthen the boundaries at the same depth and corresponding positions. They also inhibit the redundant boundaries at further depths and the same positions. This inhibition from near to far depths is called *boundary pruning* (Grossberg, 1994). Boundary pruning spares the closest surface representation that successfully fills in a closed boundary at a given set of positions. The process is illustrated in Figure 14.
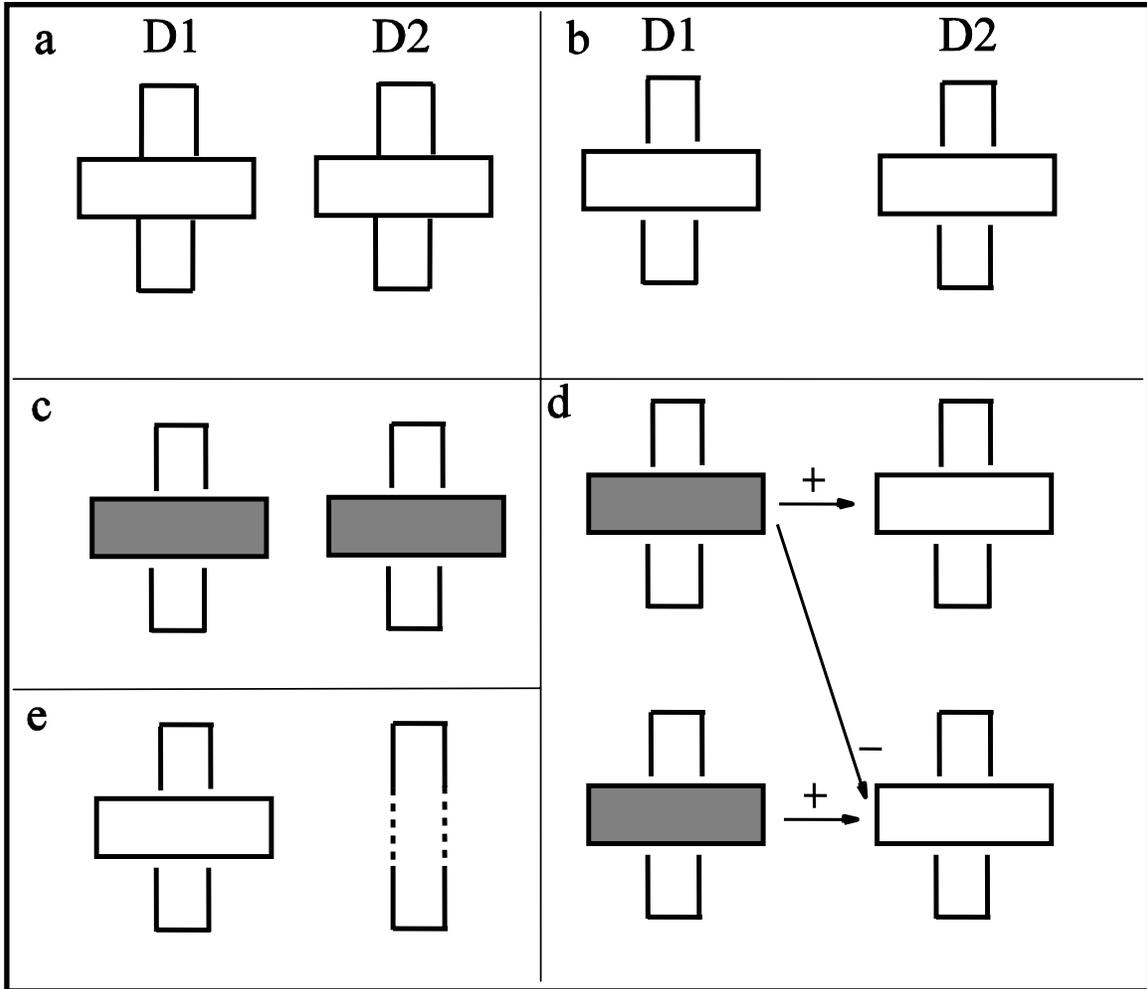
Figure 14. Boundary pruning: The initial multiple depth representation of a 2D image is pruned by filling-in and near-to-far inhibition. (a) Initial representation. (b) T junction breakup (c) Filling-in at depths D1 and D2. (d) The boundaries extracted from the FIDOs enhance the BCS signals within depth while inhibiting the BCS signals at far depths. (e) This near-to-far inhibition prunes the redundant boundaries which allow boundaries at far depth to get completed (dashed lines).

Initially, the 2D image of the two overlapping rectangles is represented at multiple depths as discussed in Section 4.2.1 (Figure 14a). The T-junctions are broken due to grouping and spatial and orientational competition, as described in Section 4.2.5 (Figure 14b). This allows filling-in to occur selectively within the horizontal rectangles at both depths D1 and D2 (Figure 14c). The contour-sensitive signals extracted from the filled-in FIDOs inhibit the boundary signals at further depths (Figure 14d). As a result, the redundant representations of the horizontal rectangle are pruned. The partially occluded vertical rectangle boundaries in Figure 14e can then be

34

amodally complete by bipole grouping (dashed lines). This example illustrates how surface filling-in can influence figure-ground segregation (Albert, 1999).

In the model, all disparity-gradient cells — positive, negative, and zero — within depth act as strong filling-in barriers for the FIDOs at the corresponding depth, and weak filling-in barriers at other depths. This boundary representation, shown in Figure 13, supports filling-in within the first square of the Necker cube, but not within the second square. This is because there is a closed boundary corresponding to the first square but there is no closed boundary for the second square. The redundant boundaries are pruned as follows: In order to ease the computational load, the boundary representation that corresponds to the successfully filled-in square directly inhibits the boundary representations at further depths, as shown in Figure 15.

(a)

(b)
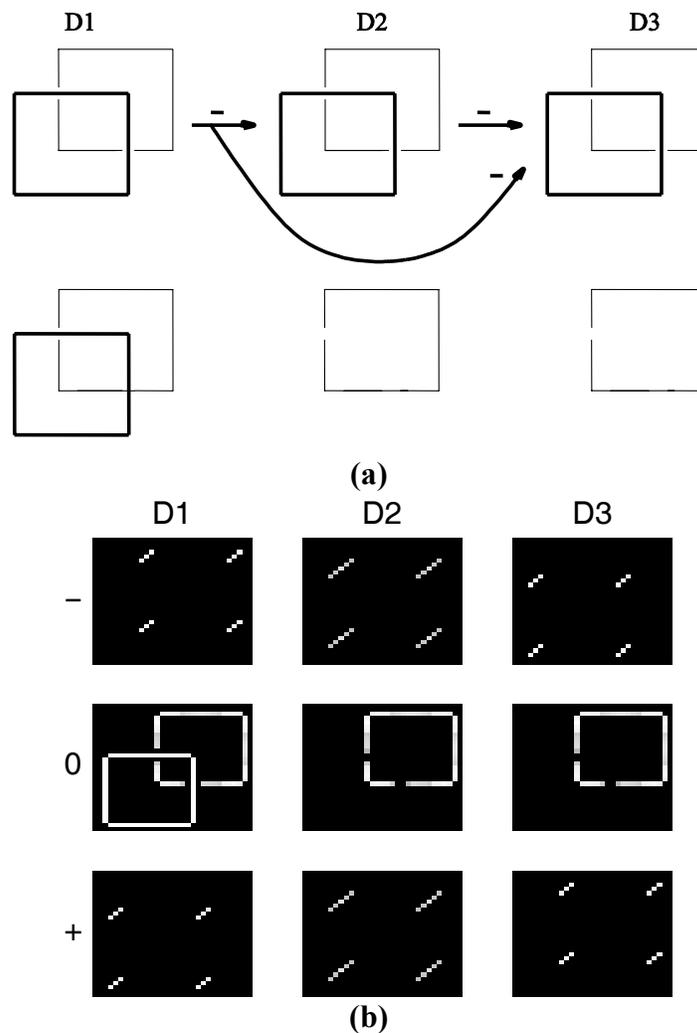
Figure 15. (a) Boundary pruning by near-to-far inhibition inhibits the boundary representation of the first square at depths D2 and D3. (b) Simulated boundaries after near-to-far inhibition.

4.2.7 Disambiguation by grouping

Once the near-to-far inhibition occurs, the first square is represented at depth D1 while the second square is represented at all depths, D1, D2, and D3. This causes an increase in activation of angle cells corresponding to the corners of the first square at depth D1, compared to other depths. This is because the angle cells corresponding to corners of the first square at depth D1 receive both bottom-up input from V1 angle cells, and horizontal input from disparity-gradient cells, whereas the angle cells at depths D2 and D3 get bottom-up input from V1 angle cells but zero horizontal input from disparity-gradient cells. Due to competition between angle cells across depths within position, the angle cells at depth D1 inhibit the corresponding angle cells at other depths. This causes the cross disparity-gradient cells near the corners of the first square to get activated more in depth D1 than at other depths. This is because the disparity-gradient cells at the corners of the first square are not activated by angle cells at depths D2 and D3, but only by colinear bipole cells. In depth D1 though, the disparity-gradient cells are activated by both angle cells and colinear bipole cells. The weak representation of the disparity-gradient cells in D2 and D3 is then inhibited by competition across disparity and within position among disparity-gradient cells. The representation of cross disparity-gradient cell activation is schematically shown in Figure 16a.
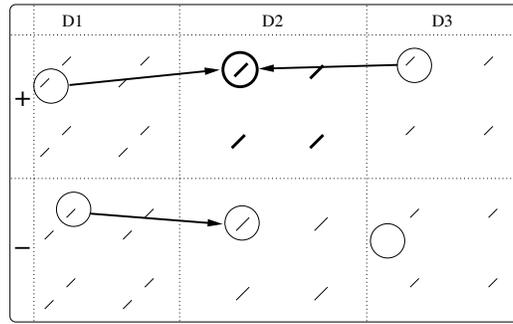
(a)
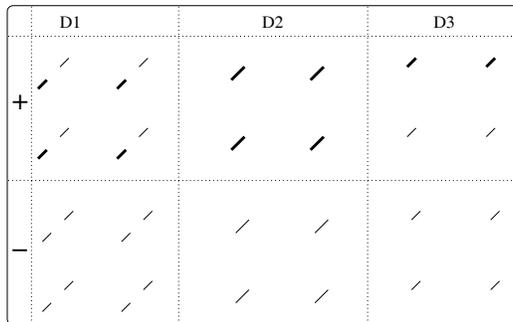
(b)
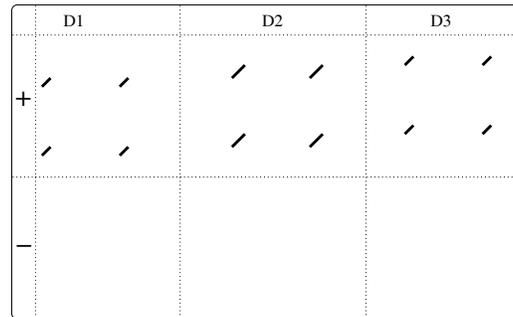
(c)

Figure 16. Positive disparity-gradient cells (+) group more strongly than negative disparity-gradient cells (-) and inhibit negative disparity-gradient cells. (a) Positive disparity-gradient cells at depth D2 get input (arrows in the figure) from other colinear positive disparity-gradient cells at depths D1 and D3. Negative disparity-gradient cells at depth D2 get input only from depth D1. (b) As a result, positive disparity-gradient cells are activated more than negative disparity-gradient cells. (c) Competition across disparity-gradients and disparity inhibits weaker representations. See text for details.

As discussed before (Figure 4), positive disparity-gradient cells group from near to far while negative disparity-gradient cells group from far to near. As shown in Figure 16a, positive disparity-gradient cells at depth D2 get input from other colinear positive disparity-gradient cells from depths D1 and D3 at nearby positions, thus increasing its activation at depth D2, as shown in Figure 16b. Negative disparity-gradient cells get input from other colinear negative disparity-gradient cells only at depth D1, but not D3, and hence the activation of the negative disparity-gradient cell is not increased. This is because negative disparity-gradient cells group from far-to-near and hence receive input from cells at far depths and cells at near depths, as illustrated in Figure 4. The increased activation of positive disparity-gradient cells together with competition across disparities and disparity-gradients leads to inhibition of negative disparity-gradient cells at all depths, as shown in Figure 16c.

(a)

(b)

(c)
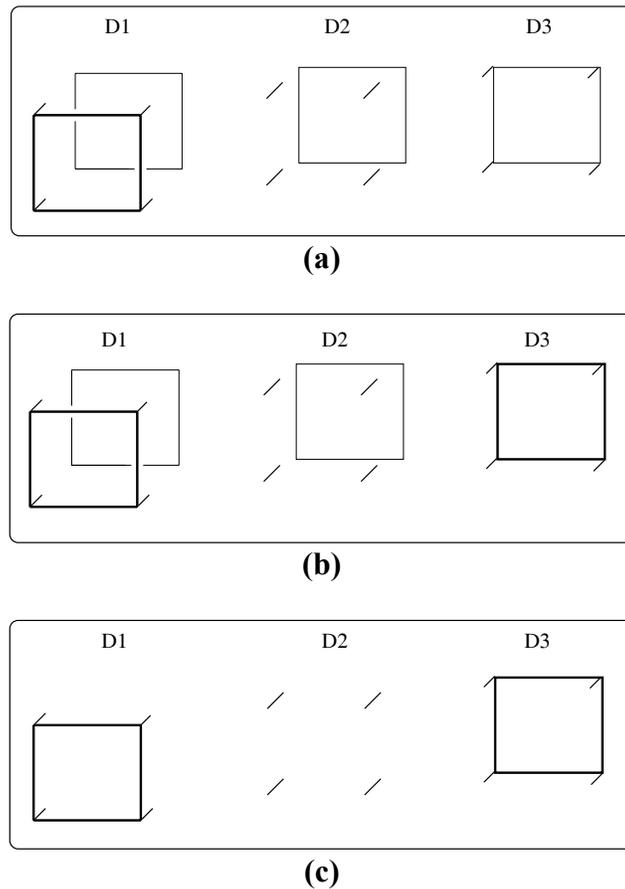
Figure 17. (a) Strong positive disparity-gradient cell activations at depth D1 increases the activation of zero disparity-gradient cells at that depth through V2 angle cells (b). The strong D1 activations enhance the activations of positive disparity-gradient cells, leading to enhanced activation of zero disparity-gradient cells of depth D3. (c) Zero disparity-gradient cells at depth D3 inhibit zero disparity-gradient cells at other depths.

Now the angle cells corresponding to the corners of the second square, at depth D3 receive more input than the corresponding angle cells at depths D1 and D2. This is because, the angle cells at depth D3 receive input from positive disparity-gradient cells and zero disparity-gradient cells, while the ones at depths D1 and D2 receive input only from zero disparity-gradient cells. The increased activation of angle cells lead to increased activation of zero disparity-gradient cells in depth D3, as shown in Figure 17. Again, due to competition between disparity-gradient cells across disparity-gradients and depths, the zero disparity-gradient cells at depth D3 inhibit the zero disparity-gradient cells at depths D1 and D2, thereby leading to a stable representation of the Necker cube. The simulated boundaries of the Necker cube are shown in Figure 18.
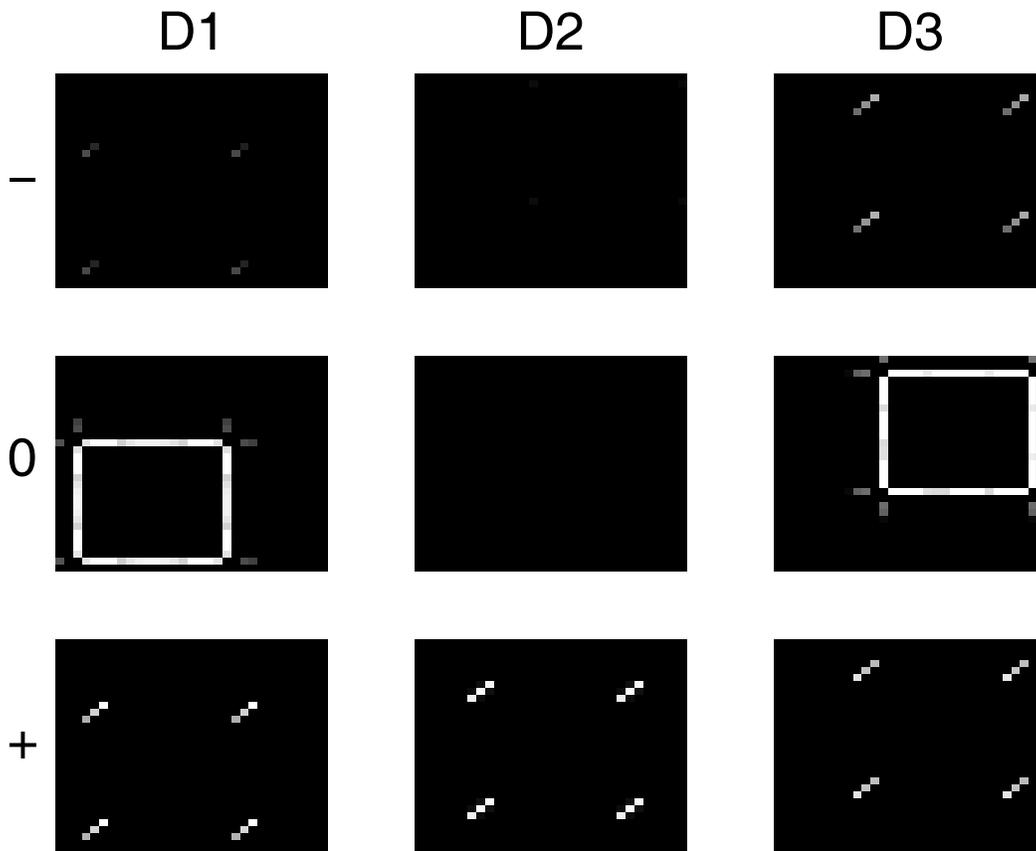
Figure 18. Model simulation of Necker cube boundaries via the process described in Figure 17. The columns D1, D2, and D3 represent different depths. The top row represents negative disparity-gradient cells (-), the middle row represents zero disparity-gradient cells (0), and the bottom row represents positive disparity-gradient cells (+). This notation is also used in subsequent figures.

In summary, the interpretation of the Necker cube involves three main processes. First, angle cells selectively activate disparity-gradient cells locally near the corners. Increased attention to a particular angle of the Necker cube lead to the asymmetric breaking of the X junction and figure-ground separation. Then, cooperative grouping among disparity-gradient cells and competition between disparity-gradient cells across disparities and disparity-gradients lead to a final boundary representation of the Necker cube.

4.3   Necker cube simulation with different attentional focus

The present simulation shows how attention to a different angle of the Necker cube can bias the interpretation of the Necker cube to form a different 3D percept. The difference in where the attentional highlight is present leads to a different figure-ground segregation. Since an angle of the second square is highlighted (angle H in Figure 9a; see Figure 19a), the zero disparity-gradient cells that represent vertical and horizontal arms of the second square are activated more than those of the first square. This causes the same cascade of events to occur for the second square that was summarized in Figures 13 and 15 for the first square. In particular, attention propagates along the boundary of the second square and leads to inhibition of the vertical and horizontal arms of the first square near the T junctions. Near-to-far inhibition leads to a representation where the second square is represented in depth D1. This representation increases the activity of negative disparity-gradient cells, as opposed to positive disparity-gradient cells in the previous simulation, thereby leading to a different 3D boundary representation of the Necker cube, as shown in Figure 19. All parameters and settings for the simulation remained the same except for the slight attentional bias in the input.
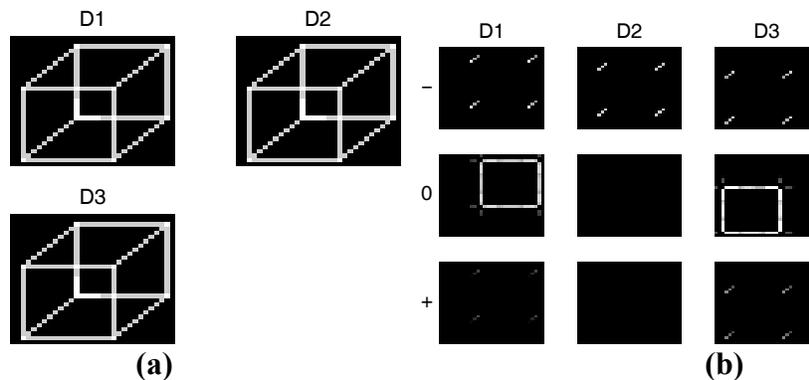
Figure 19. (a) Input with attentional highlight at vertex H (see Figure 9). (b) Model simulation of boundary representation. Note that attentional highlight at a different vertex from Figure 18 leads to a different interpretation of the Necker cube.

4.4   Necker cube simulation with different length

The depth at which the second square gets represented in the previous simulations is dependent on the length of the oblique lines of the Necker cube. This section shows how the cross disparity-gradient cell groups for a smaller Necker cube. The input to the network, shown in Figure 20a, is similar to the one used in the previous simulations except that it is smaller in size. This forces the grouping of disparity-gradient cells to group between depths D1 and D2, thereby representing the second square at depth D2 as shown in Figure 20b, instead of D3, as shown in Figure 18 for a larger Necker cube.

(a)   (b)
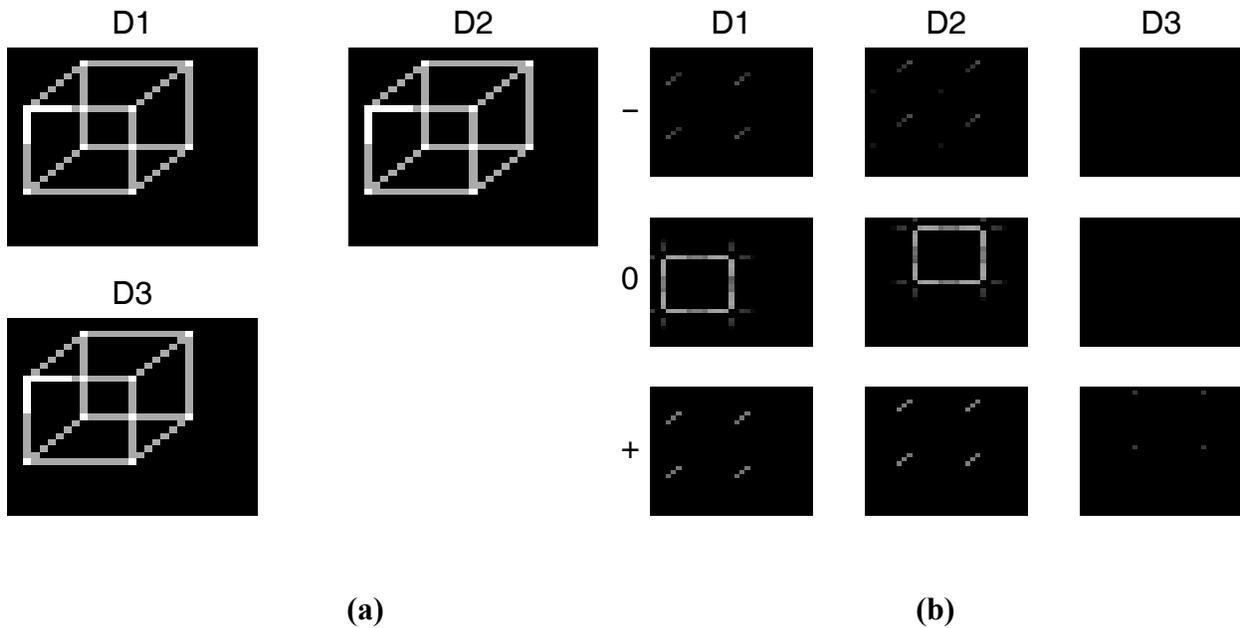
Figure 20. (a) Small Necker cube input with highlight at vertex A. (b) Model simulation of boundaries. Here the further cube surface is closer than in Figure 18.

4.5   Necker cube reversals and cortical development

The previous simulations showed how the Necker cube image generates a stable 3D boundary representation when attention is focused on a single angle at a vertex, or for that matter, when any fluctuation in boundary strength favors one representation over the other. The Necker cube percept is, however, bistable and its interpretation switch over time (Heath, Ehrlich and Orbach, 1963; Kawabata, 1986; Maier, Wilke, Leopold, Treue and Logothetis, 2001). Section 4.1 noted

that habituative transmitters are needed to develop horizontal connections in layer 2/3A of V1 into angle and colinear bipole cells, and by extension, angle and disparity-gradient cells in V2. Such habituative transmitters in V2, when they interact with the rest of the network, can also generate bistable percepts. In particular, the model switches its boundary representations through time when habituative transmitters gate the excitatory and inhibitory inputs to layer 2/3A cells of V2. This result links development to perception by showing that habituative mechanisms necessary for development can also explain bistable percepts.

Layer 2/3A cells of V2 receive long-range horizontal inputs from other layer 2/3A cells in V2 and bottom-up input from layer 2/3A of V1. Layer 2/3A cells also receive inhibitory inputs from other layer 2/3A cells of V2. Habituative transmitters are proposed to gate both the excitatory and inhibitory inputs to layer 2/3A cells. For simplicity, one transmitter is proposed to gate all excitatory inputs, both bottom-up and horizontal, and another to gate all inhibitory inputs to a cell.

Figure 21 shows the Necker cube representation of the network at different time steps. Figure 21a shows the initial interpretation of the Necker cube. This is similar to the previous simulation output discussed before, which shows how the network behaves when its transmitters are fully accumulated. After some time, the winning disparity-gradient cells habituate. In particular, the disparity-gradient cells corresponding to the first square at depth D1 and the disparity-gradient cells corresponding to the second square at depth D3 habituate. The second square get activated at depths D1 and D2 as the activation of, and inhibition from, the disparity-gradient cells corresponding to the second square at depth D3 decreases due to habituation. At the same time, the first square get activated at depths D2 and D3 as the disparity-gradient cells corresponding to the first square at depth D1 habituate. Since the activation of disparity-gradient cells corresponding to the second square is more than the activation of disparity-gradient cells corresponding to the first square at depth D1, the T-junctions are broken. Now, the network state is similar to that in Section 4.3. The model boundary representation now supports the second interpretation, as shown in Figure 21b, which is similar to the output of the model in Figure 19. After some time, the network cycles back to the initial interpretation as habituation and recovery proceed, as shown in Figure 21c. Thus, even though habituation operates locally at individual cells, the model switches between globally consistent interpretations. This is because the various cells in the network are coupled in a context-sensitive way by both intralaminar and interlaminar

feedback. Such a stable oscillation *that reproduces the correct 3D grouping* cannot be taken for granted in a distributed, hierarchical, multiple time-scale system like the present one. In our hands, all of the model mechanisms, interacting together, were needed to generate this basic result.
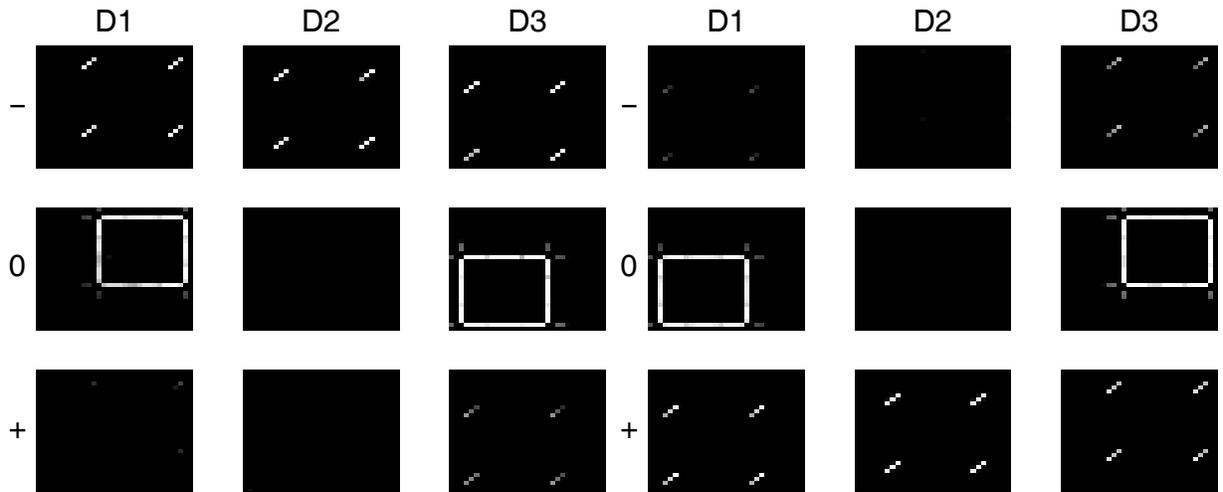
(a)   (b)
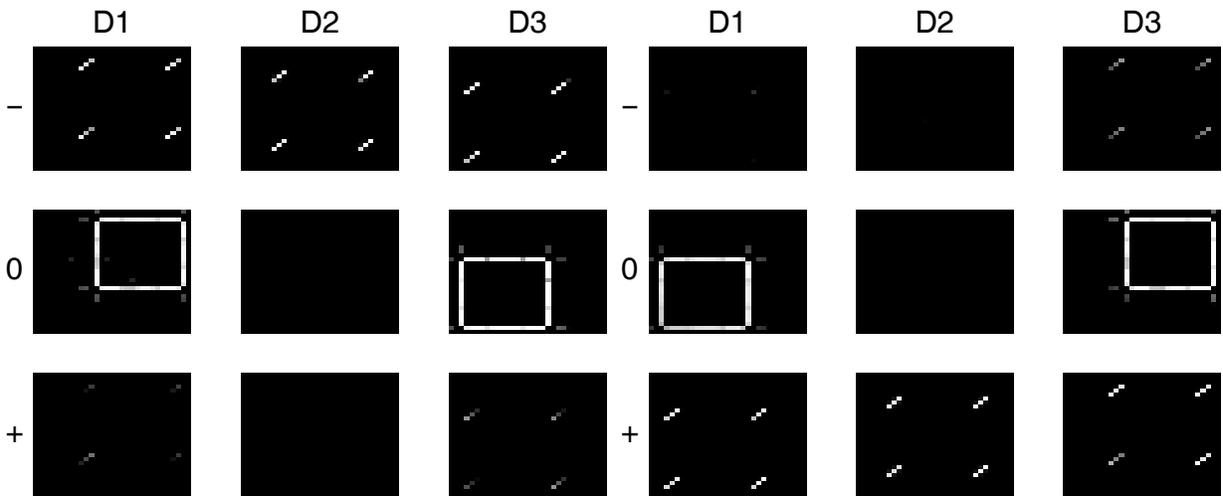
(c)   (d)

Figure 21. Necker cube reversal simulation. Model simulation of boundary representation at different times: (a) Time step 15. (b) Time step 49. (c) Time step 77. (d) Time step 100. This simulation shows that the model can cycle between the perceived 3D representations.

4.6   Model simulations of slant representation without monocular cues

A slanted rectangle, when viewed in 3D, is represented at multiple depths by matching its binocular disparities. Recent modeling has shown how such binocular matches can be carried out in layers 3B and 2/3A of V1 (Grossberg and Howe, 2002). These V1 cell responses can be grouping by V2 disparity-gradient cells to code the slant of a 3D object. This section describes simulations that illustrate this property. It is also shown how disparity-gradient cells can complete groupings across depth, as during percepts of 3D neon color displays.

Since these simulations focus on interactions between disparity-gradient cells, angle cells in both V1 and V2 are not used. Colinear bipole cells directly activate disparity-gradient cells. Three different inputs to layer 2/3A cells of V1 were generated corresponding to (1) a flat surface, (2) a slanted surface, and (3) a slanted neon surface, as shown in Figure 22. The flat surface is represented by activation within depth, the slanted surface is represented by activation across depths (Figure 5), and the slanted neon surface is represented similarly to the slanted surface case but with zero activation in depth D2. These inputs were generated to show how the disparity-gradient cells group and complete boundaries across depths, and are consistent with previous model simulations of 3D boundary and surface perception; e.g., Grossberg and Howe (2002) and Grossberg and McLoughlin (1997). The slanted surface is represented as a parallelogram made of vertical and oblique line segments (Figure 5).

4.6.1   Simulation results

The colinear bipole cells get activated by the representations at various depths for the inputs discussed before. As discussed before, colinear bipole cells at a particular depth activate all disparity-gradient cells ─ zero, positive, and negative ─ equally at their corresponding depths. Even though the initial activation of all disparity-gradient cells at corresponding depths are equal, the long-range horizontal connections between disparity-gradient cells together with short-range inhibition across depths and disparity-gradients lead to the correct 3D boundary representation.
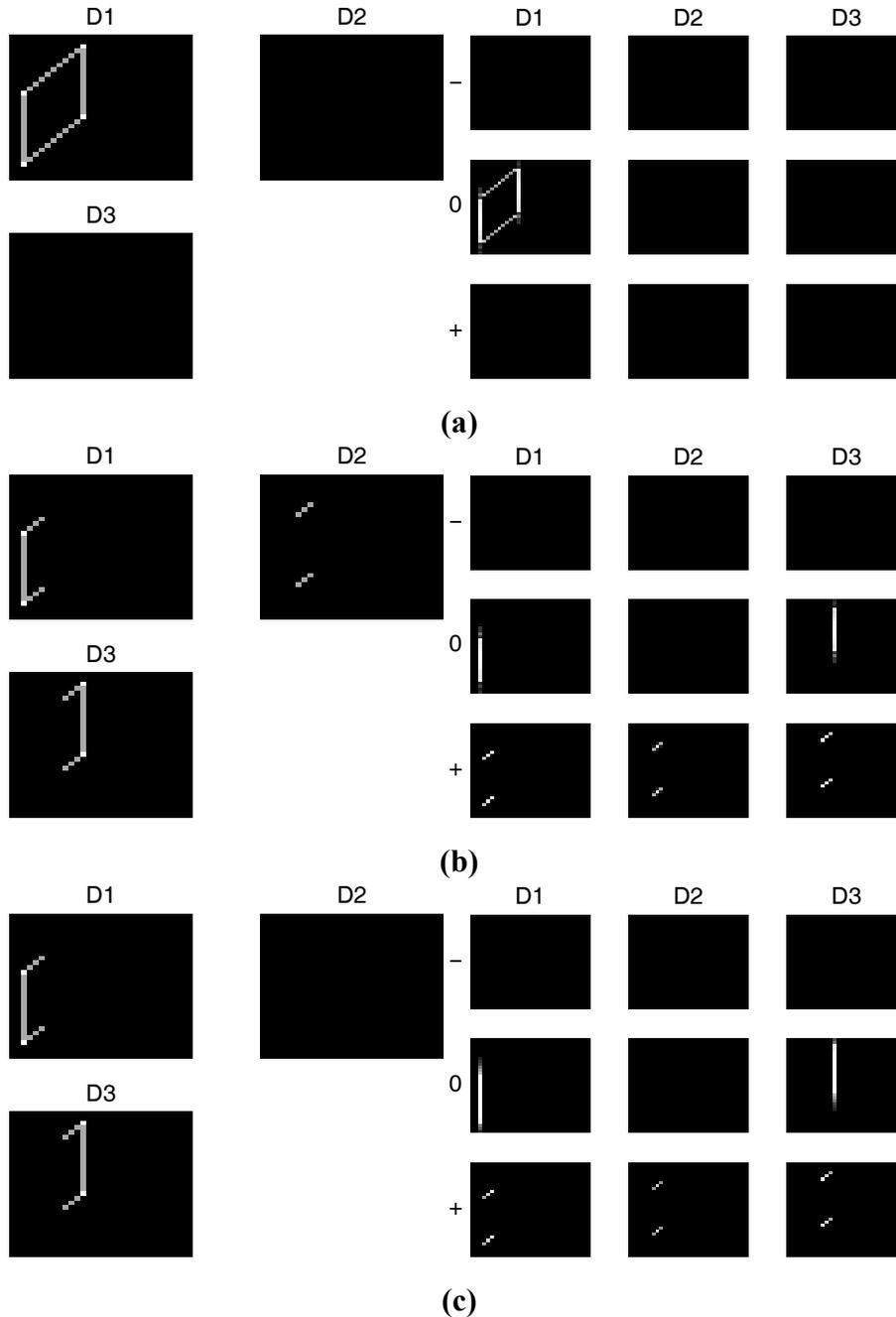
Figure 22. 3D input simulation. (a) Input and output for the at surface. Note that the output is represented by zero disparity-gradient cells within depth. (b) Input and output for the slanted surface. Note that the output is represented by zero disparity-gradient and positive disparity-gradient cells indicating the slant of the input. (c) Input and output for the slanted neon surface. Even though there is no input in depth D2, the disparity-gradient cells complete correctly at depth D2. Note the similarity between this output and the output for the slanted surface.

For the flat surface (Figure 22a), the input is represented within a single depth. The colinear bipole cells activate all disparity-gradient cells at that depth equally. Since zero disparity-

gradient cells group strongly within depth, their activation increases, and competition between disparity-gradient cells across disparity-gradients and depths, eliminate the activity of cross disparity-gradient cells. The final output, shown in Figure 22a, is represented by zero disparity-gradient cells within depth.

For the slanted surface (Figure 22b), the input is represented at multiple depths. The colinear bipole cells activate all disparity-gradient cells at the corresponding depths. Since the input is present at multiple depths, cross disparity-gradient cells group more strongly than zero disparity-gradient cells. In particular, positive disparity-gradient cells group strongly than negative or zero disparity-gradient cells. This is because positive disparity-gradient cells group from near to far, and hence cells at depth D2 get horizontal input from depth D1 and D3, whereas negative disparity-gradient cells at depth D2 do not get any horizontal input (Figure 16). Again, positive disparity-gradient cells inhibit negative and zero disparity-gradient cells through competition between disparity-gradient cells across depths. The final representation of the input by disparity-gradient cells is shown in Figure 22b.

The slanted neon surface input (Figure 22c) is similar to the slanted surface input except that there is no activation at depth D2. This representation schematically models a slanted neon surface. The output of the simulation is shown in Figure 22c. The simulation shows that the positive disparity-gradient cells correctly complete the boundary at appropriate depths, yielding a boundary like that for the slanted surface (Figure 22b).

4.6.2   Filling-in of slanted and neon surfaces

This section proposes how surface filling-in of a slanted surface occurs. As discussed in Section 2.5, a slanted or curved boundary could have gaps at certain depths through which filling-in signals may escape. In the model, the boundary signals at their preferred depth act as a strong barrier to filling-in and weaker barrier at other depths (Figure 5). Closed boundaries are hereby generated within depth and filling-in can be contained, at least partially, within depth.

So far, the model simulations have focused on boundary representations. Boundary representations need to be complemented with surface filling-in to generate visible percepts (Grossberg, 1994). Grossberg and Todorović (1988) showed that a first step in computing the relative lightness of two surfaces, while discounting the illuminant, can be achieved by cells that interact thorough a center-surround symmetric receptive fields. Two kinds of cells, on-center off-

surround (ON), and off-center on-surround (OFF), that respond to increases or decreases in intensity, respectively, are used in the model (Kandel, Schwartz and Jessell, 2000). The excitatory and inhibitory components are balanced so that the cell responses are attenuated to spatially uniform stimulation and the cell, therefore, respond preferentially to lightness borders.

The ON and OFF cells generate ON and OFF filling-in signals, as shown in Figure 23a, in response to the input image shown in the figure. These filling-in signals are then used to fill-in the ON and OFF FIDOs (Section 4.2.6). The relative lightness of a surface is represented by the balance of activation of the ON and OFF FIDOs. That is, a darker color is represented by strong activation of the OFF FIDO and weak activation of the ON FIDO. At the same time, lighter color is represented by strong activation of the ON FIDO and weak activation of the OFF FIDO. Hence, the relative strength of the ON and OFF FIDOs indicates the color of the surface, and is computed as $[ON-OFF]^+$ in model simulations.

The simulation uses the disparity-gradient boundaries shown in Figure 22 as filling-in barriers. The ON and OFF filling-in signals are not depth-sensitive and hence project to all depths. The boundary signals at each depth act as a barrier to the filling-in signals and restrict the filling-in to create visible surfaces.

The simulation output for the flat surface, along with the corresponding boundary signals, is shown in Figure 23a. The filling-in is strong at depth D1 while it weak or non-existent at depths D2 and D3. This is because the boundary signals at depth D1 project strongly to the filling-in domains at depth D1 to restrict the filling-in. At the same time, there are no boundary signals at depths D2 and D3. Only the weak projection from depth D1 acts as a barrier for the filling-in signals at depths D2 and D3, and hence filling-in at that depth is weak. The filled-in representation for the flat surface is uniform at depth D1. The slant of the surface at a position is indicated by the relative strength of the filled-in signals across depths. For example, for the flat surface, the strength of the filled-in signal is greater at depth D1 throughout the surface to represent the flatness of the surface.

The simulation output for the slanted surface and the corresponding boundary signals, are shown in Figure 23b. Here, the boundary signals are present at all depths, but at different positions. This set of filling-in barriers support weak filling-in at all depths and the slant of the surface is represented by the relative amount of filling-in across depths. The filled-in signal at depth D1 is stronger in the initial part of the surface, while the filled-in signal is stronger at depth

D3 for the final part of the surface. In the middle, the filled-in signal is equal at all depths. The slant of the surface is represented by the relative strength of the filled-in signal across depths.

A neon slanted surface can be filled-in as follows. First, the boundaries need to get completed across depths. This is achieved by grouping among disparity-gradient cells in our model. Second, the lightness inducers are present only along certain boundary segments. For filling-in of planar surfaces, even if inducers are present only along certain parts of the boundary, the filling-in signal from those parts can fill-in the whole region bounded by the boundary (Gove, Grossberg and Mingolla, 1995). But when the surface is curved, the boundaries along which the inducers are present may be at a different depth than the boundaries along which there are no inducers. Hence, the filling-in signal needs to spread across depths. In the model, the filling-in signal projects to all depths, and since a strong boundary at a given depth also acts as a weak barrier at other depths, the filling-in signals are at least partially contained at all depths. The ON and OFF filling-in signals and the resultant filled-in surface output are shown in Figure 23c. The model output is similar to that in the slanted surface case even though the boundaries and filling-in signals are not present at certain depths along the surface.

A measure of depth can be obtained from the activity of the filled-in regions. In particular, Figure 23 shows the filled-in activity at three depths, D1, D2 and D3. The depth value at a particular point can be calculated as $D = \dfrac{D1 * A_{d1} + D2 * A_{d2} + D3 * A_{d3}}{A_{d1} + A_{d2} + A_{d3}}$, where D1, D2, and D3 are the relative depth values, and A_{d1}, A_{d2} and A_{d3} are the activities of the filled regions at the respective depths. The depth value, D, which is between D1 and D3 can then indicate the relative depth of the surface (see Figure 23d).
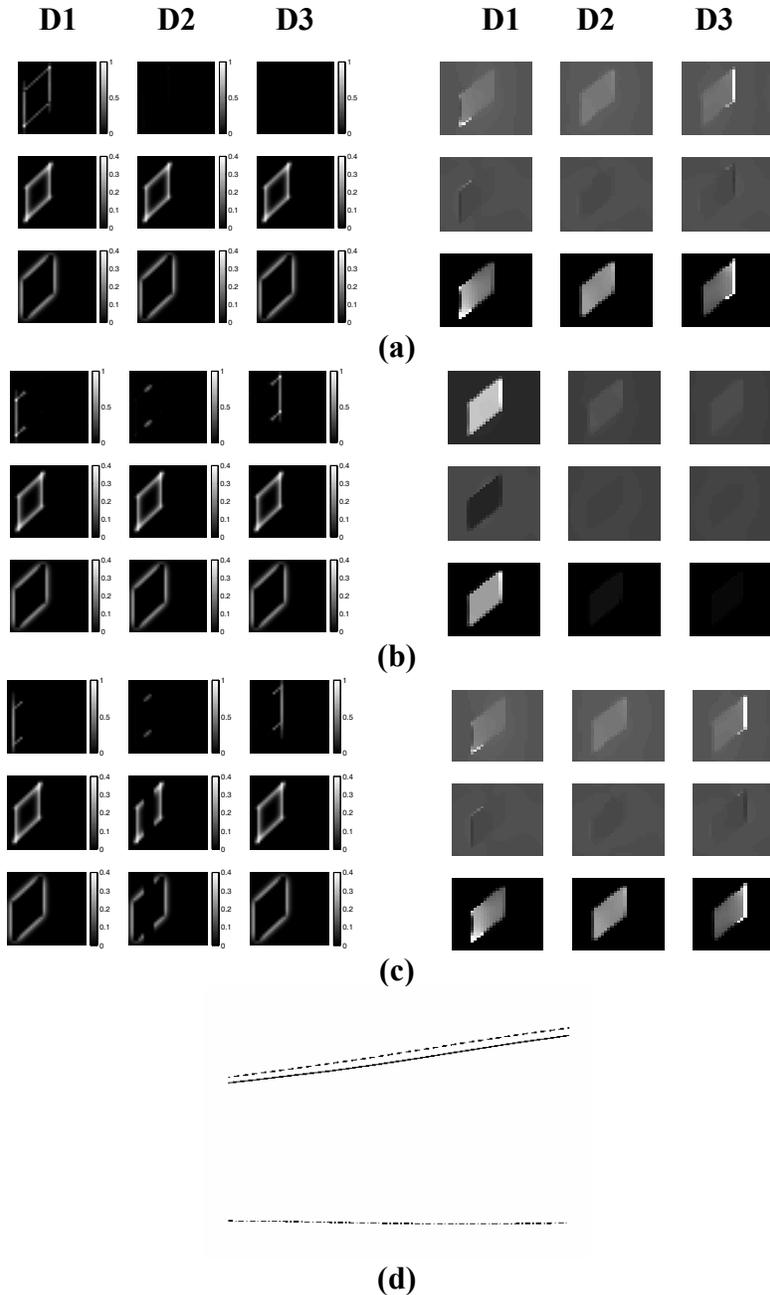
Figure 23. 3D filling-in simulations. D1, D2, and D3 represent different depths. (a) Flat surface, (b) Slanted surface, and (c) Slanted neon surface. The left part of the figures shows the boundary signals (B) for filling-in, and the ON (+) and OFF (-) filling-in signals. The right part shows the ON FIDO output (+), OFF FIDO output (-), and [ON -OFF]$^+$ output (R). For the at surface the [ON -OFF]$^+$ output is maximal at depth D1, while it is weak or non-existent at other depths. For the slanted surface, the [ON -OFF]$^+$ output is greater in the initial part of the rectangle at depth D1, greater in the middle part at depth D2, and greater in the final part at depth D3. The output for the slanted neon surface is similar to the slanted surface result. (d) The depth values calculated from the filled-in values for flat (-.), slanted (--) and slanted neon (-) surfaces. The depth value for the flat surface is constant while the depth value for the slanted surface is increasing indicating the positive slant of the surface. See text for details.

49

5   Discussion

5.1   3D grouping

This article extends 2D boundary grouping and surface formation principles to explain data about 3D slanted and curved surface percepts. 3D contour interpolation has been shown to obey similar constraints to those governing 2D grouping (Garrigan and Kellman, 2002), such as relatability (Kellman and Shipley, 1991). Our extension of 2D grouping principles to 3D clarifies the neural mechanisms that create these perceptual constraints, and provides the first mathematical description of the kernels that can be used for 3D grouping, including an explanation of how illusory contours can form across depth. This is accomplished using disparity-gradient cells that are interpreted to occur in layer 2/3A of V2. A number of studies have shown that cells in V2 code complex properties, such as figure-ground sensitive boundary coding (Zhou et al., 2000), tuning to stereo edges (von der Heydt et al., 2000), and tuning to illusory contours that is dependent on the depth of the inducers (Bakin et al., 2000). Recently, cells having similar properties to disparity-gradient cells were found in area V4 of the Macaque (Hinkle and Connor, 2001). We predict that cells having such properties may be found as early as area V2. We have proposed anatomical locations for the model cells that are needed to explain our targeted data based on the best available neurological data, and to be consistent with other theoretical constraints on models of 3D vision and figure-ground perception (Grossberg, 1994, 1997; Kelly and Grossberg, 2000). This anatomical interpretation can be directly tested. What is critical for model properties, however, is not a particular anatomical interpretation, but rather a certain ordering of model processes.

For example, using the model's ordering of processes enables its 3D grouping cells to explain percepts like the one shown in Figure 1b (Tse, 1999). In this figure, even though the lines of the cube are colinear in the 2D image, they do not complete behind the occluder. This is explained in the model as follows. When the 3D interpretation is taken into account, the oblique lines of the cube on the left are coded by positive disparity-gradient cells while the oblique lines of the cube on the right are coded by negative disparity-gradient cells. The positive and negative disparity-gradient cells do not group in the model, and hence the lines are not completed behind the occluder.

SECOND SQUARE

FIRST SQUARE

A

(a)   (b)
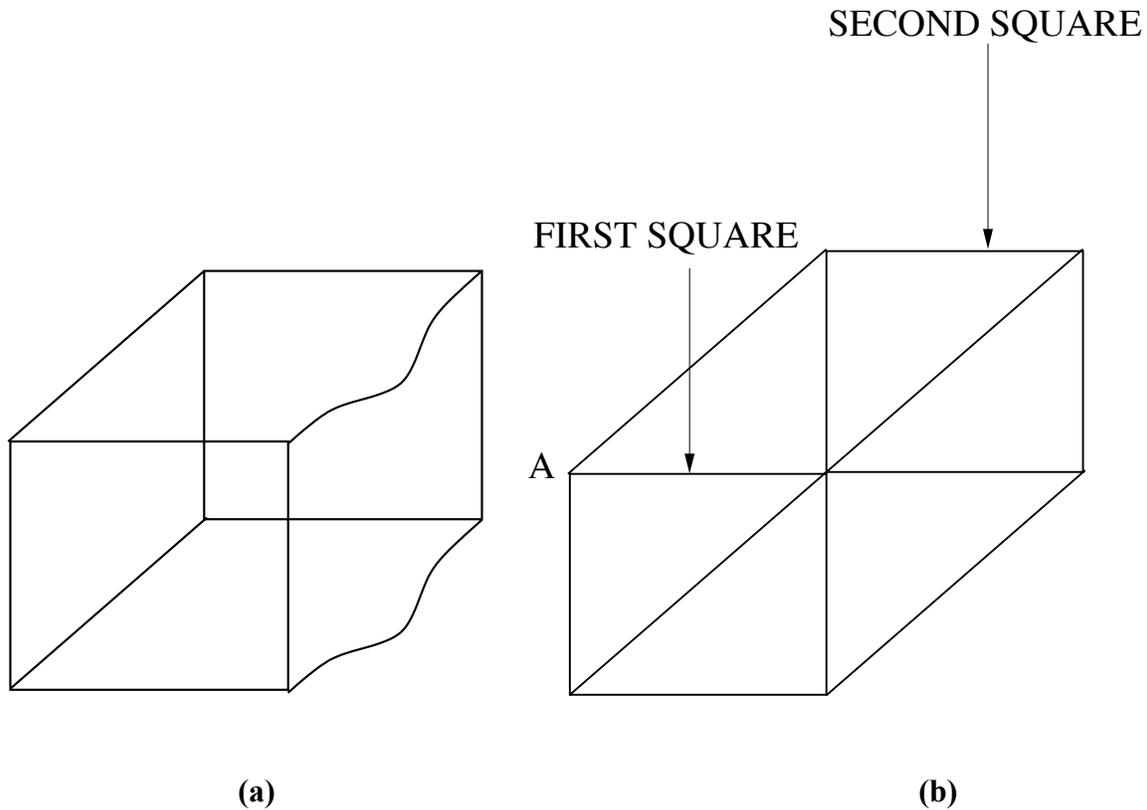
Figure 24. (a) Necker cube with curved lines. (b) Necker cube in which the two squares are brought together.

5.2   Slant aftereffects, habituation, and development

Disparity-gradient cells can also code the slant of an object or image. As noted in Section 1, the aftereffect experiments done in Ryan and Gillam (1993) showed that the slant aftereffects are mediated by cells that code slant. Lee (1999) showed that the size of the aftereffect is dependent on the disparity difference between the adapting and test surface. These results can be explained by hypothesizing cells tuned to positive and negative disparity-gradients that are also tuned to disparity. The disparity-gradient cells in the model are tuned to different disparity-gradients and are also tuned to various depths, as in Lee (1999). The model also illustrates how habituative mechanisms are needed for the development of layer 2/3A cells in V1 into angle and colinear bipole cells, and how such habituative mechanisms can also lead to bistable percepts. When habituative mechanisms interact with competing disparity-gradient cells, slant aftereffect data can also be explained. In particular, let us assume that a slanted surface with positive slant is shown as the adapting stimuli. The slanted surface will then be represented by the positive disparity-gradient cells in the model. When a test surface with zero slant is shown, the zero and

negative disparity-gradient cells are activated more than the positive disparity-gradient cells. This is because the positive disparity-gradient cells have habituated in response to the adapting stimuli. Hence, the balance of activation shifts toward negative slant, albeit slightly. Since the disparity-gradient cells in the model are also tuned to disparity, the size of aftereffect is dependent on the disparity difference between the adapt and test stimuli, as in the data of Lee (1999).

5.3   Other models of 3D interpretation of 2D images

The 3D LAMINART model embodies a detailed neural explanation of how cells that are used for 3D grouping can also be used to disambiguate ambiguous interpretations of 2D images. There are two classes of models that deal with 3D interpretation of 2D images. In the first class (Sugihara, 1986; Waltz, 1972), the edges in the 2D image are labeled as either being convex, concave, or occluded, depending on the angles that they subtend. This initial assignment of labels, and the further enforcement of a globally consistent labeling scheme by constraint satisfaction, gives rise to a stable 3D interpretation. Thus, the output of the model is a labeling of the edges in the 2D image. Even if the edges are consistently labeled, however, they still need to be represented in depth. For example, if the length of the oblique lines of a Necker cube image is increased, then, even though the angles are the same, the interpretation is different because the relative depth between the first and the second square varies. Hence, a model needs to explain where the various parts of the 2D image are in depth, and to generate a 3D surface representation that matches the human percept.

The second class of models takes a minimization approach to deal with the interpretation of a 2D image (Leclerc and Fischler, 1992; Marill, 1991). In these models, the vertices of the 2D image are moved in the 3D space such that a measure, such as standard deviation of the angles, is minimized subject to certain constraints such as planarity of the surfaces, and the compactness of the surface (Sinha and Poggio, 1996). In this approach, the whole image is taken into account and hence it is unable to deal with local variations. For example, consider a variation of the Necker cube image shown in Figure 24a. The figure shows a Necker cube-like image that has similar local information near the corners as that of the Necker cube, but is different in how these corners are connected. They are connected by curved lines rather by straight lines. The models that use a global minimization approach cannot deal with such local variations. In order to

explain the above figure, a local representation of the oblique lines in the image is needed. In the 3D LAMINART model, the interpretation is built up locally by activating disparity-gradient cells through angle cells, and global constraints emerge through grouping among disparity-gradient cells. Because, the lines are curved, the disparity-gradient cells group in such a way so as to represent the initial part of the curve by positive disparity-gradient cells and the middle part by zero disparity-gradient cells and the final part by negative disparity-gradient cells, thus indicating the change in slant of the image.

Figure 24b shows a Necker cube in which the two squares are brought together in a single vertex. This image leads to a 2D planar interpretation rather than the 3D interpretation, even though the local characteristics near the vertices are similar to those in the usual Necker cube image. Usually, this interpretation is explained by the non-accidental viewpoint hypothesis (Witkin and Tanenbaum, 1983); namely, if the 2D image is interpreted as a cube, then slight variations, such as rotation of the cube, lead to a dramatically different 2D projection. Interpreting the image as a 2D planar image leads to a more stable representation. This explanation does not, however, explain what a 3D representation is in the brain, or how a 3D perturbation of this representation that never takes place can influence it. The 3D LAMINART model explains the 2D percept as follows. Assume that there is an attentional focus at vertex A and that the first square is represented at depth D1 initially. Note that one of the corners of the first square is shared by the second square; namely, the upper right corner of the first and the lower left corner of the second squares. When the first square is represented at depth D1, then that shared corner is also represented at depth D1. This is true because the angle cells activate mostly zero disparity-gradient cells along the vertical and horizontal lines. For the same reason, other corners of the second square are also represented at depth D1. This causes the oblique lines to get grouped within depth and hence the figure is represented within depth. This explanation is similar to the model explanation of why a triangle is seen as a 2D planar image whereas a parallelogram can be seen in 3D.

5.4   Interaction of monocular and binocular cues

The 3D LAMINART model clarifies how 3D cues, such as disparity, and 2D cues, such as angles, can interact when they are present in the same image. Psychophysical data for the combination of 3D cues and 2D cues support a *weak fusion model* (Landy et al., 1995), which

argues that cues that are more reliable are weighted more than the cues that are less reliable. For example, at near distances, disparity is more reliable than at far distances. Hence, disparity information is weighted more at near distances than at far distances. On the other hand, monocular cues are more reliable for computing slant information than disparity cues (Stevens et al., 1991) and hence are weighted more than disparity in such cases. In cases where both cues are equally strong, information from both cues is averaged. In the 3D LAMINART MODEL, disparity-gradient cells in V2 are activated by both V1 colinear bipole cells and V2 angle cells. The 3D cue information comes from the disparity-sensitive, primarily vertically oriented, V1 colinear bipole cells, whereas the 2D cue information comes from V2 angle cells. These separate paths interact to give rise to a stable 3D percept. If the 2D monocular cues are stronger, they will activate the disparity-gradient cells more and hence the interpretation would be more dependent on 2D cues. If disparity information is stronger, then the disparity-gradient cells would group according to disparity and the interpretation would be consistent with disparity information. The existence of separate paths for disparity and perspective information is also supported by the results in van Ee et al. (2002). They show that when conflicting disparity and perspective information is presented, subjects see the percept specified either by disparity or by perspective, but not both. They also show that the percept is bistable, so that subjects alternately see the percept dominated by either type of cue. As discussed in the previous section, the mechanism responsible for bistability for Necker cube images in our model can also explain the bistability in displays where perspective and disparity information are presented in conflict.

APPENDIX

This section describes the model equations. First described are the equations that were used to simulate the development of colinear and non-colinear bipole cells in layer 2/3A. Then the equations for the Necker cube and 3D simulations are provided. Each model neuron is typically modeled as a single voltage compartment in which the membrane potential, $v(t)$, is given by

$$C_m \frac{dv(t)}{dt} = (E_{\text{leak}} - v(t))g_{\text{leak}} + (E_{\text{excite}} - v(t))g_{\text{excite}}(t) + (E_{\text{inhib}} - v(t))g_{\text{inhib}}(t), \tag{1}$$

where E represent reversal potentials, g_{leak} is a constant leakage conductance, and the time-varying conductances $g_{\text{excite}}(t)$ and $g_{\text{inhib}}(t)$ represent the total inputs to the cell (Grossberg, 1973; Hodgkin, 1964). The following network equations are instances of this general membrane equation, where, for simplicity, the capacitance term C_m was set equal to 1, the reversal potentials are set to: $E_{\text{excite}} = 1$, $E_{\text{inhib}} = -1$, and $E_{\text{leak}} = 0$, except where indicated. Then equation (1) can be rewritten in the form,

$$\frac{d}{dt}v = -\alpha v + (1-v)g_{\text{excite}} - (1+v)g_{\text{inhib}}, \tag{2}$$

where α is a constant decay rate, g_{excite} is the total excitatory input, and g_{inhib} is the total inhibitory input.

The differential equations were implemented in Matlab and numerically integrated using an adaptive step size Runge-Kutta 4,5 method. For computational simplicity, the equations for learning of the adaptive weights were solved at a slower time scale using Euler's method.

The developmental simulations describe only layer 2/3A of V1. The Necker cube simulations describe the dynamics of layers 2/3A of V1 and V2, as in Figure 2b.

A Development of colinear and non-colinear cells in layer 2/3A of V1

A.1 Activation Equations. The following equations were used to simulate how the horizontal connections in layer 2/3A of V1 develop into two types of cells: colinear bipole cells that link colinear line segments over short distances and non-colinear bipole cells that get tuned to angles in the image. The model self-organizes both the longer-range excitatory connections and shorter-

range inhibitory connections that are needed to realize cell selectivity to these inputs; cf., Grossberg and Williamson (2001). The layer 2/3A excitatory neurons are modeled as follows:

$$\frac{d}{dt}z_{ijkl} = -\alpha z_{ijkl} + (1 - z_{ijkl})h_{ijkl}(\gamma_1[z_{ijkl} - \tau_1]^+ + [H_{ijkl}^E + \gamma_2 I_{ijkl} - H_{ijkl}^I]^+)$$
$$- (1 + z_{ijkl})\sum_{r \neq k, t \neq l}[z_{ijrt} - \tau_2]^+,$$

(3)

where variable z_{ijkl} is the activity of the excitatory layer 2/3A cell at position (i, j), orientation k, and cell number l. Two orientations, horizontal and vertical, were used in the simulation. There are eight excitatory cells (indicated by subscript l) associated with each spatial position (i, j) and orientation k. Since two orientations, horizontal and vertical, are used in the simulation, there are sixteen excitatory cells present at each position. These sixteen cells code the various colinear and non-colinear bipole cells that will develop there.

Parameter α in (3) is the decay constant and h_{ijkl} indicates the level of habituative transmitter (see equation (7)) associated with the excitatory cell at position (i, j), orientation k and cell number l. Habituative transmitters prevent the earliest cells that learn from persistently dominating network dynamics. Term $\gamma_1[z_{ijkl} - \tau_1]^+$ in (3) describes self-excitatory feedback to the cell, where notation $[x]^+$ describes a threshold, or half-wave rectification, operation such that $[x]^+ = x$, if $x > 0$, 0 otherwise. Parameter γ_1 is the feedback gain constant. Feedback increases the activity of the winning neuron, which in turn inhibits other weakly activated cells, thereby enabling a winning cell to be selected in each cluster.

The excitatory input H_{ijkl}^E in (3) is due to long-range connections from neighboring cells in (3). It is defined by:

$$H_{ijkl}^E = \sum_{pqr}W_{pqrijkl}\ e^{-\frac{(i-p)^2 + (j-q)^2}{\sigma^2}}\frac{1}{N}\sum_t[z_{pqrt} - \tau_3]^+.$$

(4)

The horizontal connection weight is the product of a spatial Gaussian $e^{-\frac{(i-p)^2 + (j-q)^2}{\sigma^2}}$ that reflects axonal growth (Grossberg and Williamson, 2001) and an adaptive connection weight $W_{pqrijkl}$ from a cell at position (p, q) and orientation r to a cell at position (i, j), orientation k, and cell number l. Each cell receives the total input $\sum_t[z_{pqrt} - \tau_3]^+$ summed over cell number, from each

neighboring position and orientation; see Figure 8a. The total input was normalized by dividing by the total number of cells (N). Other scaling parameters work just as well. Variable I_{ijkl} in (3) is the bottom-up input, generated by hand to be consistent with previous model simulation outputs (Grossberg and Williamson, 2001; Grossberg and Howe, 2003, Grossberg and Raizada, 2000), and shown in Figure 8c. In vivo, a horizontally oriented layer 2/3A cell receives input from horizontally oriented layer 3B cells, and a vertically oriented layer 2/3A cell receives input from vertically oriented layer 3B cells (Callaway and Wiser, 1996). In the simulations, inputs are delivered directly to layer 2/3A cells. All the eight cells at a particular position and orientation receive the same bottom-up input to which small random input perturbation, in the range of 0.002 to -0.002 using the *rand* function in Matlab, was added to simulate randomness in initial cortical connections. This random perturbation is generated for each of the eight cells before the simulation and is fixed for all iterations. The bottom-up input is added to the input from the long-range horizontal connections. As a result, the activity of a layer 2/3A cell that receives bottom-up input can be modulated by the input from the long-range connections even if there is input from only one side.

Input H_{ijkl}^{I} in (3) from the inhibitory interneurons is defined by:

$$H_{ijkl}^{I} = \sum_{rfg} B_{ijrfgkl}^{+} [s_{ijrfg}]^{+},$$

(5)

where variable s_{ijrfg} is the activity of the gth inhibitory interneuron (see below) associated with the excitatory neuron at position (i, j), orientation r, and cell number f, and $B_{ijrfgkl}^{+}$ is the weight from this inhibitory interneuron to an excitatory neuron at the same position, but different orientation k, and cell number l. This inhibition balances the excitation H_{ijkl}^{E} from the long-range connections in (3) to implement the bipole property. In particular, cells which receive no bottom-up input and signals from only one side of the horizontal receptive field are not activated enough to exceed the inhibitory input and thus are not able to propagate the grouping signal any further. Cells that receive sufficiently strong horizontal excitation from both sides, however, may exceed the inhibitory input and thereby fire. Along with the inhibitory input from the interneurons, layer 2/3A excitatory cells also receive inhibition $\sum_{r \neq k, t \neq l} [z_{ijrt} - \tau_2]^{+}$ in (3) across all the cells that represent a given position (i, j). This competition enables a winning cell to be selected in each cell cluster.

As described in Section 2.2, each excitatory neuron is associated with a pool of inhibitory interneurons. The inhibition from the interneurons to the excitatory cell and the inhibition among inhibitory interneurons helps to implement the bipole property. In the developmental simulations, each excitatory cell is associated with four inhibitory interneurons corresponding to the left, right, top and bottom side of the excitatory cell (see Figure 8b). Each interneuron receives the same long-range input that is received by the excitatory cell from its corresponding side. In the Necker cube simulations below, depending on the type of cell (angle, colinear, disparity-gradient), each excitatory cell is associated with two or three inhibitory interneurons since it is assumed that these selective horizontal long-range connections have already developed.

A.2 Habituative transmitter gates. In equation (3), the total input,

$$J_{ijkl} = \gamma_1[z_{ijkl} - \tau_1]^+ + [H^E_{ijkl} + \gamma_2 I_{ijkl} - H^I_{ijkl}]^+,$$ (6)

to each layer 2/3A excitatory cell is multiplicatively gated by a habituative, or depressing, transmitter h_{ijkl} that obeys the equation:

$$\frac{1}{\delta}\frac{d}{dt}h_{ijkl} = \left(A_h + \frac{1}{1 + \theta[J_{ijkl} - \varepsilon]^+}\right)(1 - h_{ijkl}) - B_h h_{ijkl}[J_{ijkl} - \varepsilon]^+$$ (7)

(Abbott et al., 1997; Grossberg 1969, 1976b, 1980; Tsodyks et al., 1998). In (7), transmitter starts out at its maximal value 1. Term $\left(A_h + \frac{1}{1 + \theta[J_{ijkl} - \varepsilon]^+}\right)(1 - h_{ijkl})$ describes the accumulation, or recovery, of the transmitter to its maximum value 1 at the variable rate $\left(A_h + \frac{1}{1 + \theta[J_{ijkl} - \varepsilon]^+}\right)$, and term $-B_h h_{ijkl}[J_{ijkl} - \varepsilon]^+$ describes transmitter habituation at the variable rate $B_h[J_{ijkl} - \varepsilon]^+$. Other examples of rate-dependent recovery from habituation during vision are proposed to include motion perception (Grossberg and Rudd, 1992) and photoreceptor adaptation (Carpenter and Grossberg, 1981). The recovery proceeds at a slower rate when the total input, J_{ijkl}, is above the threshold, but proceeds at a faster rate when the input is below threshold ε. This property helps the neuron to recover faster when it receives less total input and does not take part in the representation of the input. In particular, suppose a neuron wins the

competition and habituates for a particular input. When another input then enables a different cell to win the competition, the previously habituated cell can recover faster and thus can recover fully before the input that the neuron codes is presented once again. In the Necker cube simulations below, this property enables the neurons that habituate when one interpretation of the cube is represented to recover fully before the neurons that represent the other interpretation habituate, thus enabling the network to switch back to the initial interpretation. Parameter B_h in (7) governs the rate of habituation. Habituation in (7) occurs at a rate proportional to the rate of release, or inactivation, of transmitter, h_{ijkl}, in (3).

The activity, s_{ijklm}, of the inhibitory interneurons is modeled as follows:

$$\frac{d}{dt} s_{ijklm} = -s_{ijklm} + H_{ijkl}^{Em} - s_{ijklm} \sum_{rfg} B_{ijrfgklm}^{-} \left[s_{ijrfg} \right]^{+}. \tag{8}$$

Each inhibitory interneuron in (8) receives part, H_{ijkl}^{Em}, of the total long-range input, H_{ijkl}^{E} in (4), to the excitatory layer 2/3A cell at its position. Since there are four inhibitory interneurons for each excitatory cell in the developmental simulations, each interneuron receives horizontal input from one of the four sides (L=left, R=right, T=top, B=bottom) of the excitatory cell; namely,

$$H_{ijkl}^{EL} = \sum_{3\frac{\pi}{4} < \theta(p-i,q-j) < 5\frac{\pi}{4},r} W_{pqrijkl} \; e^{-\frac{(i-p)^2 + (j-q)^2}{\sigma}} \frac{1}{N} \sum_{t} \left[z_{pqrt} - \tau_3 \right]^{+}, \tag{9}$$

$$H_{ijkl}^{ER} = \sum_{7\frac{\pi}{4} < \theta(p-i,q-j) < \frac{\pi}{4},r} W_{pqrijkl} \; e^{-\frac{(i-p)^2 + (j-q)^2}{\sigma}} \frac{1}{N} \sum_{t} \left[z_{pqrt} - \tau_3 \right]^{+}, \tag{10}$$

$$H_{ijkl}^{ET} = \sum_{\frac{\pi}{4} < \theta(p-i,q-j) < 3\frac{\pi}{4},r} W_{pqrijkl} \; e^{-\frac{(i-p)^2 + (j-q)^2}{\sigma}} \frac{1}{N} \sum_{t} \left[z_{pqrt} - \tau_3 \right]^{+}, \tag{11}$$

and

$$H_{ijkl}^{EB} = \sum_{5\frac{\pi}{4} < \theta(p-i,q-j) < 7\frac{\pi}{4},r} W_{pqrijkl} \; e^{-\frac{(i-p)^2 + (j-q)^2}{\sigma}} \frac{1}{N} \sum_{t} \left[z_{pqrt} - \tau_3 \right]^{+}, \tag{12}$$

where $\theta(x,y) = \tan(y/x)$. The inhibitory interneurons inhibit each other in (8) via the term $\sum_{efg} B_{ijrfgklm}^{-} \left[s_{ijrfg} \right]^{+}$. $B_{ijrfgklm}^{-}$ is the inhibitory weight from interneuron number g, at position (i,j),

orientation *r*, and cell number *f*, to an interneuron number *m*, at the same position, but orientation *k*, and cell number *l*. This recurrent inhibition among the interneurons normalizes the total inhibition that is received by the excitatory cell and helps to realize the bipole property, as described in Section 2.2.

A.3 Learning of adaptive weights. The adaptive weights, $W_{pqrijkl}$ in (3) and (9)-(12), for the layer 2/3A long-range connections were modified through learning using the equations:

$$\frac{d}{dt}W_{pqrijkl} = [z_{ijkl} - \tau_4]^+ \left(\frac{1}{N}\sum_t [z_{pqrt} - \tau_4] - W_{pqrijkl} \right). \tag{13}$$

As in the simulations of horizontal cell development in Grossberg and Williamson (2001), equation (13) is an *instar* learning law, which has become the standard law for learning self-organizing maps (Grossberg, 1976a, 1980; Kohonen, 1989). During instar learning, the activity, z_{ijkl}, in the postsynaptic target cell turns on learning, and the adaptive weight, $W_{pqrijkl}$, learns the expected value of the total signal $\frac{1}{N}\sum_t [z_{pqrt} - \tau_4]^+$ from its presynaptic source cells during the interval when the target cell is active. In order to ease the computational load, the adaptive weights for cells at a single position were learned and these weight values were used for cells in other spatial positions as well. This simplification is justified by the hypothesis that image statistics are the same across position.

An *outstar* learning law (Grossberg, 1968, 1980) was used to learn the weights, $B^+_{ijrfgkl}$ in (5) between an inhibitory interneuron and its excitatory neuron, and the weights, $B^-_{ijrfgklm}$ in (8), among the inhibitory interneurons. The weights, $B^+_{ijrfgkl}$ and $B^-_{ijrfgklm}$, have only two spatial indices (*i, j*) because they are short-range interactions whose spatial extent is limited to a single hypercolumn that is indexed by position (*i, j*). The learning law for $B^+_{ijrfgkl}$ is:

$$\frac{d}{dt}B^+_{ijrfgkl} = [s_{ijrfg}]^+ \left(H^{Eg}_{ijkl} - B^+_{ijrfgkl} \right). \tag{14}$$

It helps to create and maintain the balance between inhibition and excitation that is needed to realize the bipole grouping property. Outstar learning accomplishes this by causing the inhibitory

synaptic weight $B^+_{ijrfgkl}$ to track the expected activation, H^{Eg}_{ijkl}, of the target excitatory cells at times when its source inhibitory interneuron has positive activity $[s_{ijrfg}]^+$. Likewise, the inhibitory interneuronal weights $B^-_{ijrfgklm}$ obey the outstar equation:

$$\frac{d}{dt}B^-_{ijrfgklm} = [s_{ijrfg}]^+\left([s_{ijklm}]^+ - B^-_{ijrfgklm}\right) .$$

(15)

The weight $B^-_{ijrfgklm}$ tracks the positive activity $[s_{ijklm}]^+$ of the target inhibitory interneuron. This property enables the network to normalize the total inhibitory input H^I_{ijkl} in (3) from the interneurons, which enables layer 2/3A excitatory cells to fire if there is excitatory input on both sides sufficient to overcome this normalized inhibition.

Both the long-range horizontal excitatory connections $W_{pqrijkl}$ and short-range inhibitory connections $B^+_{ijrfgkl}$ and $B^-_{ijrfgklm}$ develop from zero initial values in the model; that is, $W(0) = B^+(0) = B^-(0) = 0$. The parameter values are $\alpha = 0.5$, $\gamma_1 = 5$, $\gamma_2 = 0.1$, $\varepsilon = 0.1$, $\sigma = 7$, $\tau_1 = 0.35$, $\tau_2 = 0.15$, $\tau_3 = 0.03$, $\tau_4 = 0.2$, $\delta = 0.1$, $\theta = 1000$, $A_h = 0.1$, and $B_h = 1.0$. The simulation results are shown in Figure 8d. Each subplot in the figure depicts the weights for each of the sixteen cells at a single spatial position. In particular, each subplot shows the connection weights from cells at position (p, q) within a 11x11 neighborhood. The first eight subplots ($l=1$, …8) in the first two rows show the connection weights from vertically oriented cells, $W_{pq0ijkl}$, to cells at position (i,j), orientation k, and cell number l. These cells receive long-range input from the neighboring horizontally and vertically oriented cells. The weight from horizontally oriented cells is indicated by the horizontal lines, while the weight from vertically oriented cells is indicated by vertical lines. Weights from both horizontally and vertically oriented cells are indicated by oblique lines. The length of the line indicates the strength of the connection. The second eight subplots in the last two rows shows the connection weights from horizontally oriented cells, $W_{pq1ijkl}$.

B Dynamics of 3D Necker cube Bistability

The Necker cube simulations were done using the laminar circuit shown in Figure 2b with the activity equations describing interactions within and between layers 2/3A of V1 and V2. As

described in Section 4.2, V1 layer 2/3A of the model contains colinear and non-colinear bipole cells (angle cells), while V2 layer 2/3A contains disparity-gradient and angle cells. Below, V1 colinear bipole cells are indicated by the letter c, V1 non-colinear bipole cells by letter n, V2 angle cells by letter a, and disparity-gradient cells by letter g. Letters (i,j) and (p,q) indicate 2D positions, d and e disparities, k and r orientations, m and o disparity-gradients, and l and t angle cell types. Inhibitory interneurons are indicated by letter s, and the inhibitory interneuron number by letters u and v. In the developmental simulations, there are four inhibitory interneurons (equation (8)) for each side (top, left, right, and bottom) of the excitatory cell. Here, we use only two (left and right for horizontally oriented cell or top and bottom for vertically oriented cell) or three (for angle cell) inhibitory interneurons, since it is assumed that these selective horizontal long-range connections have already developed. Four orientations [vertical $(k = 1)$, $45°$ oblique $(k = 2)$, horizontal $(k = 3)$, and $135°$ oblique $(k = 4)$] eight angles corresponding to the eight corners of the Necker cube $(l = 1,...,8)$, three disparities $(d = D1, D2, D3)$, and three disparity-gradients [positive $(m = 0)$, zero $(m = 1)$, and negative $(m = 2)$] were used in the simulations. Both excitatory and inhibitory habituative transmitters were used in layer 2/3A of V2 for the Necker cube simulations. The inhibitory habituation helps the neurons of the alternate interpretation of the Necker cube to get activated before the activity of neurons representing the present interpretation goes below threshold. This property helps the network to make the switch to various interpretations faster. Habituation was not used in layer 2/3A of V1 as there are no competing interactions between which the network can switch in layer 2/3A of V1 and hence the network would reach an equilibrium state and stay in that state even if habituation was used.

B.1 V1 colinear bipole cell activation equations. Activity c_{ijkd} obeys the equation:

$$\frac{d}{dt}c_{ijkd} = -\alpha_c c_{ijkd}$$

$$+ \left(1 - c_{ijkd}\right)\left(\gamma_1\left[\sum_v H^{Ec}_{ijkdv} + \gamma_2 I_{ijkd} - H^{Ic}_{ijkd}\right]^+ + \gamma_3\left[c_{ijkd} - \beta_c\right]^+ + \gamma_3\sum_o\left[g_{ijkdo} - \beta_g\right]^+\right) \quad (16)$$

$$- \left(1 + c_{ijkd}\right)\left(C^P_{ijkd} + C^S_{ijkd}\right) \ .$$

In (16) α_c is the decay rate. Term H^{Ec}_{ijkdv} describes excitatory input from the long-range connections in layer 2/3A of V1; namely:
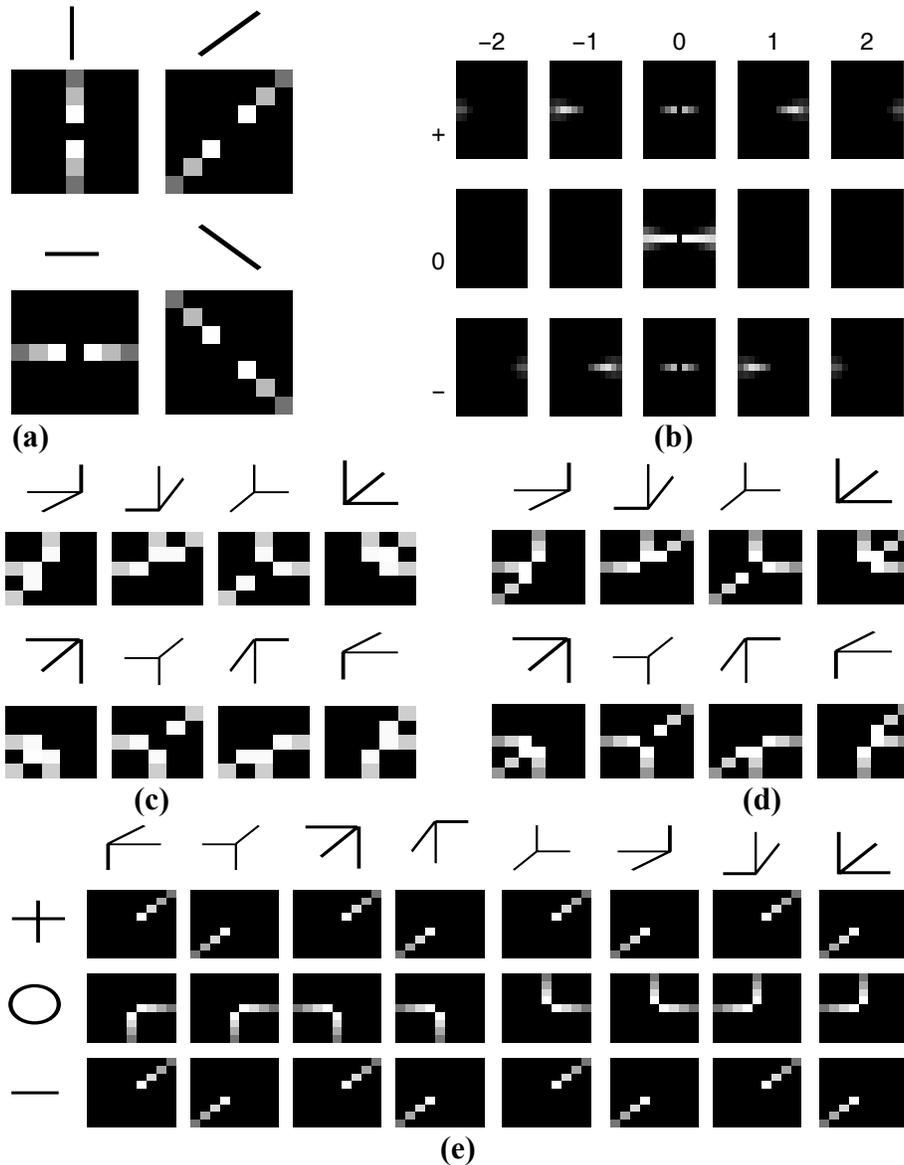
Figure 25. (a) Kernels for V1 colinear bipole cells. The four subplots show the connection weights for vertical, horizontal and two oblique oriented cells from other cells of similar orientation but different positions. (b) Kernels for the disparity-gradient cells. The figure shows the kernel for horizontally oriented disparity-gradient cells. For simplicity, the kernels are shown only across depths and not across disparity-gradients and orientations. The rows indicate different disparity-gradient cells—positive, zero, and negative—while the columns indicate different relative depths with respect to the depth of the disparity-gradient cell. The left columns indicate near relative depths while the right columns indicate far relative depths. (c) Kernels for V1 angle cells. The eight subplots depict the kernels for eight different angles of the Necker cube. The kernels show the summed weight across orientations for each angle. (d) Kernels for V2 angle cells. The kernels are the same as for V1 angle cells but slightly longer. (e) Activation of disparity-gradient cells by angle cells. The columns indicate the activation of 75 disparity-gradient cells by each of the eight angle cells. The rows represent the activation of positive, zero, and negative disparity-gradient cells by the corresponding angle cells.

63

$$H^{Ec}_{ijkdv} = \sum_{pq} W^c_{pqijkv} \left[c_{pqkd} - \zeta_c \right]^+, \tag{17}$$

where W^c_{pqijkv} in (17) is the long-range connection weight from side v [left ($v = o$) and right ($v = 1$) for colinear bipole cells] in layer 2/3A of V1 from colinear bipole cells at position (p,q), and orientation k to colinear bipole cells at position (i,j) and orientation k. The connection weights in (17) are defined for the horizontal orientation (k=3) as follows:

$$W^c_{pqij31} = \left[\text{sign}(i - p)exp\left(-\left(\frac{(i-p)^2}{\sigma^2_p} + \frac{(j-q)^2}{\sigma^2_q} \right) \right) \right]^+ \tag{18}$$

and

$$W^c_{pqij32} = \left[\text{sign}(p - i)exp\left(-\left(\frac{(i-p)^2}{\sigma^2_p} + \frac{(j-q)^2}{\sigma^2_q} \right) \right) \right]^+, \tag{19}$$

where $\text{sign}(x) = $ -1, if $x < 0$, 1 otherwise. The connection weights for other orientations are obtained by appropriate rotation. These weights are represented graphically in Figure 25a. Term I_{ijkd} in (16) is the bottom-up input from layer 3B cell at position (i,j) orientation k, and disparity d, generated by hand as described in Section 4.2.1. As for the developmental simulations, the bottom-up input is added to the input from the long-range connections. The activity of layer 2/3A cell that receive bottom-up input can hereby be modulated by the input from the long-range connections even if there is input from only one side.

Input H^{Ic}_{ijkd} in (16) is the inhibitory signal from the inhibitory interneurons, and is defined by:

$$H^{Ic}_{ijkd} = \sum_{rv} B^{cc+}_{kr} \left[s^c_{ijrdv} \right]^+ + \sum_{tv} B^{nc+}_{ktv} \left[s^n_{ijtdv} \right]^+. \tag{20}$$

Variable s^c_{ijrdv} in (20) represents the activity of inhibitory interneuron number v associated with the excitatory layer 2/3A colinear bipole cell at position (i,j), disparity d, and orientation r. Variable s^n_{ijtdv} in (20) represents the activity of inhibitory interneuron number v associated with an excitatory non-colinear (angle) bipole cell in layer 2/3A of V1 at position (i,j), disparity d, and angle type t. The inhibitory interneurons of a layer 2/3A colinear bipole cell of orientation r inhibit an excitatory colinear bipole cell of orientation k with weight B^{cc+}_{kr}, and the inhibitory

interneuron number v of layer 2/3A non-colinear bipole cell of angle type t inhibits an excitatory colinear bipole cell of orientation k with weight B_{ktv}^{nc+}. This inhibition from the inhibitory interneurons helps to maintain the balance between excitation and inhibition to enforce the bipole property; see equation (3).

Term $\gamma_3 [c_{ijkd} - \beta_c]^+$ in (16) is the self-excitatory feedback. Term $\gamma_3 \sum_o [g_{ijkdo} - \beta_g]^+$ in (16) is the feedback input from V2 disparity-gradient cells at position (i,j), orientation k, and disparity d summed across disparity-gradients o, to V1 colinear bipole cells. The inhibitory input, C_{ijkd}^P, at the same position and disparity from other angle and colinear bipole cells is defined by:

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

(21)

where n_{ijtd} is the activity of V1 layer 2/3A angle cell of angle type t at the same position (i,j) and disparity d. Since both the colinear bipole cell and angle cell receive same bottom-up input, this inhibition and the horizontal input from long-range connections help to disambiguate ambiguous activation of colinear bipole cells and angle cells such that angle cells are activated at corners and colinear bipole cells are activated along straight edges. Term C_{ijkd}^S in (16) is the inhibitory input from spatial and orientational competition across position and orientation but within disparity; namely,

$$C_{ijkd}^S = \sum_r Y_{ijrd} K(r - k),$$

(22)

where the orientation kernel $K(x) = 1$, if $|x| = 2$, 0 otherwise, and notation $|x|$ indicates the absolute value of x. In particular, term $K(r - k)$ in (22) is non-zero only if orientations r and k are perpendicular. Spatial competition term Y_{ijrd} in (22) is defined by:

$$Y_{ijrd} = \sum_{pq} [c_{pqrd} - \beta_c]^+ exp\left(-\left((i - p)^2 + (j - q)^2 \right) \right).$$

(23)

The spatial and orientational competition enables cells of perpendicular orientation to inhibit each other across positions to achieve T-junction sensitivity without using T-junction operators (Grossberg, 1994; Kelly and Grossberg, 2000).

The activity of the inhibitory interneurons are defined by:

$$\frac{d}{dt} s_{ijkdu}^c = \delta_I \left[-s_{ijkdu}^c + H_{ijkdu}^{E_c} - s_{ijkdu}^c \left(\sum_{rv} B_{kr}^{-cc} [s_{ijrdv}^c]^+ + \sum_{tv} B_{ktv}^{nc-} [s_{ijtdv}^n]^+ \right) \right],$$

(24)

where s^c_{ijkdu} represent the activity of the inhibitory interneuron number u associated with the excitatory cell at position (i,j), orientation k and disparity d. Term δ_I in (24) determines the rate at which the activity changes. Term H^{Ec}_{ijkdu} in (24) is as defined in (17). B^{cc-}_{kr} is the inhibitory weight between inhibitory interneurons of colinear bipole cells and B^{nc-}_{ktv} is the inhibitory weight between inhibitory interneurons of non-colinear and colinear bipole cells. This recurrent inhibition among the interneurons helps to normalize the total inhibition received by the excitatory cell to help implement the bipole property. The values for the weights are $B^{cc+}_{kr} = 0.33$, if $k = r$; 0.01, otherwise; $B^{nc+}_{ktv} = 0.33$, if the orientation of the arm v of the angle t is k, 0.01; otherwise; $B^{cc-} = 0.33$, and $B^{nc-} = 0.33$. The parameter values are $\alpha_c = 0.6$, $\gamma_1 = 0.6$, $\gamma_2 = 1.67$, $\gamma_3 = 0.15$, $\gamma_4 = 5$, $\beta_g = 0.3$, $\zeta_c = 0.02$, $\sigma_p = 3$, $\sigma_q = 0.3$, $\beta_c = 0.3$, $\beta_n = 0.3$, and $\delta_I = 3$.

B.2 V1 angle cell activation equations. The activity, n_{ijld}, of V1 layer 2/3A angle cell at position (i,j), angle type l, and disparity d, is defined by:

$$\frac{d}{dt} n_{ijld} = -\alpha_n n_{ijld}$$
$$+ \left(1 - n_{ijld}\right)\left(\gamma_1\left[\sum_v H^{En}_{ijldv} + \gamma_2 I^N_{ijd} - H^{In}_{ijld}\right]^+ + \gamma_3\left[n_{ijld} - \beta_n\right]^+ + \gamma_3\left[a_{ijld} - \beta_a\right]^+\right) \quad (25)$$
$$- \left(1 - n_{ijld}\right)N^P_{ijld}.$$

Eight different angle cells ($l=1, \dots 8$) corresponding to the eight corners of the cube (Figure 9) are used in the simulation. Other type of angles were not included to simplify the simulations. Since each angle cell gets activated most for a particular angle, including other types of angle cells will not change the result of the simulation. Parameter α_n in (25) is the decay rate. Term H^{En}_{ijldv} in (25) define the input received from the long-range connections for branch v of the angle cell, and is defined by:

$$H^{En}_{ijldv} = \sum_{pqr} W^n_{pqrijlv}\left[c_{pqrd} - \zeta_c\right]^+, \quad (26)$$

where term $W^{n}_{pqrijlv}$ is the connection weight from colinear bipole cell at position (p,q) and orientation r, to an angle cell of angle type l at position (i,j) for branch v of the angle cell. The connection weights are similar to the ones defined for colinear bipole cells in equations (18) and (19). In particular, the connection weights for branch 1 of the angle cell of angle type $l = 1$ (angle A in Figure 9), from horizontally oriented colinear bipole cells ($(r = 3)$), are defined as follows,

$$W^{n}_{pq3ij11} = \left[sign(i-p)\exp\left(-\left(\frac{(i-p)^2}{\sigma_p^2}+\frac{(j-q)^2}{\sigma_p^2}\right)\right)\right]^+. \tag{27}$$

The connection weights for the other branches and other angle cells are obtained by appropriate rotation, and are shown graphically in Figure 25c. Input $I^{N}_{ijd} = \sum_r I_{ijrd}$ in (25) sums the bottom-up input I_{ijrd} across orientation r at position (i,j) and disparity d. The angle cells receive the same bottom-up input as that of the colinear bipole cells, summed across orientations. The bottom-up input to V1 non-colinear bipole cells is added to the input from the horizontal long-range connections. As a result, horizontal interactions can modulate cell response.

Input H^{In}_{ijld} in (25) is the inhibitory input from the interneurons, and is defined by:

$$H^{In}_{ijld} = \sum_{tv} B^{nn+}_{ltv}\left[s^{n}_{ijtdv}\right]^+ + \sum_{rv} B^{cn+}_{lrv}\left[s^{c}_{ijrdv}\right]^+. \tag{28}$$

There are three inhibitory interneurons associated with each excitatory angle cell in layer 2/3A of V1. Variable s^{n}_{ijtdv} in (28) is the activity of the inhibitory interneuron number v associated with a non-colinear (angle) cell at position (i,j), angle type t, and disparity d. Variable s^{c}_{ijrdv} in (28) represent the activity of the inhibitory interneuron number v associated with the excitatory layer 2/3A colinear bipole cell at the same position and orientation r. Term B^{nn+}_{ltv} is the inhibitory weight from the interneuron number v associated with an angle cell of angle type t to an excitatory angle cell of angle type l, and B^{cn+}_{ltv} is the inhibitory weight from the interneuron number v associated with a colinear bipole cell of orientation r to an excitatory angle cell of angle type l. The inhibition from the inhibitory interneurons helps to maintain the balance between excitation and inhibition to enforce the bipole property.

Term $\gamma_3 [n_{ijld} - \beta_n]^+$ in (25) is the self-excitatory feedback and term $\gamma_3 [a_{ijld} - \beta_a]^+$ in (25) is the feedback input from V2 angle cells. Angle cells in layer 2/3A of V1 also receive inhibitory input from other angle cells and colinear bipole cells, at the same position and disparity, but across orientations and angles, depicted by the term, N_{ijld}^P in (25), which is defined by:

$$N_{ijld}^P = \gamma_4 \left(\sum_{t \neq l} [n_{ijtd} - \beta_n]^+ + \sum_r [c_{ijrd} - \beta_c]^+ \right). \tag{29}$$

In (29), variable n_{ijtd} is the activity of the V1 angle cell at position (i,j), angle type t and disparity d, and c_{ijrd} is the activity of the V1 colinear bipole cell at the same position and disparity, but orientation r. The inhibition helps to disambiguate ambiguous activation of colinear bipole cells and angle cells such that angle cells are activated at corners and colinear bipole cells are activated along straight edges.

The activity of the inhibitory interneurons is defined by:

$$\frac{d}{dt} s_{ijldu}^n = \delta_I \left[-s_{ijldu}^n + H_{ijldu}^{En} - s_{ijldu}^n \left(\sum_{tv} B^{nn-} [s_{ijtdv}^n]^+ + \sum_{rv} B^{cn-} [s_{ijrdv}^c]^+ \right) \right], \tag{30}$$

where δ_I is the rate at which the activity changes, and H_{ijldu}^{En} is defined in (26). Term B^{nn-} is the inhibitory weight between inhibitory interneurons of angle cells. B^{cn-} is the inhibitory weight from inhibitory interneurons of colinear bipole cells to angle cells. This inhibition normalizes the total inhibition and helps to implement the bipole property. The values for the weights are $B_{ltv}^{nn+} =$ 0.33, if angle l and t share the same branch of that of the inhibitory interneuron v, 0.01, otherwise; $B_{lrv}^{cn+} = 0.33$, if the orientation of the arm v of the angle l is r, 0.01, otherwise; $B^{nn-} =$ 0.33, and $B^{cn-} = 0.33$. The parameter values are $\alpha_n = 0.6$, $\gamma_1 = 0.6$, $\gamma_2 = 1.67$, $\gamma_3 = 0.15$, $\gamma_4 = 5$, $\beta_a = 0.3$, $\zeta_c = 0.02$, $\sigma_p = 4$, $\sigma_q = 0.3$, $\beta_n = 0.3$, $\beta_c = 0.3$, and $\delta_I = 3$.

B.3 V2 angle cells

B.3.1 Activation equations. Layer 2/3A of V2 in the model contains angle cells and disparity-gradient cells. As with V1 layer 2/3A, V2 layer 2/3A contains eight different angle cells

corresponding to the eight corners of the cube. The activity, a_{ijld}, of V2 layer 2/3A angle cell at position (i,j), angle type l, and disparity d, is defined by:

$$\frac{d}{dt}a_{ijld} = \delta_E\left(-\alpha_a a_{ijld} + \left(1-a_{ijld}\right)h_{ijld}^{aE}\left[\sum_v H_{ijldv}^{Ea} + \gamma_1 I_{ijld} - H_{ijld}^{Ia}\right]^+ - \left(1+a_{ijld}\right)A_{ijld}^P\right),$$

(31)

where δ_E is the rate at which the activity of V2 angle cell changes. Term α_a in (31) is the decay rate and h_{ijld}^{aE} in (31) indicates the level of excitatory habituative transmitter (see equation (38)) associated with this angle cell. Habituative transmitters which were useful for the self-organization of angle and colinear cells in the developmental simulations enable the network to switch between alternate interpretations of the Necker cube. Term H_{ijldv}^{Ea} in (31) represents long-range connections in layer 2/3A of V2 from disparity-gradient cells for each branch v of the angle cell. It is defined by:

$$H_{ijldv}^{Ea} = \sum_{pqr}W_{pqrijlv}^{a}\sum_o\left[g_{pqrdo} - \zeta_g\right]^+,$$

(32)

where g_{pqrdo} in (32) represents the activity of disparity-gradient cells at position (p,q), orientation r, disparity d and disparity-gradient o, and $W_{pqrijlv}^{a}$ is the connection weight from the disparity-gradient cell at position (p,q) and orientation r, to an angle cell of angle type l, at position (i,j), for branch v of the angle cell. The connection weights are the same as described for V1 angle cells but slightly longer. In particular, the connection weights for an angle cell branch 1, angle type $l=1$, from horizontally oriented disparity-gradient cells ($r=3$), are defined as follows:

$$W_{pq3ij11}^{a} = \left[sign(i-p)\exp\left(-\left(\frac{(i-p)^2}{\sigma_p^2}\right) + \left(\frac{(j-q)^2}{\sigma_p^2}\right)\right)\right]^+.$$

(33)

The connection weights for the other branches are obtained by appropriate rotation and are shown graphically in Figure 25d. Term $I_{ijld} = [n_{ijld} - \beta_n]^+$ in (31) is the bottom-up input from V1 angle cells at the same position (i,j), angle type l, and disparity d. As for the V1 layer 2/3A

neurons, the bottom-up input is added to the input from long-range connections. Term H_{ijld}^{la} in (31) is the inhibitory input from the interneurons and is defined by:

$$H_{ijld}^{la} = \sum_{tv} B_{ltv}^{aa+} \left[s_{ijtdv}^{a} \right]^{+} + \sum_{rv} B_{lrv}^{gn+} \sum_{o} \left[s_{ijrdov}^{g} \right]^{+}. \tag{34}$$

As described for V1 angle cells, each V2 angle cell is associated with three inhibitory interneurons. Variable s_{ijtdv}^{a} in (34) is the activity of the inhibitory interneuron number v at position (i,j) angle type t, and disparity d, and variable s_{ijrdov}^{g} in (34) is the activity of the inhibitory interneuron number v associated with the disparity-gradient cell at the same position and disparity, but disparity-gradient o and orientation r. Term B_{ltv}^{aa+} is the inhibitory weight from the interneuron number v associated with an excitatory V2 angle cell of angle type t to an excitatory angle cell of angle type l, and B_{lrv}^{ga+} is the inhibitory weight from the interneuron number v associated with a disparity-gradient cell of orientation r to an excitatory angle cell of angle type l. The inhibition from the inhibitory interneurons helps to maintain the balance between excitation and inhibition to enforce the bipole property.

Angle cells in layer 2/3A of V2 also receive inhibitory input from other angle cells and disparity-gradient cells at the same position, but across disparities and angles, defined by term A_{ijld}^{P} in (31):

$$A_{ijld}^{P} = \gamma_3 \left(\sum_{t \neq l, e \neq d} h_{ijte}^{aI} \left[a_{ijte} - \beta_a \right]^{+} + \sum_{reo} h_{ijreo}^{gI} \left[g_{ijreo} - \beta_g \right]^{+} \right), \tag{35}$$

where variable a_{ijte} is the activity of a V2 angle cell at position (i,j), angle type t and disparity e, and variable g_{ijreo} is the activity of a disparity-gradient cell at the same position and disparity, but orientation r, and disparity-gradient o. Term h_{ijte}^{aI} in (35) represents the inhibitory habituative transmitter associated with the angle cell at position (i,j), angle type t and disparity e; see equation (39). Similarly, term in (35) represents the inhibitory habituative transmitter associated with a disparity-gradient cell at position (i,j), orientation r, disparity e, and disparity-gradient o; see equation (40). Both excitatory and inhibitory habituative transmitters were used in the Necker cube simulations. The inhibitory habituation helps the network to switch to an alternate interpretation before the activity of the neurons representing the present interpretation goes below the threshold for inhibition. The activity of the inhibitory interneurons is defined by:

$$\frac{d}{dt} s^a_{ijldu} = \delta_I \left[-s^a_{ijldu} + H^{Ea}_{ijldu} - s^a_{ijldu} \left(\sum_{tv} B^{aa-} \left[s^a_{ijtdv} \right]^+ + \sum_{rv} B^{ga-} \sum_o \left[s^g_{ijrdov} \right]^+ \right) \right]. \tag{36}$$

Term δ_I is the rate at which activity changes, and H^{Ea}_{ijldu} is defined in (32). Term B^{aa-} is the inhibitory weight between inhibitory interneurons of angle cells and B^{ga-} is the inhibitory weight between inhibitory interneurons of angle and disparity-gradient cells. The recurrent inhibition among the interneurons normalizes the total inhibition and helps to implement the bipole property. The values for the weights are $B^{aa+}_{ltv} = 0.4$, if angle l and t share the same branch of that of the inhibitory interneuron v; 0.01, otherwise; $B^{ga+}_{lrv} = 0.4$, if the orientation of the arm v of the angle l is r; 0.01, otherwise; $B^{aa-} = 0.4$, and $B^{ga-} = 0.4$. The parameter values are $\delta_E = 0.5$, $\alpha_a = 0.8$, $\gamma_1 = 1.2$, $\gamma_2 = 0.7$, $\gamma_3 = 2.0$, $\beta_n = 0.2$, $\zeta_g = 0.1$, $\sigma_p = 15$, $\sigma_q = 0.1$, $\beta_a = 0.25$, $\beta_g = 0.1$, and $\delta_I = 3$.

B.3.2 Habituation of excitatory input. In equation (31), the total input,

$$J^a_{ijld} = \left[\sum_v H^{Ea}_{ijldv} + \gamma_1 I_{ijld} - H^{Ia}_{ijld} \right]^+ \tag{37}$$

to each layer 2/3A angle cell is multiplicatively gated by habituative transmitter, h^{aE}_{ijld}, that obeys the following equation:

$$\frac{d}{dt} h^{aE}_{ijld} = \left(A_h + \frac{1}{1 + \theta \left[J^a_{ijld} - \varepsilon \right]^+} \right) \left(1 - h^{aE}_{ijld} \right) - B_h h^{aE}_{ijld} \left[J^a_{ijld} - \varepsilon \right]^+. \tag{38}$$

The interpretation of (38) is the same as that of (7). The parameter values are $A_h = 0.0065$, $B_h = 0.0585$, $\theta = 1000$, and $\varepsilon = 0.3697$.

B.3.3 Habituation of inhibitory input from V2 angle cells. The inhibitory input, $\sum_{t \neq l, e \neq d} h^{aI}_{ijte} \left[a_{ijte} - \beta_a \right]^+$, from other V2 angle cells in (35) is gated by an inhibitory habituative transmitter, h^{aI}_{ijld}, whose dynamics are defined by:

$$\frac{d}{dt} h^{aI}_{ijld} = \left(A_h + \frac{1}{1 + \theta \left[a_{ijld} - \beta_a \right]^+} \right) \left(1 - h^{aI}_{ijld} \right) - B_h h^{aI}_{ijld} \left[a_{ijld} - \beta_a \right]^+. \tag{39}$$

The habituative transmitter dynamics are the same as described before, except that the habituation occurs at a rate proportional to the total signal $[a_{ijld} - \beta_a]^+$ that the transmitter gates. The parameter values are $A_h = 0.005$, $A_h = 0.03$, $\theta = 1000$, and $\beta_a = 0.25$.

B.3.4 Habituation of inhibitory input from V2 disparity-gradient cells. The inhibitory input, $\sum_{reo} h_{ijreo}^{gI}[g_{ijreo} - \beta_g]^+$, from V2 disparity-gradient cells in (35) is gated by inhibitory habituative transmitter, h_{ijkdm}^{gI}, whose dynamics are defined as follows:

$$\frac{d}{dt} h_{ijkdm}^{gI} = \left(A_h + \frac{1}{1 + \theta[g_{ijkdm} - \beta_g]^+} \right)\left(1 - h_{ijkdm}^{gI}\right) - B_h h_{ijkdm}^{gI}\left[g_{ijkdm} - \beta_g\right]^+ . \tag{40}$$

The habituative transmitter dynamics are the same as described before for V2 angle cells except that the habituation occurs at a rate proportional to the total signal $[g_{ijkdm} - \beta_g]^+$ that the transmitter gates. The parameter values are $A_h = 0.005$, $B_h = 0.045$, and $\theta = 1000$.

B.4 V2 disparity-gradient cells

B.4.1 Activation Equations. As described in Section 4.2, three different types of disparity-gradient cells corresponding to negative, positive and zero disparity gradients, were used in the simulation. The activity, g_{ijkdm}, of a V2 layer 2/3A disparity-gradient cell at position (i,j), orientation k, disparity d, and disparity-gradient m, is defined by:

$$\frac{d}{dt} g_{ijkdm} = \delta_E \left(-\alpha_g g_{ijkdm} + \left(1 - g_{ijkdm}\right) h_{ijkdm}^{gE}\left[\sum_v H_{ijkdmv}^{Eg} + \gamma_1 I_{ijkdm}^g - H_{ijkdm}^{Ig} \right]^+ \right.$$
$$\left. - \left(1 + g_{ijkdm}\right)\left(G_{ijkdm}^P + G_{ijkdm}^s \right) \right), \tag{41}$$

where δ_E determines the overall rate at which the activity of the neuron changes. Term α_g is the decay rate, and h_{ijkdm}^{gE} is the excitatory habituative transmitter; see equation (51). V2 disparity-gradient cells receive long-range input from other disparity-gradient cells in nearby positions and disparities. Term H_{ijkdmv}^{Eg} in (41) is the input from the branch v of the disparity-gradient cells:

$$H_{ijkdmv}^{Eg} = \sum_{pqe} W_{pqeijkdmv}^g\left[g_{pqkem} - \zeta_g\right]^+ . \tag{42}$$

The connection weight, $W^g_{pqeijkdmv}$ in (42), is described in detail in Section B.4.4 below. In particular, $W^g_{pqeijkdmv}$ is the connection weight from a disparity-gradient cell at position (p,q), orientation k, disparity e, and disparity-gradient m, for branch v of a disparity-gradient cell at position (i,j), orientation k, disparity d, and disparity-gradient m. Term I^g_{ijkdm} in (41) is the input from V1 colinear bipole cells and V2 angle cells:

$$I^g_{ijkdm} = \gamma_2 \left[c_{ijkd} - \beta_c \right]^+ + \gamma_3 \sum_{pqt} W^{ga}_{pqtijkm} \left[a_{pqtd} - \zeta_a \right]^+, \tag{43}$$

where $\left[c_{ijkd} - \beta_c \right]^+$ is the bottom-up input from V1 bipole colinear cells and $\gamma_3 \sum_{pqt} W^{ga}_{pqtijkm} \left[a_{pqtd} - \zeta_a \right]^+$ is the input from V2 angle cells. The V2 angle cells selectively activate the disparity-gradient cells, as described in Section 3. In particular, $W^{ga}_{pqtijkm}$ defines the weight from an angle cell at position (p,q) and angle type t, to a disparity-gradient cell at position (i,j), orientation k, and disparity-gradient m. The connection weight from an angle cell of angle type $t = 1$ to a horizontally oriented ($k = 3$), zero disparity-gradient cell ($m = 1$) is defined by:

$$W^{ga}_{pq1ij31} = \left[\text{sign}(i-p) exp\left(-\left(\frac{(i-p)^2}{\sigma_p^2} + \frac{(j-q)^2}{\sigma_p^2} \right) \right) \right]^+. \tag{44}$$

The connection weights for the other orientations are obtained by appropriate rotation and are shown in Figure 25e. Terms γ_2 and γ_3 in (43) control how much of the binocular input from V1 colinear bipole cells and monocular input from V2 angle cells can affect the disparity-gradient cells; see Section 5.4. As for the V2 angle cells, the bottom-up input is added to the input from the long-range connections.

Term H^{Ig}_{ijkdm} in (41) is the inhibitory input from the interneurons:

$$H^{Ig}_{ijkdm} = \sum_{rov} B^{gg+}_{krmo} \left[s^g_{ijrdov} \right]^+ + \sum_{tv} B^{ag+}_{ktv} \left[s^a_{ijtdv} \right]^+ \tag{45}$$

As with the V1 colinear bipole cells in (16), each V2 disparity-gradient cell is associated with two inhibitory interneurons. Variable s^g_{ijrdov} is the activity of the inhibitory interneuron number v associated with the disparity-gradient cell at position (i,j), orientation r, disparity d, and disparity-gradient o. Variable B^{gg+}_{krmo} in (45) is the connection weight from inhibitory interneurons

of a disparity-gradient cell of disparity-gradient o and orientation r to a disparity-gradient cell of disparity-gradient m and orientation k at the same position and disparity. Similarly, variable B_{ktv}^{ag+} in (45) is the weight from inhibitory interneuron number v of an angle cell type t to a disparity-gradient cell of orientation k at the same position and disparity. The inhibition from the interneurons balances the excitation from long-range connections to implement the bipole property.

V2 disparity-gradient cells also receive inhibitory input from the same position due to other disparity-gradient cells of different disparity gradient and disparity, and from V2 angle cells. Term G_{ijkdm}^{P} in (41) represents this inhibitory input:

$$G_{ijkdm}^{P} = \gamma_4 \left(\sum_{te} h_{ijte}^{aI} \left[a_{ijte} - \beta_a \right]^+ + \sum_{e \neq d, o \neq m} h_{ijkeo}^{gI} \left[g_{ijkeo} - \beta_g \right]^+ \right). \tag{46}$$

This inhibition helps to disambiguate ambiguous activation of the disparity-gradient cells and V2 angle cells in layer 2/3A of V2. Term h_{ijte}^{aI} is the inhibitory habituation at V2 angle cells; see equation (39). Term h_{ijkeo}^{gI} is the inhibitory habituation at V2 disparity-gradient cells; see equation (40). Term G_{ijkdm}^{S} in (41) is the inhibitory input from spatial and orientational competition across position and orientation, but within disparity:

$$G_{ijkdm}^{S} = \sum_{r} Y_{ijrdm} K(r-k), \tag{47}$$

where $K(x) = 1 = 1$, if $|x| = 2$; 0 otherwise. Term Y_{ijrdm} in (47) is defined by:

$$Y_{ijrdm} = \sum_{pq} \sum_{o} h_{pqrdo}^{gI} \left[g_{pqrdo} - \theta_g \right]^+ exp\left(-\left((i-p)^2 + (j-q)^2 \right) \right), \tag{48}$$

which is analogous to (23). As in (22), the spatial and orientational competition enables cells of perpendicular orientation to inhibit each other across positions to achieve T-junction sensitivity without T-junction operators (Grossberg, 1994).

The activity of the inhibitory interneurons are defined by:

$$\frac{d}{dt} s_{ijkdmu}^{g} = \delta_I \left[-s_{ijkdmu}^{g} + H_{ijkdmu}^{Ed} - s_{ijkdmu}^{g} \left(\sum_{rov} B^{gg-} \left[s_{ijrdov}^{g} \right]^+ + \sum_{tv} B^{ag-} \left[s_{ijtdv}^{a} \right]^+ \right) \right], \tag{49}$$

where δ_I is the rate at which the activity of the inhibitory interneuron changes, and s^g_{ijkdmu} represent the activities of inhibitory interneuron number u associated with the disparity-gradient cell at position (i,j), orientation k, disparity d, and disparity-gradient m. Term H^{Eg}_{ijkdmu} is defined in (42); B^{gg-} is the inhibitory weight between inhibitory interneurons of disparity-gradient cells; and B^{ag-} is the inhibitory weight between inhibitory interneurons of angle cells and disparity-gradient cells. This inhibition among the interneurons helps to normalize the total inhibition received by the excitatory cell so as to implement the bipole property. The values for the weights are $B^{gg+}_{krmo} = 1.2$, if $k = r$ and $m = o$; 0.01, otherwise; $B^{ag+}_{ktv} = 0.4$, if the orientation of the arm v of the angle t is k; 0.01, otherwise; $B^{gg-} = 1.2$ and $B^{ag-} = 0.4$. The parameter values are $\delta_E = 0.5$, $\alpha_g = 0.8$, $\gamma_1 = 1.7$, $\gamma_2 = 3$, $\gamma_3 = 1$, $\gamma_4 = 2$, $\beta_c = 0.2$, $\zeta_a = 0.1$, $\zeta_g = 0.05$, $\beta_g = 0.1$, $\beta_a = 0.25$, and $\theta_g = 0.15$, and $\delta_I = 3$.

B.4.2 Habituation of excitatory input. In equation (41), the total input,

$$J^g_{ijkdm} = \left[\sum_v H^{Eg}_{ijkdmv} + \gamma_1 I^g_{ijkdm} - H^{Ig}_{ijkdm} \right]^+ \tag{50}$$

to each layer 2/3A disparity-gradient cell is multiplicatively gated by habituative transmitter, h^{gE}_{ijkdm}, that obeys the following equation:

$$\frac{d}{dt} h^{gE}_{ijkdm} = \left(A_h + \frac{1}{1 + \theta\left[J^g_{ijkdm} - \varepsilon\right]^+} \right)\left(1 - h^{gE}_{ijkdm}\right) - B_h h^{gE}_{ijkdm} \left[J^g_{ijkdm} - \varepsilon\right]^+, \tag{51}$$

as in (7). The parameter values are $A_h = 0.005$, $B_h = 0.030$, $\theta = 1000$, and $\varepsilon = 0.64$.

B.4.3 Habituation of inhibitory input. The inhibitory inputs in G^P_{ijkdm} of equation (46) habituate via the same transmitter equations h^{aI}_{ijld} and h^{gI}_{ijkdm} as in (39) and (40) because the level of habituation is determined only by the source angle and disparity-gradient cells, respectively.

B.4.4 Disparity-gradient cell kernels. These kernels generalize 2D bipole kernels to enable 3D groupings of slanted and curved contours. The connection weight, $W^g_{ijd00k1mv}$ in (42), from a disparity-gradient cell at position (i,j), disparity d, orientation k, and disparity-gradient m for the

left and right branches ($v = 1$ and $v = 2$) of a disparity-gradient cell at position $(0,0)$, zero disparity, orientation k and disparity-gradient m, is defined as follows:

$$W^g_{ijd\,0030m1} = \left[H_{ijkdm} \right]^+ ,$$ (52)

and

$$W^g_{ijd\,0030m2} = \left[-H_{ijkdm} \right]^+ .$$ (53)

The connection weight for other orientations and disparity-gradients are obtained by appropriate rotation. Term H_{ijkdm} is, in turn, defined by:

$$H_{ijkdm} = sign(i) e^{-\Gamma_{ij}} e^{-\Gamma_k} e^{-\Gamma_{kk}} e^{-\Gamma_d} e^{-\Gamma_m} e^{-\Gamma_{mm}} ,$$ (54)

and consists of six terms that determine how the weight values vary as a function of the differences in distance, orientation, disparity, and disparity-gradient between the source and the target cells. This kernel generalize the bipole kernel used in Gove et al. (1995) to the case of 3D grouping of both slanted and curved contours. The individual terms in (54) have a multiplicative effect on the final value of the weight. Because each term is an exponential, all the terms in (54) can be combined within a single exponential term

$$H_{ijkdm} = sign(i) exp\left(-\Gamma_{ij} - \Gamma_k - \Gamma_{kk} - \Gamma_d - \Gamma_m - \Gamma_{mm} \right) .$$ (55)

The first term, $e^{-\Gamma_{ij}}$ in (54), describes how the connection weight decreases as a Gaussian function of the distance between the two cells. Thus:

$$\Gamma_{ij} = \frac{(i+j)^2}{2\sigma_{ij}^2} .$$ (56)

The second term, $e^{-\Gamma_k}$, decreases as a Gaussian function of the orientation of the position (i,j) of the target cell with respect to the preferred horizontal grouping of the source cell at position $(0,0)$. Thus:

$$\Gamma_k = \frac{K^2}{2\sigma_k^2} ,$$ (57)

where K is the orientation of position (i,j) with respect to $(0,0)$, namely:

$$K = \tan^{-1} \left(\frac{i}{\frac{i^2 + j^2}{2j} - j} \right) .$$ (58)

The third term, $e^{-\Gamma_{kk}}$, decreases as a Gaussian function of the difference between the preferred orientation k of the cell at position (i,j), with respect to K:

$$\Gamma_{kk} = \frac{(k-K)^2}{2\sigma_{kk}^2}. \tag{59}$$

The two terms, Γ_k and Γ_{kk}, enable the network to complete boundaries smoothly in a way that satisfies the relatability conditions of Kellman and Shipley (1991). The remaining three terms help to realize 3D grouping: The fourth term, $e^{-\Gamma_d}$, decreases as a Gaussian function of the difference in disparity between the source and the target cell:

$$\Gamma_d = \frac{d^2}{2\sigma_d^2}. \tag{60}$$

The fifth term, $e^{-\Gamma_m}$, decreases as a Gaussian function of the difference in disparity gradient of the source cell at position $(0,0)$ and disparity 0 with respect to the position (i,j) and disparity d of the target cell. Thus:

$$\Gamma_m = \frac{M^2}{2\sigma_m^2}, \tag{61}$$

where M is the disparity gradient of (i,d) with respect to $(0,0)$ in the space-disparity plane:

$$M = tan^{-1}\left(\frac{i}{\frac{i^2+d^2}{2d}-d}\right). \tag{62}$$

Since the kernel is defined for disparity-gradient cells of horizontal orientation, index j is not used in the above equation, as $j=0$ for horizontal orientation. The sixth term, $e^{-\Gamma_{mm}}$, decreases as a Gaussian function of the difference between the preferred disparity gradient m of the cell at position (i,j), with respect to M:

$$\Gamma_{mm} = \frac{(m-M)^2}{2\sigma_{mm}^2}. \tag{63}$$

The two terms Γ_m and Γ_{mm} are similar to terms Γ_k and Γ_{kk} except that the former two operate in the space-disparity domain while the latter two operate in 2D space. The parameter values are $\sigma_{ij}=10$, $\sigma_k=0.5$, $\sigma_{kk}=0.3$, $\sigma_d=4$, $\sigma_m=0.1$, σ_{mm} and $=0.1$.

B.5 Surface filling-in equations. The boundaries represented by disparity-gradient cells act as a barrier to the 3D filling-in process in V4. The filling-in equations generalize those used in Grossberg and Todorović (1988) by developing the proposal in Grossberg (1994) that the boundary signal at a particular depth acts as a barrier to filling-in signal at that depth, as well as a weak barrier at other depths; see Figure 5.

The model LGN (see Figure 2a) discounts the illuminant and computes Weber-law modulated and normalized estimates of image contrasts above an adaptation level; see Grossberg (1980) and Mingolla et al. (1999). To accomplish this, the LGN ON and OFF activities obey on-center off-surround, and off-center on-surround, shunting networks, respectively:

$$\frac{d}{dt}x_{ij}^+ = -\alpha_x x_{ij}^+ + \left(U - x_{ij}^+\right)P_{ij} - \left(L + x_{ij}^+\right)Q_{ij} \tag{64}$$

and

$$\frac{d}{dt}x_{ij}^- = -\alpha_x x_{ij}^- + \left(U - x_{ij}^-\right)P_{ij} - \left(L + x_{ij}^-\right)Q_{ij}, \tag{65}$$

where α_x is the decay constant and U and $-L$ are the upper and lower activity bounds, respectively. The on-center, P_{ij}, and the off-surround, Q_{ij}, are defined by Gaussian kernels:

$$P_{ij} = \sum_{pq} C_{pq} I_{i+p,j+q} \tag{66}$$

and

$$Q_{ij} = \sum_{pq} S_{pq} I_{i+p,j+q} \quad , \tag{67}$$

where term I_{ij} is the input, and

$$C_{pq} = \frac{A_1}{2\pi\sigma_c^2} exp\left\{-\frac{p^2 + q^2}{2\sigma_c^2}\right\} \tag{68}$$

and

$$S_{pq} = \frac{A_2}{2\pi\sigma_s^2} exp\left\{\frac{p^2 + q^2}{2\sigma_s^2}\right\}. \tag{69}$$

The width of the center and surround are described by the parameters σ_c and σ_s. At equilibrium, (64) and (65) become:

$$x_{ij}^+ = \frac{\left[\sum_{pq}\left(UC_{pq} - LS_{pq}\right)I_{i+p,j+q}\right]^+}{\alpha_x + \sum_{pq}\left(C_{pq} + S_{pq}\right)I_{i+p,j+q}} \tag{70}$$

and

$$x_{ij}^- = \frac{\left[\sum_{pq}\left(US_{pq} - LC_{pq}\right)I_{i+p,j+q}\right]^+}{\alpha_x + \sum_{pq}\left(C_{pq} + S_{pq}\right)I_{i+p,j+q}}. \tag{71}$$

The difference of these ON and OFF activities is computed to generate opponent output signals:

$$X_{ij}^+ = \left[x_{ij}^+ - x_{ij}^-\right]^+ \tag{72}$$

and

$$X_{ij}^- = \left[x_{ij}^- + x_{ij}^+\right]^+, \tag{73}$$

where X_{ij}^+ is the ON LGN output and X_{ij}^- is the OFF LGN output. The activity of the filling-in cells is defined by:

$$\frac{d}{dt}F_{ijd}^+ = -\alpha_f F_{ijd}^+ + \sum_{(p,q)\in N}\left(F_{pqd}^+ - F_{ijd}^+\right)\psi_{ijpqd} + X_{ij}^+ \tag{74}$$

and

$$\frac{d}{dt}F_{ijd}^- = -\alpha_f F_{ijd}^- + \sum_{(p,q)\in N}\left(F_{pqd}^- - F_{ijd}^-\right)\psi_{ijpqd} + X_{ij}^-, \tag{75}$$

where F_{ijd}^+ is the activity of the cell the ON FIDO (see Section 4.2.6) at position (i,j) and disparity d, and F_{ijd}^- is the activity of the corresponding cell in the OFF FIDO. Term α_f is the decay rate, and the boundary-gated diffusion coefficient ψ_{ijpqd} is defined by:

$$\psi_{ijpqd} = \frac{\delta}{1 + \varepsilon\left(Z_{ijd} + Z_{pqd} + \theta_1\left(Z_{ijL} + Z_{pqL}\right) + \theta_2\left(Z_{ijM} + Z_{pqM}\right)\right)}, \tag{76}$$

where

$$Z_{ijd} = \sum_{ro}\left[g_{ijrdo} - \theta_f\right]^+ \tag{77}$$

is the boundary signal at position (i,j) and disparity d that creates resistive barriers to the diffusion process at that depth when it is activated by disparity-gradient cell signals $[g_{ijrdo} - \theta_f]^+$. Term N in (74) and (75) consists of four nearest neighbors to a cell. The boundary signals in (77) from other depths, Z_{ijL} and Z_{ijM}, act as a weak barriers to the diffusion. Parameters θ_1 and θ_2 control how much of the boundary signal at a particular depth can influence the filling-in signals at different depths. Since only three depths are used in the simulations (D1, D2, and D3; see Section 4.2), subscript d refers to the boundary signal at the same depth, L refers to the boundary signal at the nearest depth, and M refers to the boundary signal at the second nearest depth. The terms δ and ε control how much a boundary signal can block the diffusion process. The parameter values are $\alpha_f = 0.1$, $\delta = 100$, $\varepsilon = 1000$, $\theta_1 = 0.1$, $\theta_2 = 0.05$, $\theta_f = 0.1$, $\alpha_x = 100$, $U = 50$, $L = 50$, $A_1 = 1$, $A_2 = .03361$, $\sigma_c = 0.5$, and $\sigma_s = 1.5$.

REFERENCES

Abbott, L. G., Varela, J. A., Sen, K., and Nelson, S. B. (1997). Synaptic depression and cortical gain control. *Science*, *275*, 220-224.

Albert, M. K. (1999). Surface formation and depth in monocular scene perception. *Perception*, *28*, 1361-1372.

Amir, Y., Harel, M., and Malach, R. (1993). Cortical hierarchy reflected in the organization of intrinsic connections in Macaque monkey visual cortex. *Journal of Comparative Neurology*, *334*, 19-46.

Bakin, J. S., Nakayama, K., and Gilbert, C. D. (2000). Visual responses in monkey areas V1 and V2 to three-dimensional surface configurations. *Journal of Neuroscience*, *20*(21), 8188-8198.

Bosking, W. H., Zhang, Y., Schofield, B., and Fitzpatrick, D. (1997). Orientation selectivity and the arrangement of horizontal connections in the tree shrew striate cortex. *Journal of Neuroscience*, *17*(6), 2112-2127.

Bringuier, 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.

Callaway, E. M. (1998). Local circuits in primary visual cortex of Macaque monkey. *Annual Review of Neuroscience*, *21*, 47-74.

Callaway, E. M., and Wiser, A. K. (1996). Contributions of individual layer 2-5 spiny neurons to local circuits in Macaque primary visual cortex. *Visual Neuroscience*, *13*, 907-922.

Carman, G. J., and Welch, L. (1992). Three-dimensional illusory contours and surfaces. *Nature*, *360*(10), 585-587.

Carpenter, G. and Grossberg, S. (1981). Adaptation and transmitter gating in vertebrate photoreceptors. *Journal of Theoretical Neurobiology*, *1*, 1-42. Reprinted in: S. Grossberg (Ed.) (1987) The Adaptive Brain. Amsterdam: North Holland.

Crook, J. M., Engelmann, R., and Lowel, S. (2002). Gaba-inactivation attenuates colinear facilitation in cat primary visual cortex. *Experimental Brain Research*, *143*, 295-302.

DeAngelis, G. C. (2000). Seeing in three dimensions: the neurophysiology of stereopsis. *Trends in Cognitive sciences*, *4*(3), 80-90.

Garrigan, P. and Kellman, P. J. (2002). Perception of 3D illusory contours: A study of contour relatability in depth. In *Vision sciences society abstract 357*, San Diego, USA.

Gove, A., Grossberg, S., and Mingolla, E. (1995). Brightness perception, illusory contours, and corticogeniculate feedback. *Visual Neuroscience*, *12*, 1027-1052.

Grosof, D. H., Shapley, R. M., and Hawken, M. J. (1993). Macaque V1 neurons can signal "illusory contours". *Nature*, *365*, 550-552.

Grossberg, S. (1968). Some nonlinear networks capable of learning a spatial pattern of arbitrary complexity. *Proceedings of National academy of Sciences*, *59*, 368-372.

Grossberg, S. (1969). On the production and release of chemical transmitters and related topics in cellular control. *Journal of Theoretical Biology*, *22*, 325-364.

Grossberg, S. (1973). Contour enhancement, short-term memory and constancies in reverberating neural networks. *Studies in Applied Mathematics*, *52*, 217-257.

Grossberg, S. (1976a). Adaptive pattern classification and universal recoding, I: Parallel development and coding of neural feature detectors. *Biological Cybernetics*, *23*, 121-134.

Grossberg, S. (1976b). Adaptive pattern classification and universal recoding II: Feedback, expectation, olfaction, and illusions. *Biological Cybernetics*, *23*, 187-202.

Grossberg, S. (1980). How does a brain build cognitive code? *Psychological Review*, *87*, 1-51.

Grossberg, S. (1984). Outline of a theory of brightness, color, and form perception. In Degreef, E. van Buggenhaut, J. (Eds.), *Trends in mathematical psychology*, pages 5559-5586. Elsevier, North Holland.

Grossberg, S. (1987). Cortical dynamics of three-dimensional form, color, and brightness perception: II. binocular theory. *Perception and Psychophysics*, *41*(2), 117-158.

Grossberg, S. (1994). 3-D vision and figure-ground separation by visual cortex. *Perception and Psychophysics*, *55*(1), 48-120.

Grossberg, S. (1997). Cortical dynamics of three-dimensional figure-ground perception of two-dimensional pictures. *Psychological Review*, *104*(3), 618-658.

Grossberg, S. (1999). How does the cerebral cortex work? learning, attention and grouping by the laminar circuits of visual cortex. *Spatial Vision*, *12*(2), 163-185.

Grossberg, S. (2000). The complementary brain: Unifying brain dynamics and modularity. *Trends in Cognitive Sciences*, *4*, 233-246.

Grossberg, S., and Howe, P. D. L. (2003). A laminar cortical model of stereopsis and three-dimensional surface perception. *Vision Research, 43*, 801-829.

Grossberg, S., Hwang, S., and Mingolla, E. (2002). Thalamocortical dynamics of the McCollough effect: Boundary-surface alignment through perceptual learning. *Vision Research, 42*(10), 1259-1286.

Grossberg, S., and Kelly, F. J. (1999). Neural dynamics of binocular brightness perception. *Vision Research*, *39*, 3796-3816.

Grossberg, S., and McLoughlin, N. P. (1997). Cortical dynamics of three-dimensional surface perception: Binocular and half-occluded scenic images. *Neural Networks*, *10*(9), 1583-1605.

Grossberg, S., and Mingolla, E. (1985a). Neural dynamics of form perception: Boundary completion, illusory figures, and neon color spreading. *Psychological Review*, *92*, 173-211.

Grossberg, S., and Mingolla, E. (1985b). Neural dynamics of perceptual grouping: Textures, boundaries and emergent segmentations. *Perception and Psychophysics*, *38*, 141-171.

Grossberg, S., and Mingolla, E. (1987). Neural dynamics of surface perception: Boundary webs, illuminants, and shape-from-shading. *Computer Vision, Graphics and Image processing*, *37*, 116-165.

Grossberg, S., Mingolla, E., and Ross, W. D. (1997). Visual brain and visual perception: how does the cortex do perceptual grouping? *Trends in Neurosciences*, *92*(2), 106-111.

Grossberg, S., and Raizada, R. (2000). Contrast-sensitive perceptual grouping and object-based attention in the laminar circuits of primary visual cortex. *Vision Research*, *40*(10-12), 1413-1432.

Grossberg, S., and Rudd, M. E. (1992). Cortical dynamics of visual motion perception: short-range and long-range apparent motion. *Psychological Review*, *99*(1), 78-121.

Grossberg, S., and Seitz, A. (2003). Laminar development of receptive fields, maps, and columns in visual cortex: The coordinating role of the subplate. *Cerebral Cortex*, *13*, 852-863.

Grossberg, S., and Todorović, D. (1988). Neural dynamics of 1-D and 2-D brightness perception: a unified model of classical and recent phenomena. *Perception and Psychophysics*,

43, 241-277.

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 grouping and learning. *Cerebral Cortex*, *11*(1), 37-58.

Grunewald, A., and Grossberg, S. (1998). Self-organization of binocular disparity tuning by reciprocal corticogeniculate interactions. *Journal of Cognitive Neuroscience*, *10*, 199-215.

Guy, G., and Medioni, G. (1996). Inferring global perceptual contours from local features. *International Journal of Computer Vision*, *20*(1/2), 113-133.

He, Z. J., and Nakayama, K. (1995). Visual attention to surfaces in three-dimensional space. *Proceedings of the National Academy of Sciences*, *92*, 11155-11159.

Heath, H. A., Ehrlich, D., and Orbach, J. (1963). Reversibility of the Necker Cube: II. Effects of various activating conditions. *Perceptual and Motor Skills*, *17*, 539-546.

Heeger, D. J. (1992). Normalization of cell responses in cat striate cortex. *Visual Neuroscience*, *9*(2), 181-197.

Hegde, J., and Van Essen, D. C. (2000). Selectivity of Complex Shapes in Primate Visual Area V2. *Journal of Neuroscience*, *20*, RC61 (1-6).

Hinkle, D., and Connor, C. E. (2001). Three-dimensional orientation tuning in Macaque area V4. In *Society for Neuroscience Abstr., 286.7*, San Diego, USA.

Hodgkin, A. L. (1964). *The conduction of the nervous impulse*. Springfield, IL: Charles C. Thomas.

Holliday, I. E., and Braddick, O. J. (1991). Pre-attentive detection of a target defined by

stereoscopic slant. *perception*, *20*, 355-362.

Hupé, J. M., James, A. C., Payne, B. R., Lomber, S. G., Girard, P., and Bullier, J. (1998). Cortical feedback improves discrimination between figure and background by V1, V2 and V3 neurons. *Nature*, *394*, 784-787.

Ito, M., and Gilbert, C. D. (1999). Attention modulates contextual influences in the primary visual cortex of alert monkeys. *Neuron*, *22*, 593-604.

Janssen, P., Vogels, R., and Orban, G. A. (2000). Three-Dimensional Shape Coding in Inferior Temporal Cortex. *Neuron*, *27*, 385-397.

Julesz, B. (1971). *Foundations of Cyclopean Perception*. Chicago: The University of Chicago Press.

Kandel, E. R., Schwartz, J. H., and Jessell, T. M. (2000). *Principles of neural science* (4th ed.). Chicago: 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*(4), 843-856.

Kawabata, N. (1986). Attention and depth perception. *Perception*, *15*, 563-572.

Kellman, P. J., and Shipley, T. F. (1991). A theory of visual interpolation in object perception. *Cognitive Psychology*, *23*, 141-221.

Kellman, P. J. (2003). Interpolation processes in the visual perception of objects. *Neural Networks*, *16*, 915-923.

Kelly, F., and Grossberg, S. (2000). Neural dynamics of 3-D surface perception: figure-ground

separation and lightness perception. *Perception and Psychophysics*, *62*(8), 1596-1618.

Kohonen, T. (1989). *Self-organization and associative memory, 3rd edition*. Berlin: Springer-Verlag.

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. *Cerebral Cortex*, *9*(4), 406-413.

Landy, M. S., Maloney, L. T., Johnston, E. B., and Young, M. (1995). Measurement and modeling of depth cue combination: in defense of weak fusion. *Vision Research*, *35*(3), 389-412.

Leclerc, Y. G., and Fischler, M. A. (1992). An Optimization-Based Approach to the Interpretation of Single Line Drawings as 3-D Wire Frames. *International Journal of Computer Vision*, *9*(2), 113-136.

Lee, B. (1999). Aftereffects and the representation of stereoscopic surfaces. *Perception*, *28*, 1155-1169.

Lee, S.-H., and Blake, R. (2002). V1 activity is reduced during binocular rivalry. *Journal of Vision*, *2*(9), 618-626. http://journalofvision.org/2/9/4/, DOI 10.1167/2.9.4.

Liinasuo, M., Kojo, I., Häkkinen, J., and Rovamo, J. (2000). Neon color spreading in three-dimensional illusory objects in humans. *Neuroscience Letters*, *281*, 119-122.

Maier, A., Wilke, M., Leopold, D. A., Treue, S., and Logothetis, N. K. (2001). Parallel perception of multiple visually bistable patterns. In *Society for Neuroscience Abstr., 165.15*, San Diego, USA.

Marill, T. (1991). Emulating the human interpretation of line-drawings as three-dimensional

objects. *International Journal of Computer Vision*, *6*(2), 147-161.

Markram, H., and Tsodyks, M. (1996). Redistribution of synaptic efficacy between neocortical pyramidal neurons. *Nature*, *382*, 807-810.

Marr, D., and Poggio, T. (1976). Cooperative computation of stereo disparity. *Science*, *194*, 283-287.

McLoughlin, N. P., and Grossberg, S. (1998). Cortical computation of stereo disparity. *Vision Research*, *38*(1), 91-99.

Mingolla, E., Ross, W., and Grossberg, S. (1999). A neural network for enhancing boundaries and surfaces in synthetic aperture radar images. *Neural Networks*, *12*, 499-511.

Nakayama, K., and Shimojo, S. (1992). Experiencing and perceiving visual surfaces. *Science*, *257*, 1357-1363.

Nakayama, K., and Silverman, G. H. (1986). Serial and parallel processing of visual feature conjunctions. *Nature*, *320*, 264-265.

Neumann, H., and Stiehl, H. S. (1990). Toward a computational architecture for monocular pre-attentive segmentation. In Eckmiller, R., Hartmann, G., Hauske, G. (Eds.), *Parallel processing in neural systems and computers (Proc. Int. Conf. on Parallel processing in neural systems and computers, tenth cybernetics congress of the DGK, Dusseldorf, Mar. 19-21, 1990)*, pages 491-494. North-Holland, Amsterdam.

Nguyenkim, J. D., and DeAngelis, G. C. (2001). MT neurons are selective for 3D surface orientation defined by disparity gradients. In *Society for Neuroscience Abstr., 165.8*, San Diego, USA.

Olson, S., and Grossberg, S. (1998). A neural network model for the development of simple and

complex cell receptive fields within cortical maps of orientation and ocular dominance. *Neural Networks*, *11*, 189-208.

Ozhawa, I. (1998). Mechanisms of stereoscopic vision: The disparity energy model. *Current Opinion in Neurobiology*, *8*, 509-515.

Paradiso, M. A., and Nakayama, K. (1991). Brightness perception and filling-in. *Vision Research*, *31*, 1221-1236.

Pasupathy, A., and Connor, C. E. (1999). Responses to Contour Features in Macaque area V4. *Journal of Neurophysiology*, *82*, 2490-2502.

Pasupathy, A., and Connor, C. E. (2001). Shape representation in area V4: Position-specific tuning for boundary conformation. *Journal of Neurophysiology*, *86*, 2505-2519.

Pessoa, L., and Neumann, H. (1998). Why does the brain fill-in? *Trends in Cognitive Sciences*, *2*(11), 422-424.

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. *Behavioral and Brain Sciences*, *21*(6), 723-802.

Poggio, G. F. (1972). Spatial properties of neurons in striate cortex of unanesthetized Macaque monkey. *Investigative Ophthalmology*, *11*, 369-377.

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.

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. *Visual Cognition*, *8*(3/4/5), 341-466.

Ramsden, B. M., Hung, C. P., and Roe, A. W. (2001). Real and illusory contour processing in Area V1 of the primate -- a cortical balancing act. *Cerebral Cortex, 11*, 648-665.

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.

Roelfsema, P. R., and Spekreijse, H. (1999). Correlates of gradual spread of attention over a traced curve in Macaque area V1. In *Society for Neuroscience 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.

Ryan, C., and Gillam, B. (1993). A proximity-contingent stereoscopic depth aftereffect: evidence for adaptation to disparity gradients. *Perception*, *22*, 403-418.

Sakata, H., Taira, M., Kusunoki, M., Murata, A., Tsutsui, K., Tanaka, Y., Shein, W. N., and Miyashita, Y. (1999). Neural representation of three-dimensional features of manipulation objects with stereopsis. *Experimental Brain Research*, *128*, 160-169.

Schmidt, K. E., Lowel, R., and Singer, W. (1997). The perceptual grouping criterion of colinearity is reflected by anisotropies of connections in the primary visual cortex. *European Journal of Neuroscience*, *9*, 1083-1089.

Seyama, J., Takeuchi, T., and Sato, T. (2000). Tilt dependency of slant aftereffect. *Vision Research*, *40*, 349-357.

Shevelev, I. A. (1998). Second-order features extraction in the cat visual cortex: Selective and invariant sensitivity of neurons to the shape and orientation of crosses and corners. *BioSystems*, *48*, 195-204.

Sinha, P., and Poggio, T. (1996). Role of learning in three-dimensional form perception. *Nature*, *384*(5), 460-463.

Stevens, K. A., Lees, M., and Brookes, A. (1991). Combining binocular and monocular curvature features. *Perception*, *20*, 425-440.

Sugihara, K. (1986). *Machine Interpretation of Line Drawings*. Cambridge, MA, USA: MIT Press.

Taira, M., Tsutshi, K.-I., Jiang, M., Yara, K., and Sakata, H. (2000). Parietal neurons represent surface orientation from the gradient of binocular disparity. *Journal of neurophysiology*, *83*, 3140-3146.

Thomas, O. M., Cumming, B. G., and Parker, A. J. (2002). A specialization for relative disparity in V2. *Nature Neuroscience*, *5*(5), 472-478.

Tootell, R. B. H., and Hamilton, S. L. (1989). Functional anatomy of the second visual area (V2) in the Macaque. *Journal of Neuroscience*, *9*(8), 2620-2644.

Tse, P. U. (1999). Volume completion. *Cognitive Psychology*, *39*, 37-68.

Tsodyks, M., Pawleslik, K., and Markram, H. (1998). Neural networks with dynamic synapses. *Neural Computation*, *10*(4), 821-835.

van Ee, R., van Dam, L. C. J., and Erkelens, C. J. (2002). Bi-stability in perceived slant when binocular disparity and monocular perspective specify different slant. *Journal of Vision*, *2*(9), 597-607. http://journalofvision.org/2/9/2, DOI 10.1167/2.9.2.

Viswanathan, L., and Mingolla, E. (1999). Dynamics of attention in depth: Evidence from multi-element tracking. Technical Report CAS/CNS-TR-99-010, Dept. of Cognitive and Neural Systems, Boston University, Boston, MA.

von der Heydt, R., and Peterhans, E. (1989). Mechanisms of contour perception in monkey visual cortex. I. lines of pattern discontinuity. *Journal of Neuroscience*, *9*, 1731-1748.

von der Heydt, R., Peterhans, E., and Baumgartner, G. (1984). Illusory contours and cortical neuron responses. *Science*, *224*, 1260-1262.

von der Heydt, R., Zhou, H., and Friedman, H. S. (2000). Representation of stereoscopic edges in monkey visual cortex. *Vision Research*, *40*, 1955-1967.

Waltz, D. L. (1972). Generating semantic descriptions from drawings of scenes with shadows. In Winston, P. H. (Ed.), *The Psychology of Computer Vision*, pages 19-113. McGraw-Hill Book Company - New York.

Williams, L. R., and Jacobs, D. W. (1995). Stochastic completion fields: A neural model of illusory contour shape and salience. Technical Report 95-11, NEC Research Institute, Princeton, NJ 08540, USA.

Witkin, A., and Tanenbaum, J. (1983). On the role of structure in vision. In Beck, J. (Ed.), *Human and Machine Vision*, pages 481-543. Academic Press, New York.

Zhou, H., and Friedman, H. S., von der Heydt, R. (2000). Coding of border ownership in monkey visual cortex. *Journal of Neuroscience*, *20*(17), 6594-6611.