Vision Research 106 (2015) 64–80

Contents lists available at ScienceDirect

Vision Research

journal homepage: www.elsevier.com/locate/visres

A neural model of border-ownership from kinetic occlusion

Oliver W. Layton [a,b], Arash Yazdanbakhsh [b,*]

[a] Department of Cognitive Science, Rensselaer Polytechnic Institute, 110 8th Street, Troy, NY 12180, USA
[b] Center for Computational Neuroscience and Neural Technology, Boston University, 677 Beacon Street, Boston, MA 02215, USA

ARTICLE INFO

Article history:
Received 28 May 2014
Received in revised form 29 October 2014
Available online 11 November 2014

Keywords:
Border-ownership
Figure–ground
Accretion/deletion
Kinetic edge
Motion
Inter-areal connection

ABSTRACT

Camouflaged animals that have very similar textures to their surroundings are difficult to detect when stationary. However, when an animal moves, humans readily see a figure at a different depth than the background. How do humans perceive a figure breaking camouflage, even though the texture of the figure and its background may be statistically identical in luminance? We present a model that demonstrates how the primate visual system performs figure–ground segregation in extreme cases of breaking camouflage based on motion alone. Border-ownership signals develop as an emergent property in model V2 units whose receptive fields are nearby kinetically defined borders that separate the figure and background. Model simulations support border-ownership as a general mechanism by which the visual system performs figure–ground segregation, despite whether figure–ground boundaries are defined by luminance or motion contrast. The gradient of motion- and luminance-related border-ownership signals explains the perceived depth ordering of the foreground and background surfaces. Our model predicts that V2 neurons, which are sensitive to kinetic edges, are selective to border-ownership (magnocellular B cells). A distinct population of model V2 neurons is selective to border-ownership in figures defined by luminance contrast (parvocellular B cells). B cells in model V2 receive feedback from neurons in V4 and MT with larger receptive fields to bias border-ownership signals toward the figure. We predict that neurons in V4 and MT sensitive to kinetically defined figures play a crucial role in determining whether the foreground surface accretes, deletes, or produces a shearing motion with respect to the background.

1. Introduction

Prey, such as frogs, moths, fish, and snakes, demonstrate adaptations in the appearance of their body to evade detection from predators (Osorio & Srinivasan, 1991). Successful concealment from predators is often achieved through camouflage, when the visual markings on the prey's body cause the animal to be grouped with, rather than stand out from, the surroundings (Fig. 1a). Camouflage is often only effective so long as the animal remains stationary, because predators and humans alike cannot detect stationary objects that resemble their surroundings in texture, color, and luminance. However, when a previously invisible animal breaks camouflage by sudden motion, humans rapidly perceive a figure at a different depth from the surroundings, even if the texture is identical (Fig. 1b). When a figure moves in front of a similarly textured background, it is said to produce *kinetic occlusion* (Cutting, 1997). No reliable luminance contrast exists between the figure and background; there is only the relative motion between the texture patterns separated by a kinetically defined edge (*kinetic edge*). How do humans perceive the figure at a different depth than the background (*figure–ground segregation*), despite the figure and background possessing statistically identical luminance patterns? The neural mechanisms underlying kinetic occlusion and figure–ground segregation are unclear. We present a neural model that elucidates how the visual system performs figure–ground segregation from kinetic occlusion.

When considering the motion of a figure over a similarly textured background, two scenarios naturally arise. Texture belonging to the figure may kinetically occlude or disocclude some of the texture belonging to the background (Fig. 1b, top panel). In this case, the foreground texture deletes or accretes new background texture, respectively (*accretion/deletion*). Humans are more likely to perceive the texture that moves with or is correlated with the kinetic edge over time as a surface in the foreground (Gibson, Kaplan, & Reynolds, 1969; Kaplan & Gibson, 1969; Regan & Beverley, 1984; Yonas, Craton, & Thompson, 1987). The top panel of Fig. 1b shows an example of a texture ('G') moving perpendicularly to the stationary texture on the left ('F'). Texture elements are deleted upon arriving at the kinetic edge. Humans likely perceive 'F' as the foreground surface.

* Corresponding author.
E-mail address: yazdan@bu.edu (A. Yazdanbakhsh).

http://dx.doi.org/10.1016/j.visres.2014.11.002
0042-6989/

Alternatively, the texture belonging to the figure may move, but not occlude or disocclude the background texture. This may occur when the texture belonging to the figure produces a *shearing* motion parallel to its boundary (Fig. 1b, bottom panel). Humans are more likely to perceive the faster-moving texture as the foreground surface (Regan & Beverley, 1984; Royden, Baker, & Allman, 1988). In the example depicted in the bottom panel of Fig. 1b, the texture elements on the right move vertically ('F'), and the texture on the left is stationary ('G'). Humans likely perceive 'F' as the foreground surface.

Single cell data indicate that the primate visual system has developed adaptations to detect prey breaking from camouflage. Neurons in primate visual area V2 demonstrate selectivity to oriented kinetic edges than cannot be explained by a selectivity to the motion component directions (Chen et al., 2012; Gharat & Baker, 2012; Marcar et al., 2000). In other words, the neurons appear to be sensitive to the accretion/deletion or shearing motion of the dots moving within the receptive field, rather than the mere presence of motion. von der Heydt and colleagues have shown that neurons in V2 exhibit tuning to *border-ownership* in displays consisting of a square defined by moving dots that either appear in the foreground ("object") or background ("window") (von der Heydt, Qiu, & He, 2003). For example, when the square appears in the foreground, either the dots within the square remain stationary and the exterior dots move, or the square moves and the exterior dots remain stationary. Texture accretion and/or deletion occur in both scenarios. When the kinetically defined border of the square appeared on the preferred side of the neuron's receptive field tuned to border-ownership, the response was higher than if the motion in the receptive field was the same, but the receptive field was centered on the other side of the square. Border-ownership sensitivity to the borders of figures defined by luminance contrast with the background has been well established (Friedman, Zhou, & von der Heydt, 2003; Qiu & von der Heydt, 2005; Zhou, Friedman, & von der Heydt, 2000). In our proposed model, a subpopulation of cells in V2, which is sensitive to oriented kinetic edges, also codes border-ownership. We propose that border-ownership provides a mechanism for the visual system to determine the depth ordering in scenes composed of kinetically defined figures.

Although V4 has been classically considered an area that processes static form, shape, and color, it recently has been shown to have a topographic map of motion selectivity (Tanigawa, Lu, & Roe, 2010). Not only do V4 neurons respond to kinetic edges (Mysore et al., 2006), but 10–20% of neurons demonstrate sensitivity to kinetically defined shapes (Mysore et al., 2008). The population of neurons also responds to the shapes defined by luminance contrast. This suggests that V4 is involved in figure–ground segregation, regardless of whether the figure is defined by luminance contrast or kinetic occlusion. Unlike in V4, neurons in MT do not exhibit sensitivity to kinetic edges (Marcar et al., 2000), but do exhibit tuning to moving kinetically defined shapes (Handa et al., 2008). Neurons in MT have been shown to be sensitive to long-range motion in a uniform direction (Born, 2000), but the response diminishes when multiple motion directions appear within the receptive field (Snowden et al., 1991). This indicates that MT neurons may be involved in processing kinetic occlusion, but the neurons signal characteristics about a uniformly moving surface rather than the edges.

We present a neural model that explains how the primate visual system detects a figure when it breaks camouflage from a similarly textured background. The model is the only model we are aware of that performs border-ownership assignment of kinetically defined figures. Unlike many existing approaches to figure–ground segregation from kinetic occlusion that require the explicit detection of kinetic edges, border-ownership assignment arises in the proposed model as an emergent property of dynamical feedforward and feedback interactions across areas V2, MT, and V4. Our model makes the key prediction that neurons in V2 that have been shown to be sensitive to kinetic edges are also sensitive to border-owner-

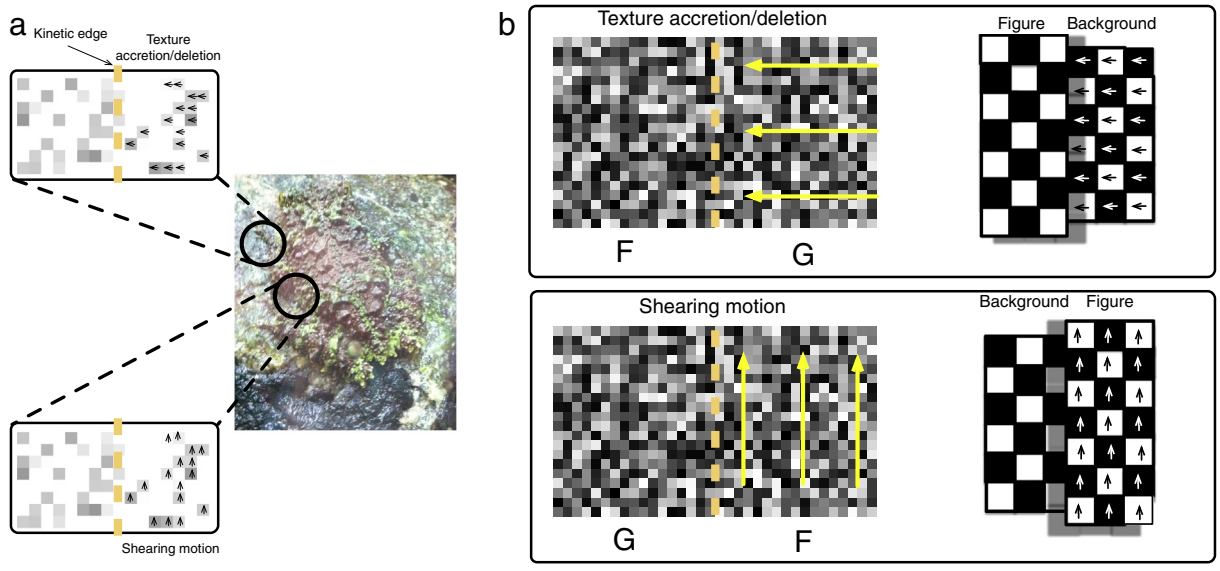

Fig. 1. Accretion/deletion of texture (top row) and shearing motion (bottom row) often indicate kinetic occlusion in nature. (a) A mossy frog (*Theloderma corticale*) is difficult to segment from the background when stationary, unless it breaks camouflage. If part of the animal moves underneath a stationary surface, the occlusion of the animal by the surface is optically specified through the deletion of texture (top panel). Movement parallel to the kinetic boundary (dashed yellow line) between the animal and the stationary background often occurs when the animal is in the foreground (bottom panel). (b) Psychophysical displays that express the same ordering in depth, as corresponding panels in (a), of two adjacent surfaces established by texture accretion/deletion (top panel; Kaplan & Gibson, 1969) and shearing motion (bottom panel; Royden, Baker, & Allman, 1988). In both displays, one surface is seen in the foreground (F) and one is seen in the background (G). The surface to the left of the kinetic edge is stationary, and the surface on the right moves as indicated by the yellow arrows. The kinetic edges remain stationary over time. A schematic depiction of the depth ordering perceived by human subjects is shown in the right panels. Top: The moving texture is deleted at the kinetic edge. Bottom: The texture on the right side moves parallel to the kinetic edge. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

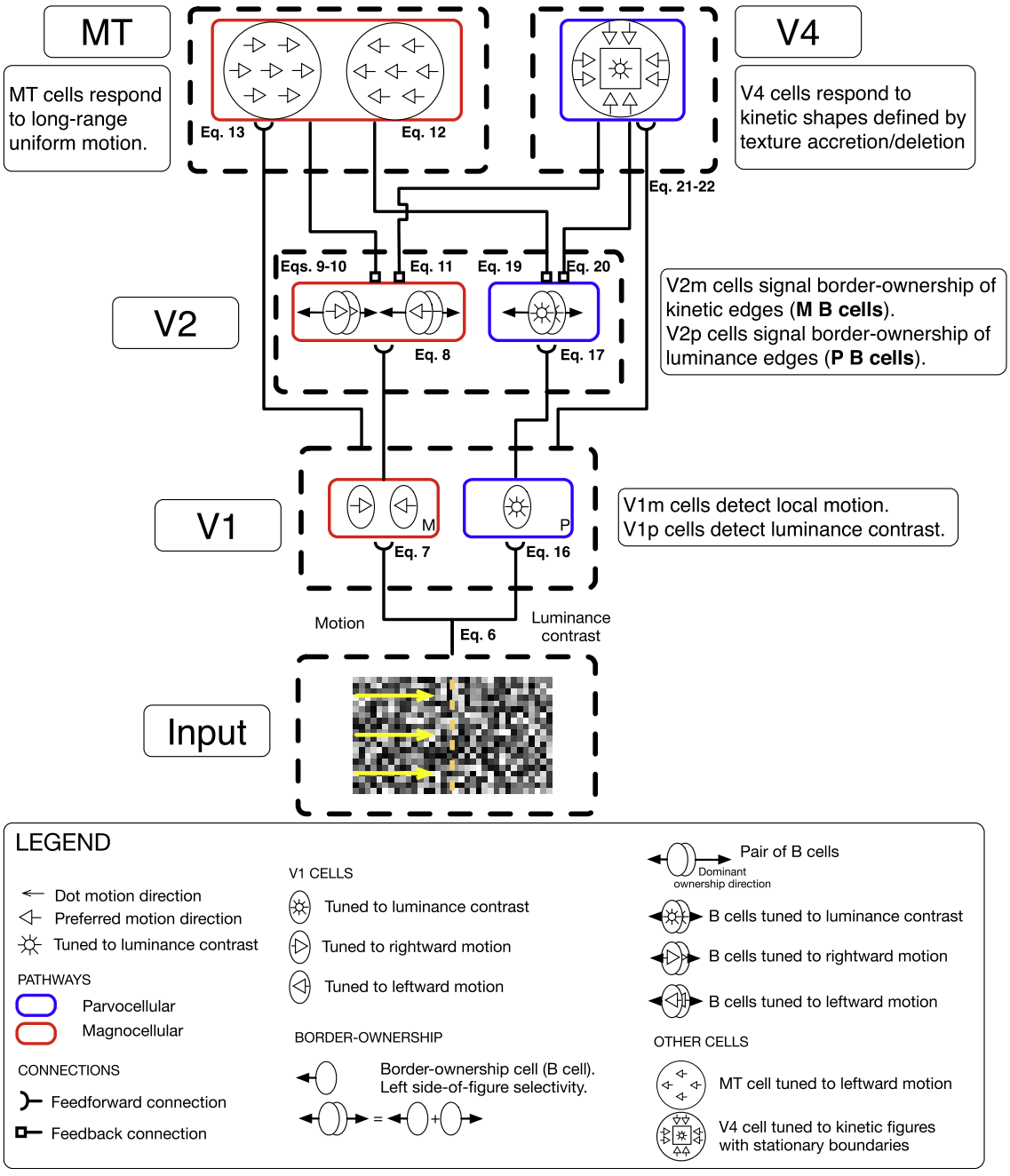

Fig. 2. An overview of the proposed model. The model contains a magnocellular pathway (M; red) that processes motion signals and a parvocellular pathway (P; blue) that processes luminance contrast signals. The model consists of three major stages. First, motion is detected in magnocellular V1 (V1 m) and luminance contrast is detected in parvocellular V1 (V1p). Second, units in MT pool over motion signals to yield selective responses to large regions of uniform motion. Units in V4 integrate luminance contrast signals and motion signals perpendicular to the receptive field center (i.e. accretion/deletion signals) to detect kinetic figures. V4 units are *driven* by luminance contrast, and *enhanced* by the accretion/deletion signals. Third, border-ownership signals arise in V2 due to feedback from units in V4 and MT. Border-ownership cells in magnocellular (MB cells) V2 are independently tuned to a border-ownership direction and motion direction and those in parvocellular (PB cells) V2 are tuned to a border-ownership direction and luminance contrast. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

ship (*magnocellular border-ownership cells*), and complement populations of neurons known to be sensitive to border-ownership of edges defined by luminance contrast (*parvocellular border-ownership cells*). V2 border-ownership sensitive units code the relative depth across the population through an activity gradient between the most active units: the closest depth is represented by border-ownership units that produce the largest activity peak, the second closest depth is represented by units that produce the second largest peak, and so forth. In the present article, we focus on challenging camouflage-breaking displays wherein the foreground and background textures are both composed of randomly distributed gray scale values (e.g. Fig. 1b).

2. Methods

The proposed model of border-ownership from kinetic occlusion is schematized in Fig. 2. The model clarifies how the magnocellular (red) and parvocellular (blue) pathways in the primate visual system interact to give rise to border-ownership signals

nearby the edges of kinetically defined figures. Both pathways are considered because camouflage-breaking figures may have either *stationary* or *moving* kinetic boundaries. Surfaces that are composed of stationary dots (e.g. F in Fig. 1a, top panel) have luminance contrast cues, and surfaces that are composed of moving dots (e.g. G in Fig. 1a, top panel) have motion cues. The magnocellular pathway (Fig. 2, left) processes motion-related signals, while the parvocellular pathway (Fig. 2, right) processes signals related to luminance contrast. The objective of the model is to clarify how border-ownership signals emerge based on the known connectivity between visual areas, despite whether the figure and background are defined by luminance contrast or motion. The parvocellular pathway is based on a simplified version of a previous model that focuses on border-ownership of figures defined by luminance contrast (Layton, Mingolla, & Yazdanbakhsh, 2012). Simplifications are possible because the present article focuses on figure–ground segregation in random dot displays (e.g. Fig. 1b).

The proposed model consists of three basic stages: detection of luminance contrast in the parvocellular pathway and motion in the magnocellular pathway (model V1), spatial pooling of motion and luminance contrast signals (model MT and V4), and cross-cue magnocellular–parvocellular pathway interactions (feedback from model MT and V4 to model V2). Border-ownership signals develop in model V2.

2.1. Model V1

Units in the parvocellular pathway of model V1 detect luminance contrast-based edges in the visual display to emulate complex cells in primary visual cortex (Hubel & Wiesel, 1968; Ringach, 2002). These units are depicted in Fig. 2 by the single ellipse with a superimposed 'sun' (Eq. (16)).

Model units in the magnocellular pathway of V1 are tuned to a particular direction of motion (Eq. (7)). These units are depicted in Fig. 2 by the single ellipse with a superimposed arrow, which indicates the preferred motion direction. For simplicity, we assume that dots move up, down, left or right at unit speed across successive frames of input. The model uses a Reichardt circuit to detect motion (Eq. (6); Egelhaaf, Borst, & Reichardt, 1989; Van Santen & Sperling, 1985). A Reichardt or correlation-based motion detection mechanism is consistent with the arrival in primate V1 of signals from the lateral geniculate nucleus (LGN), following suitable conduction delays (Nowak et al., 1995).

2.2. Model V2

Once parvocellular and magnocellular V1 units detect luminance contrast and motion, respectively, the signals propagate to distinct populations of parvocellular B cells (*PB cells*; Eq. (17)) and magnocellular B cells (*MB cells*; Eq. (8)) in V2 with comparable selectivities. For example, model V1 units tuned to leftward motion project to MB cells in V2, which are also tuned to leftward motion.

At every spatial location, there are a number of B cells that code border-ownership. There are B cells that are sensitive to the orientation of the kinetic or luminance contrast-based edge. We simulate a pair of B cells selective to border-ownership to either side θ of the oriented edge (Eq. (17)). B cells shown in Fig. 2, represented by the stack of ellipses, are sensitive to border-ownership of vertically oriented edges (left versus right). The solid black arrows to either side of the ellipses indicate the B cell's preferred side of border-ownership. For example, the B cell in Fig. 2 corresponding to the foreground ellipse selectively responds when the boundary of a figure enters the receptive field and the figure is to the left. B cells compete with one another within each border-ownership orientation axis (e.g. vertical). The larger of the two solid black arrows attached to each stack shows the dominant border-ownership direction. For the displays considered in the present work, kinetic edges are either vertical or horizontal. Therefore, we simulate pairs of B cells tuned to border-ownership along the vertical and horizontal axes.

In addition, B cells selectively respond to the type of edge in the receptive field. MB cells code border-ownership of kinetic edges and PB cells code border-ownership of luminance contrast defined edges. MB cells and PB cells compete with one another. The winning B cell codes, at a particular position, the direction of ownership and the type of edge in the receptive field. For example, the dominant MB cell in the leftmost pair of ellipses depicted in the model V2 box of Fig. 2 signals ownership to the left of a vertically oriented kinetic edge. The existence of distinct MB cells and PB cells is a major prediction of the proposed model. Note that the model proposes that border-ownership signals emerge in model V2 due to feedback from areas MT and V4 (magnocellular–parvocellular interaction; Eqs. (9)–(11) and (19), (20)). It is not clear whether the changing bottom-up signal in a small B cell receptive field is due to the accretion or deletion of dots or the movement of existing dots. Feedback from units in MT and V4 that have larger receptive fields disambiguates the local signals.

2.3. Model MT

Approximately half of MT cells *in vivo*, particularly in the input layers, respond to wide-field, coherent motion in a particular direction (Allman, Miezin, & McGuinness, 1985; Tanaka et al., 1986). MT receives most of its feedforward input from strongly direction-selective cells in V1 and V2 thick cytochrome oxidase cells, and MT cells adopt much of their motion sensitivity from their inputs (Born & Bradley, 2005). V1 and V2 neurons predominately project to layer 4 of MT, where the wide-field motion cells reside (Anderson & Martin, 2001). However, feedback to V2 may only target a distinct population of neurons that do not send feedforward projections to MT (Rockland & Knutson, 2000). In other words, the connectivity between V2 and MT does not appear to be reciprocal.

The model proposes that MT is involved in processing kinetic shapes and border-ownership assignment. Model MT units (Eq. (12)) respond to surfaces defined by moving dots (Handa et al., 2008; Marcar et al., 1995) and are functionally similar to wide-field MT cells. Motion sensitive units in model V1 project to units in MT that have the same motion direction preference (Figs. 3a and 4a). MT units spatially pool over magnocellular V1 cells, and have five times larger receptive fields. For example, V1 cells sensitive to leftward motion project to MT units tuned to leftward motion. Model MT sends feedback to a distinct population of units in V2 (Eqs. (9), (10) and (19); Rockland & Knutson, 2000). In our simulations, we implemented MT units with two different RF sizes ($r = 2$ pixels, $r = 5$ pixels), and therefore feedback to V2 comes from a spatially offset location. A summary of these feedforward connections from model V1 is shown in Fig. 4a.

The model makes a number of propositions regarding the feedback sent from MT to V2. This feedback is partly responsible for the border-ownership signals that emerge in model V2. First, feedback projections from model MT units are inhibitory and target B cells in model V2. Feedback from MT plays a role in kinetic figure–ground segregation and may be involved in the selectivity of V2 neurons to kinetic edges (Hupé et al., 2001). The influence of feedback from MT to V1 and V2 is not known. However, our model is compatible with excitatory inter-areal connections targeting inhibitory interneurons, which are known to exist in V2 (Angelucci et al., 2002; Lund, Angelucci, & Bressloff, 2003). Second, feedback from model MT may either target cells with a border-ownership preference toward or away from the center of the MT receptive field (Fig. 3b). For MB cells that have the *same* direction of motion preference as the MT cell, feedback targets the B cell with a side-of-fig-

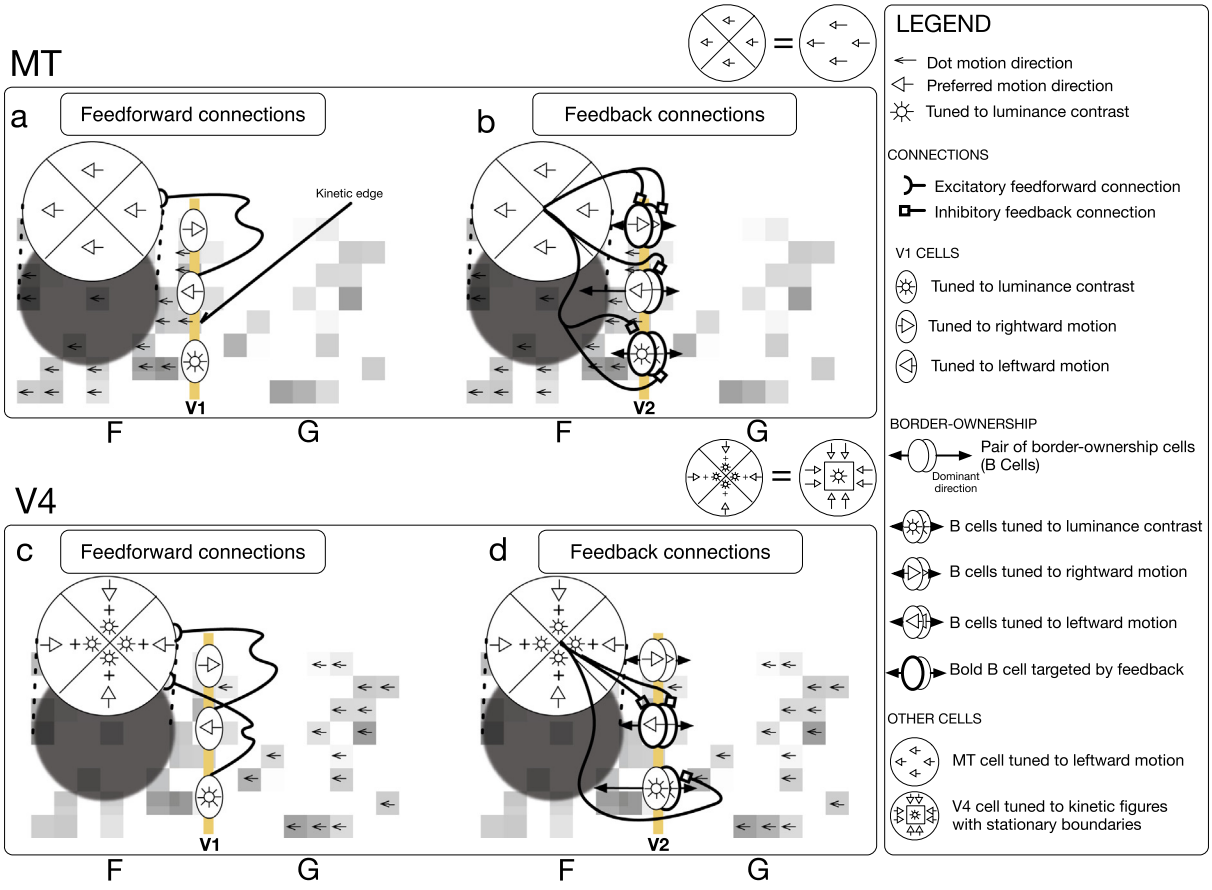

Fig. 3. Overview of the feedforward (left panels) and feedback (right panels) connectivity of model areas MT (top row) and V4 (bottom row). For the MT and V4 units, the kinetic edge (yellow dashed line) is within the right quadrant of the receptive field. Each stack of ellipses schematizes B cells that code border-ownership in opponent directions. The solid arrows attached to the left and right ellipse in the stack indicates the preferred side-of-figure of each B cell. The outlined arrows indicate the preferred motion direction. F and G label the foreground and background surfaces, respectively. (a) A MT unit that is selective to leftward motion receives feedforward input from units in V1 tuned to leftward motion. (b) The MT unit sends inhibitory feedback (outlined square connections) to distinct populations of B cells in V2 to enhance border-ownership signals toward the MT unit's receptive field center in B cells tuned to leftward motion. The border-ownership of the kinetic edge is assigned to the left surface by the MB cells tuned to leftward motion and border-ownership (the dominant border-ownership direction, indicated by the longer solid arrow) because the opponent B cell is inhibited through feedback from MT. PB cells and MB cells tuned to different motion directions are also inhibited by MT. (c) Feedforward luminance contrast signals from V1 drive the response of V4 units. V1 units that code the presence of leftward motion within the right receptive field quadrant (accretion) enhance the V4 unit's response. (d) The V4 unit sends feedback to V2 B cells to enhance the border-ownership signal toward the V4 receptive field center. The opponent PB cell with a rightward side-of-figure preference and MB cells tuned to directions of motion orthogonal to the V4 receptive field center (accretion/deletion) are inhibited. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

ure selectivity *away* from the MT cell's receptive field center (Eq. (10)). For MB cells that have the *different* direction of motion preference as the MT cell, feedback nonspecifically targets B cells that are sensitive to any side-of-figure direction (Eq. (9)). Similarly, MT units inhibit PB cells with a side-of-figure selectivity *toward* and *away* from the MT cell's receptive field center (Fig. 3b; Eq. (19)). Together, these feedback connections promote border-ownership signals to develop toward the surface on which the MT cell's receptive field is centered. By suppressing PB cells and MB cells sensitive to different motion directions than the MT cell, border-ownership signals will be enhanced in MB cells that have the same motion direction preference as the MT cell. A summary of these feedback connections from model MT is shown in Fig. 4b.

2.4. Model V4

A large proportion of dorsal V4 neurons demonstrate selectively to kinetic figures defined by the motion contrast between a surface of stationary dots and another with moving dots (Mysore et al., 2008). The neurons also responded to figures formed by opponent dot motion (texture accretion/deletion). In both cases, the kinetic boundaries remained stationary over time. Neural responses were much the same, despite whether the figure was defined by luminance contrast or moving dots, which indicates a cue-invariant figure response (Mysore et al., 2006). V4 cells receive their main feedforward input from and send feedback to V1 and V2 (Ungerleider et al., 2008). The input from V1 comes from cells that represent the foveal portion of the visual field in both cytochrome oxidase blob and interblob regions (Nakamura et al., 1993).

In the proposed model, V4 units selectively respond to kinetically defined figures with stationary boundaries, similar to dorsal V4 neurons. The model proposes that V4 neurons are involved in a circuit that assigns border-ownership of kinetic edges. Units in the parvocellular pathway in V1, which detect edges defined by luminance contrast, project to V4 (Fig. 3c; Eq. (21)). V1 units in the magnocellular pathway, which are sensitive to motion, also project to V4. The model proposes that V4 neurons demonstrate sensitivity to accretion and deletion of texture, which typically arises in nature when a figure breaks camouflage. This sensitivity is implemented in the model by restricting inputs to V4 from magnocellular V1 cells that have preferred direction of motion that runs toward or away from the center of the V4 receptive field.

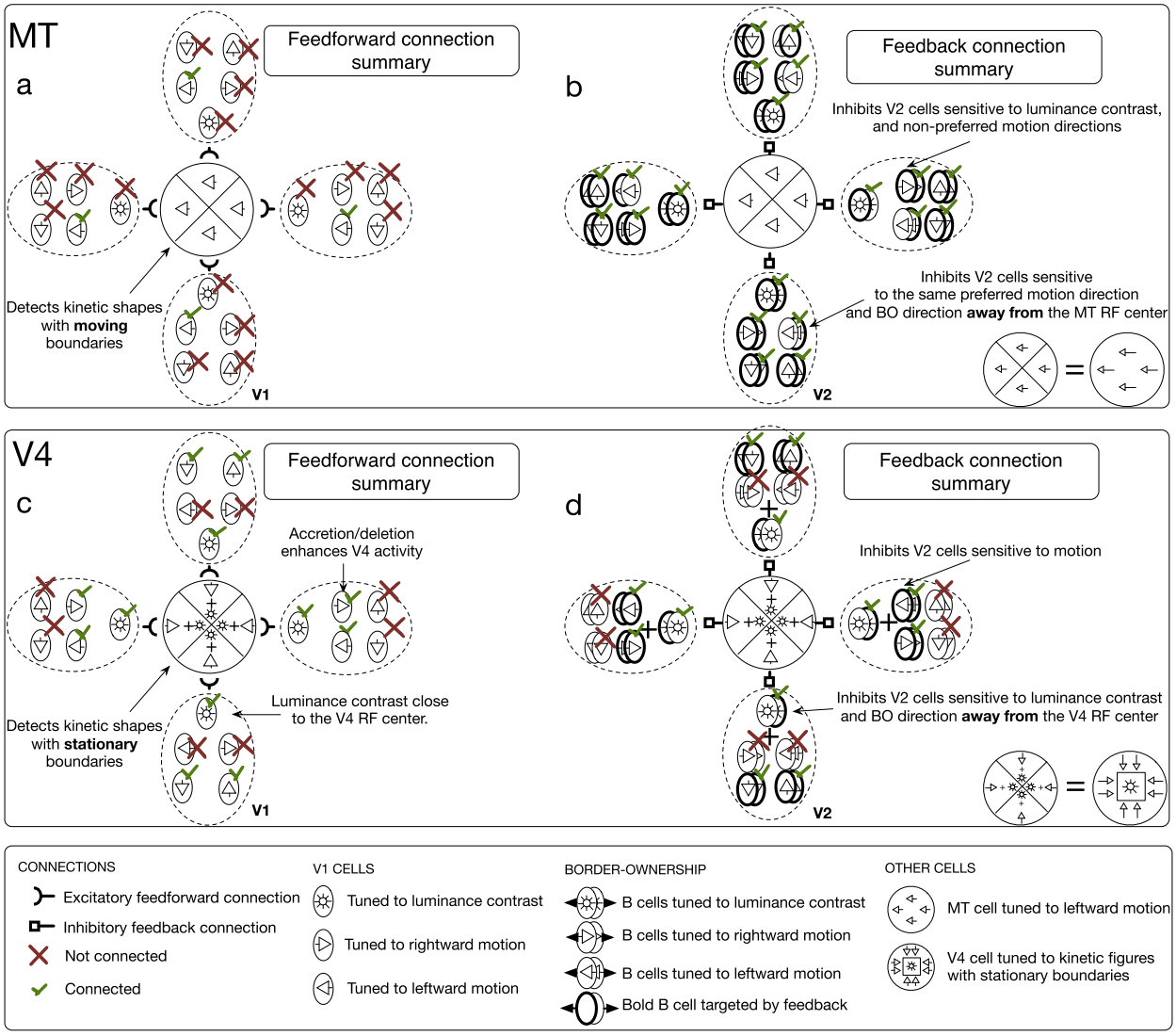

Fig. 4. Complete feedforward (left column) and feedback (right column) connectivity of model areas MT (top row), V4 (bottom row). The green checkmarks attached to a V1 or V2 unit indicate that the feedforward or feedback connection to or from V4 or MT exists, and the red 'x' indicates the connection does not exist. V1 or V2 units surrounded by a dashed elliptical contour have the same receptive field position. (a) The feedforward connectivity between V1 to MT. Units in model MT receive inputs from V1 tuned to the same direction of motion in all quadrants of the receptive field (left direction selective unit shown). (b) The feedback connections from MT to V2 B cells. The MT unit inhibits MB cells tuned to the same motion direction that have a side-of-figure selectivity away from the MT receptive field center. PB cells and MB cells tuned to different motion directions are inhibited. (c) The feedforward connectivity between V1 and V4. Luminance contrast signals *drive* the V4 unit response. Each quadrant of the V4 unit receptive field receives input from V1 cells tuned to motion perpendicular to the V4 receptive field center (accretion/deletion). (d) The feedback connections from V4 to V2 B cells. The V4 unit inhibits PB cells that have a side-of-figure selectivity away from the V4 receptive field center. MB cells tuned to perpendicular motion directions in each quadrant of the V4 receptive field are inhibited. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

In other words, V4 units only receive feedforward motion inputs that signal motion perpendicular to a stationary kinetic boundary (Fig. 3c; Eq. (22)). The response of V4 units is *driven* by inputs from parvocellular V1 (luminance contrast) and *enhanced* by the presence of texture accretion/deletion via inputs from magnocellular V1 (motion). As in model MT, we simulated units with two different RF sizes ($r = 2$ pixels, $r = 5$ pixels). A summary of feedforward connections to V4 is shown in Fig. 4c.

Similar to model MT, feedback projections from model V4 units are inhibitory and target B cells in model V2 with a border-ownership preference toward or away from the center of the V4 receptive field. Feedback targets PB cells with a side-of-figure selectivity *away* from the V4 unit's receptive field center (Fig. 3d; Eq. (20)). This enhances border-ownership signals when the V4 unit is centered on a figure defined by a static textured surface and surrounded by texture accretion/deletion. Feedback also targets the population of MB cells that has a preferred direction of motion that runs toward or away from the center of the V4 receptive field (Eq. (11)). The MB cell with a side-of-figure preference *toward* and *away* from the V4 receptive field center is inhibited. This promotes border-ownership signals to develop toward the center of the kinetic figure defined by stationary kinetic edges. A summary of feedback connections from model V4 to V2 is shown in Fig. 4d.

3. Results

To quantify the strength of border-ownership signals, we compute (Eq. (1)) the vectorial modulation index (VMI) (Zhou, Friedman, & von der Heydt, 2000). Values range from −1 to 1. Negative values indicate border-ownership assignment to the left, positive values indicate border-ownership to the right, and zero means a lack of border-ownership modulation between units with

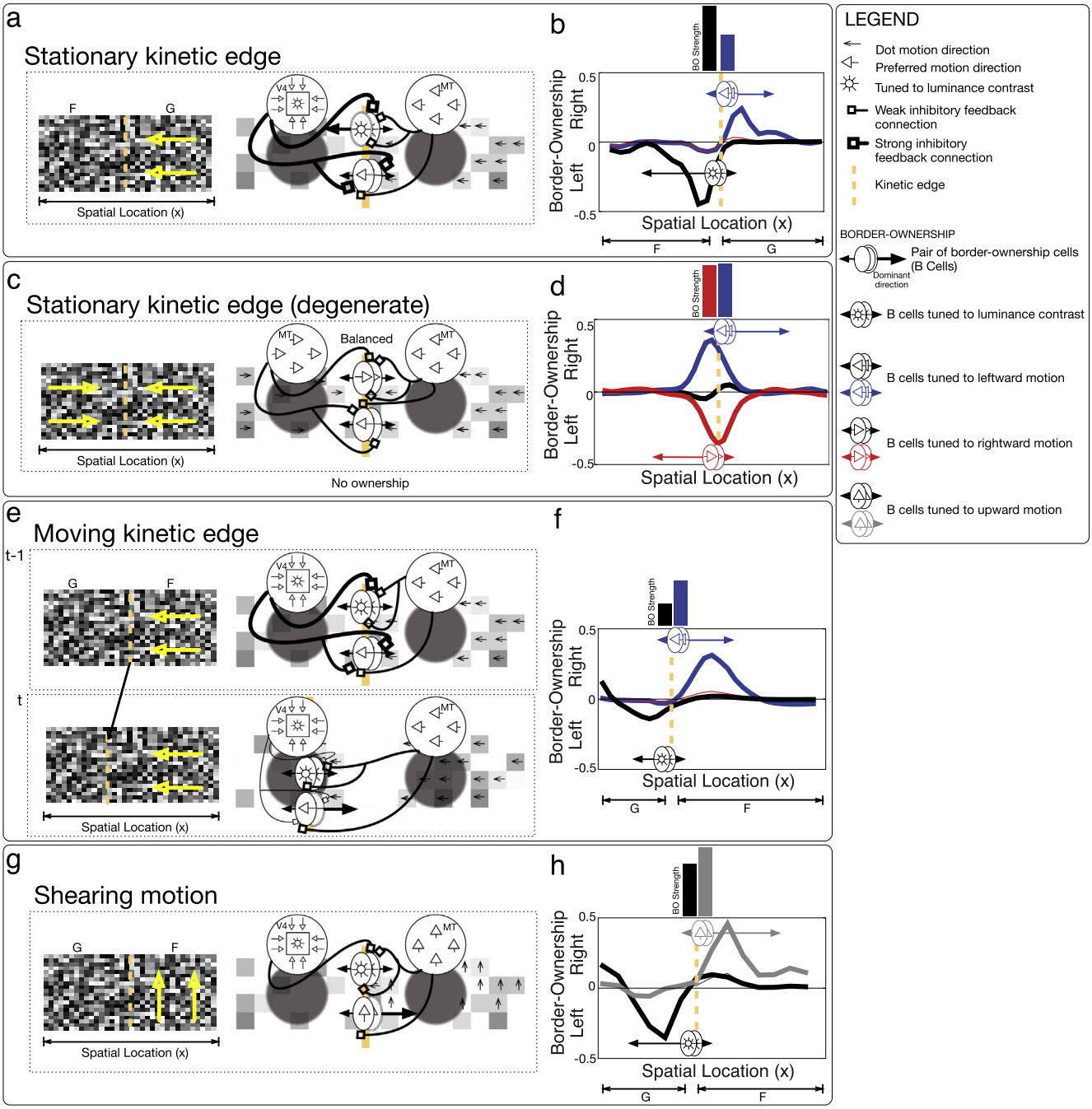

Fig. 5. Simulation results for displays containing a stationary kinetic edge (a and b), a stationary kinetic edge with no perceived depth ordering (c and d), a moving kinetic edge (e and f), and shearing motion (g and h). The yellow arrows indicate the direction of motion of textures. The surface perceived in the foreground by human subjects is labeled 'F', and the background surface is labeled 'G'. The left column depicts each visual display, the middle column depicts the mechanisms in the model that result in correct border-ownership assignment, and the right column plots the simulation results of the model. The border-ownership signals for MB (colored curves) and PB cells (black curves) are obtained by averaging the VMI obtained for one-dimensional cross-sections (x) of the visual display. The relative magnitude of border-ownership signals is indicated by lengths of colored bars labeled 'BO strength' and arrows attached to schematic B cells superimposed on the kinetic edge (yellow dashed line). The MB cell or PB cell with the largest signal assigns border-ownership of the kinetic edge to the surface in the direction of the preferred side-of-figure. For example, the border-ownership signal produced in (b) by PB cells (black) is greater than that of the MB cell population tuned to leftward motion. Therefore, border-ownership of the kinetic edge is assigned to the left surface, coded by the PB cells. Border-ownership signals produced by PB cell and MB cell populations match ordinal depth percepts of human observers. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

side-of-figure preferences to either side of a kinetic edge. VMI values reported in the text are averaged over 20 runs of the model.

$$\boldsymbol{VMI} = \frac{\boldsymbol{b}_{right} - \boldsymbol{b}_{left}}{\boldsymbol{b}_{right} + \boldsymbol{b}_{left}} \quad (1)$$

The model contains B cells in the magnocellular (MB cells) and parvocellular (PB cells) pathways. MB cells are sensitive to a border-ownership direction and a direction of motion. *The sensitivity to these attributes is independent*. For example, a MB cell may be selective to leftward motion and leftward border-ownership, and another may be selective to leftward motion and rightward border-ownership. PB cells are sensitive to a border-ownership direction and an orientation of an edge. Because kinetic random dot displays contain motion contrast, multiple antagonistic MB and PB cell populations are simultaneously active nearby a kinetic edge. The populations compete to assign border-ownership of a

kinetic edge. The border-ownership of a kinetic edge is determined by selecting the side-of-figure preference of the maximally active B cell.

Here we compare border-ownership signals produced by model B cells, human figure–ground perception, and neurophysiological properties of B cells.

3.1. Stationary kinetic edge

Fig. 5a depicts a visual display containing deletion of a moving texture at a stationary boundary, and Fig. 5b shows the results of a model simulation. The model mechanisms that give rise to the border-ownership signals are schematically depicted in Fig. 5a. Human observers report seeing the left surface as in front of the one on the right (Kaplan & Gibson, 1969). As shown in Fig. 5b, both PB cells (black) and MB cells tuned to leftward motion (blue) generate border-ownership modulation near the kinetic edge. The mean border-ownership signals across horizontal cross-sections of the display, perpendicular to the vertical kinetic edge, are plotted. The border-ownership signal produced by the PB cells (VMI = −0.47, black curve) is greater than that produced by the MB cells sensitive to leftward motion (VMI = 0.26, blue curve). This means that the PB cells coding leftward border-ownership (black arrows) are more active than those coding rightward border-ownership (blue arrows). The difference in activity (modulation) is greater among the PB cells than among the MB cells. Therefore, the model assigns border-ownership of the kinetic edge to the left stationary surface, which is coded by the PB cells with a leftward side-of-figure preference. This is consistent with the percepts of human observers that the stationary surface is in the foreground. The relative magnitude of the peak modulations is indicated by the height of the rectangles on the top of Fig. 5b, and also by the length of the arrows attached to the schematized B-cells superimposed on the kinetic edge.

The right panel of Fig. 5a shows the model mechanisms responsible for the border-ownership assignment. Above the left stationary surface, a prototypical V4 unit is shown, and above the right moving surface, a MT unit is shown. The kinetic edge is inside the receptive fields of both units. Recall that when texture accretion/deletion occurs within a V4 unit's receptive field, the response is enhanced. Therefore, the feedback signal V4 sends to V2 border-ownership cells is stronger (thick connections) than that sent from MT cells (thin connections). The PB cell that has a left side-of-figure preference receives the least inhibition compared to its opponent with a right side-of-figure preference, and therefore yields the highest activity (Fig. 5a).

The VMI values for the display shown in Fig. 5a, and those for other displays shown in Fig. 5, are on average <0.5. The VMI measures the strength of border-ownership signals, by comparing the difference in firing rates between units coding border-ownership in opposing directions. Model VMI values agree with response ratios computed on *in vivo* V2 border-ownership cells (von der Heydt, Qiu, & He, 2003; Zhou, Friedman, & von der Heydt, 2000).

When a kinetic edge is stationary, and deletion occurs at both sides (Fig. 5c), humans tend not to perceive a consistent ordering of the surfaces in depth (Kaplan & Gibson, 1969). In other words, the kinetic edge does not belong to any one surface. Model simulations show that the peak border-ownership signal produced in the MB cell population tuned to leftward motion (VMI = 0.36, blue curve) is similar to that produced by the MB cell population tuned to rightward motion (VMI = −0.39, red curve) near the kinetic edge. The similarity in the modulation across the MB cell populations suggests that the model does not reliably assign border-ownership of the kinetic edge in either direction.

As depicted in the right panel of Fig. 5c, the presence of motion in both surfaces recruits MT cells tuned to corresponding motion directions. MT units tuned to rightward motion on the left side of kinetic edge and units tuned to leftward motion on the right side are equally active and send symmetric feedback to V2 MB cells. Within each MB cell population, one B cell receives twice the inhibition that its opponent receives (two versus one inhibitory input). Because in MB cell populations this inhibition is balanced, no differences in border-ownership modulation emerge *across* the MB cell populations.

3.2. Moving kinetic edge

Fig. 5f shows a model simulation of a display containing deletion of a static texture at a moving kinetic boundary (Fig. 5e). Human observers report seeing the moving surface (right) in front of the surface whose texture is deleted (left) (Kaplan & Gibson, 1969; Regan & Beverley, 1984). The kinetic edge belongs to the moving surface. As shown in Fig. 5f, MB cells tuned to leftward motion and PB cells both demonstrate nontrivial border-ownership modulation. The magnitude of the peak modulation in the MB cell population (VMI = 0.34, blue curve) is greater than that of the PB cell population (VMI = −0.16, black curve). The model assigns border-ownership of the kinetic edge to the right surface, coded by the MB cells with right side-of-figure selectivities (blue arrow).

The top panel of Fig. 5e shows the state of the model at time $t-1$, and the bottom panel shows the state at a later time t. The state shown at $t-1$ represents what would occur if the kinetic edge does not move, similar to the scenario in Fig. 5a. Due to the presence of a static texture on the left, V4 units are most active in the left surface. Because the texture in the right surface is moving, MT units tuned to leftward motion are most active there. Because the V4 cell detects the presence of deletion within the receptive field, its response is enhanced, and it consequently sends stronger feedback to V2 units than the MT unit on the other side of the kinetic edge. However, the kinetic edge sweeps through the V4 receptive field between times $t-1$ and t, replacing stationary texture with leftward moving dots within the receptive field, which decreases the V4 activity. The luminance cues in the stationary texture drive the V4 unit, and without them, the activity of the V4 unit decreases, which results in weaker feedback to V2. Between $t-1$ and t the MT unit had leftward motion within the receptive field, so the feedback signal to V2 remains equally strong. At time t, the MB cell tuned to leftward motion that has a right side-of-figure preference receives the least inhibition relative to its opponent coding the other border-ownership direction. Therefore, border-ownership is assigned to the right by the MB cells.

3.3. Shearing motion

When the texture in one surface moves parallel to the orientation of the kinetic edge (shearing motion), humans tend to report seeing the surface associated with the faster motion in the foreground (Royden, Baker, & Allman, 1988). For a display containing a static texture on the left and vertical motion to the right of a vertically oriented kinetic edge (Fig. 5g), human subjects tend to perceive the right surface with the moving texture in the foreground.

Fig. 5h shows the simulation results of the display with shearing motion. MB cells sensitive to upward motion and PB cells sensitive to the stationary texture both produce border-ownership signals. The magnitude of the peak modulation in the MB cells tuned to upward motion (VMI = 0.49, gray curve) is greater than that of the PB cells (VMI = −0.39, black curve). The model selects the MB cells tuned to upward motion with side-of-figure preferences as the winning population, and border-ownership is assigned to the right (gray arrow). This result is a consequence of the connectivity predicted by the model (Fig. 4), which remains

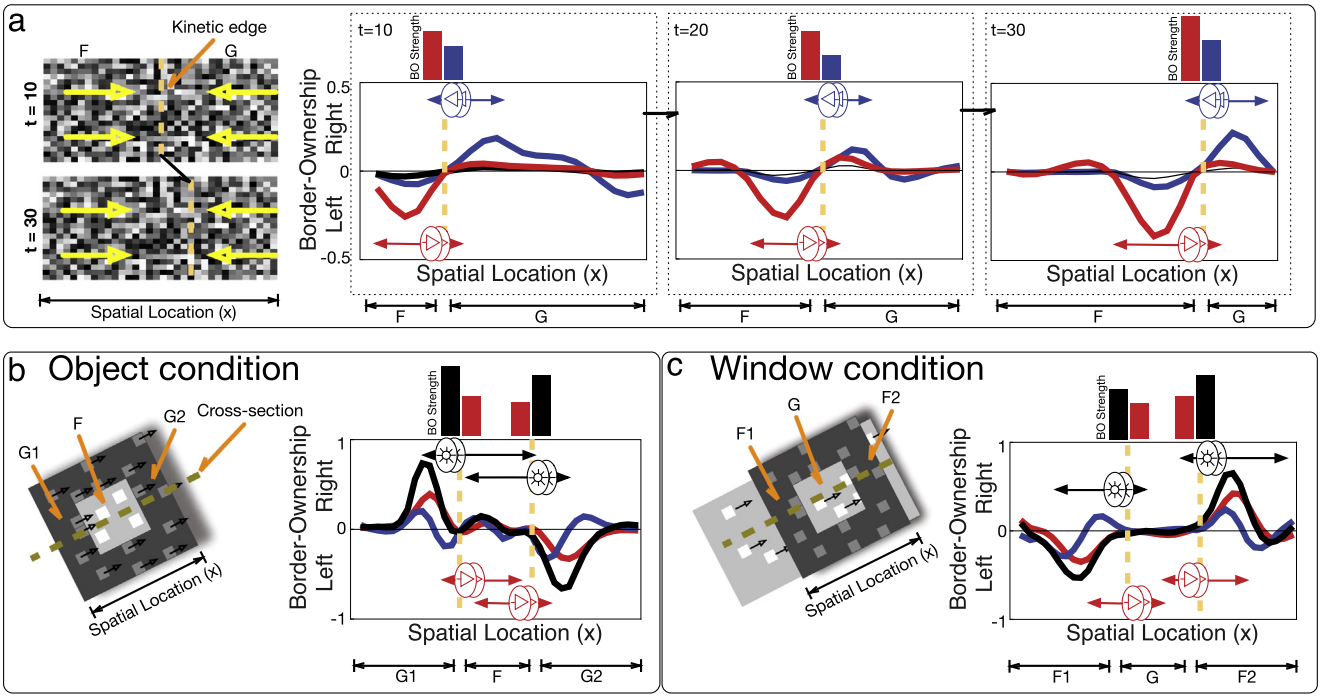

Fig. 6. (a) Simulation results at different times ($t = 10$, $t = 20$, $t = 30$) for a display containing a moving kinetic edge. The ordinal relationship between the border-ownership signals produced by MB cells tuned to rightward motion (red) and MB cells tuned to leftward motion (blue) is preserved over time. The border-ownership signals that assign the kinetic edge to the left surface closely track the kinetic edge over time. (b) Simulation of the object condition, wherein a centrally located square region is in the foreground (F) and the surround region ('moat', G1 and G2) is in the background. The left and right sides of the moat as they relate to the simulation plot are indicated by G1 and G2, respectively. (c) Simulation of the window condition, wherein a centrally located square region is in the background (G) and moat is in the foreground (F1 and F2). Border-ownership signals produced by PB cell and MB cell populations match ordinal depth percepts of human observers. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

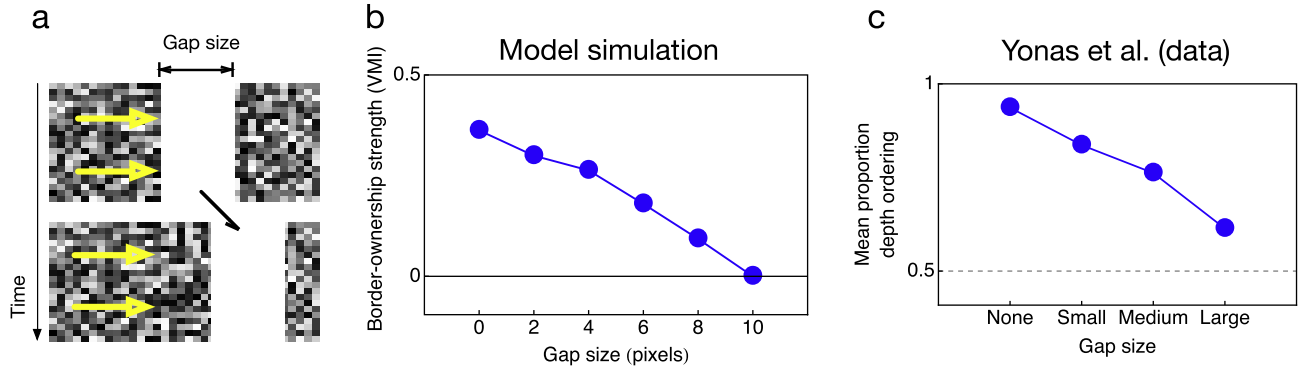

Fig. 7. A comparison of border-ownership signals produced by the model (b) and the proportion of time human subjects in the study of Yonas et al. indicated the existence of a depth ordering in the visual display (c). The visual displays were similar to those shown in Fig. 5e (moving kinetic edge), except a white gap was placed in front of the kinetic edge such that texture accretion/deletion did not occur. A border-ownership strength of zero indicates that the model cannot reliably assign border-ownership. A mean proportion depth ordering of 0.5 indicates chance performance.

the same throughout all simulations. Note that the difference in the border-ownership signal strength (VMI) in the case of shearing motion is smaller than obtained in the other cases above, except for the degenerate case (Fig. 5c) when border-ownership is not reliably assigned by the model (Fig. 5d). The smaller difference in VMI indicates a weaker border-ownership effect in the shearing motion case, which is consistent with human psychophysical data (Royden, Baker, & Allman, 1988).

The model yields a different border-ownership assignment in the case of shearing motion compared to that of the stationary kinetic edge (Fig. 5a) due to the dynamics of V4 units. As depicted in Fig. 5g, V4 units send feedback to PB cells, but not to MB cells. Recall that V4 units are driven by luminance inputs, and the response is enhanced when accretion/deletion occurs within the receptive field. Shearing motion does not result in accretion/deletion, so no enhancement of the V4 unit's response occurs. A V4 unit only sends feedback to V2 units tuned to luminance and motion attributes from which it received input. Since units in V1 tuned to vertical motion do not project to the V4 unit, V2 MB cells tuned to vertical motion are not inhibited due to feedback from V4. The MB cells with a right side-of-figure selectivity receive no inhibition and therefore are most active. Border-ownership is assigned to the moving surface.

3.4. Position of the border-ownership signal

When a kinetic edge moves, the position of the peak border-ownership modulation in the model dynamically tracks the

position of the edge over time. To demonstrate this property, we plotted (Fig. 6a) the model activity at three points of time ($t = 10$, $t = 20$, $t = 30$). In this simulation, a kinetic edge moves from left to right. Texture in the left surface (F) moves to the right and accretes, while texture in the right surface (G) moves to the left and is deleted. As depicted on the three rightmost panels side of Fig. 6a, MB cell populations tuned to leftward and rightward motion show border-ownership modulation nearby the kinetic edge. When $t = 10$, the VMI of the MB cells tuned to rightward motion is −0.28 (red curve), which is greater in absolute magnitude than that of the population tuned to leftward motion (VMI = 0.21, blue curve). Over time, the ordinal relationship between the peak modulation across the populations (red peak versus blue peak) remains constant, which results in border-ownership assignment toward F, but the absolute magnitudes fluctuate as the signals evolve. By $t = 20$ relative border-ownership modulations between the two populations (i.e. absolute differences between the VMI) stabilized to approximately 0.15.

3.5. Object and window conditions

We tested the proposed model on visual displays that contain richer kinetic figures than the kinetic random dot displays considered so far. In the object (Fig. 6b) and window conditions (Fig. 6c), a centrally located square surface is surrounded by a 'moat.' Dots in either the square or the moat move, and are biased in their mean luminance (von der Heydt, Qiu, & He, 2003). In the object condition, the square region appears in the foreground (F). In the window condition, the square region is in the background (G) and the moat is in the foreground. In simulations, the mean luminance of the square was greater than that of the moat. The object and window displays contain a combination of strong luminance contrast and motion cues to perform figure–ground segregation. The boundary between the square and moat is composed of four kinetic and luminance contrast based edges. As in the above simulations, we report the mean border-ownership modulation averaged across cross-sections of the display, indicated by the orange dashed lines.

In the simulation of the object condition (Fig. 6b), PB cells (black curve) and MB cells tuned to rightward motion (red curve) demonstrate the greatest border-ownership modulation. The absolute peak magnitudes were generally greater in both the object and window conditions, compared to previous simulations of two adjacent surfaces. PB cells garnered the greatest modulation (VMI = 0.77 left kinetic edge, −0.67 right kinetic edge), followed by the MB cells tuned to rightward motion (VMI = 0.44 left kinetic edge, −0.37 right kinetic edge). The positive sign of the VMI around the left kinetic edge and the negative sign of the VMI around the right kinetic edge indicate *inward* border-ownership signals toward the interior of the square within each B cell population. The PB cells demonstrate the greatest border-ownership modulation at the kinetic edges, and therefore are selected as the winners by the model. As indicated by the direction and lengths of the arrows attached to the schematic PB and MB cells in Fig. 6b, border-ownership of the kinetic edges, which also have luminance contrast, is assigned toward the interior of the square.

In the window condition simulation (Fig. 6c), PB cells also produce the greatest border-ownership modulation (VMI = 0.56 left kinetic edge, 0.71 right kinetic edge), followed by MB cells tuned to rightward motion (VMI = 0.4 left kinetic edge, 0.47 right kinetic edge). However, the sign of the border-ownership signals at each kinetic edge switches compared to the object condition. In the object condition, the kinetic edges are assigned *inwardly* to the square, but in the window condition, the kinetic edges are assigned *outwardly* to the moat. There is a decrease the relative VMI magnitudes garnered in the PB and MB cell populations in the window case compared to the object case, consistent with the physiological findings (von der Heydt, Qiu, & He, 2003).

The increase in VMI in simulations of the object condition compared to the window condition occurred due to greater enhancement of V4 responses. In the object condition, V4 units centered on the stationary square have accretion/deletion at two kinetic edges within the RF. In the window condition, the active V4 units are centered on the moat, but only one kinetic edge enters the RF, resulting in weaker activation. Border-ownership modulation due to feedback from V4 is weaker in the window condition compared to the object condition, yielding a smaller VMI.

3.6. Gap condition

Fig. 7 compares human figure–ground judgments (Fig. 7c) with border-ownership signals (VMI) produced by the model (Fig. 7b) when a uniform white gap is introduced between two random dot surfaces (Fig. 7a). The display is similar to the one with the moving kinetic edge shown in Fig. 5e, except texture accretion/deletion no longer occurs due to the gap. The human data in Fig. 7c are replotted from a study performed by Yonas, Craton, and Thompson (1987). In the study, a small, medium, and large-sized gap was introduced between the surfaces. Subjects were asked whether they perceived a depth ordering in the displays. Fig. 7c shows the percent of the time subjects indicated that a depth ordering was present. 100% indicates a depth ordering always was perceived and 50% indicates that subjects did not reliably see a depth ordering. When there was accretion/deletion (no gap), human subjects saw a depth ordering between the left and right surfaces 98.8% of the time. When a small, medium, and large sized gap was introduced, humans saw a depth ordering 87.2%, 78.8%, and 62.5% of the time, respectively. As the gap size increased, the likelihood of reporting a depth ordering approached chance (50%).

For small gap sizes, border-ownership signals are similar in magnitude to those obtained without a gap (Fig. 5f). When the gap size becomes large, the border-ownership modulation drops to 0 because the gap becomes comparable in size to the largest receptive field size of V4 and MT units used in the simulation. A zero border-ownership modulation is analogous to chance judgments of the presence of a depth ordering in the Yonas experiment because reliable border-ownership signals cannot be established. The decrease seen in both curves as a function of gap size suggests that units with limited receptive field sizes are important for kinetic figure–ground segregation.

4. Discussion

We have proposed a neural model of border-ownership assignment from kinetic occlusion in the primate visual system. The

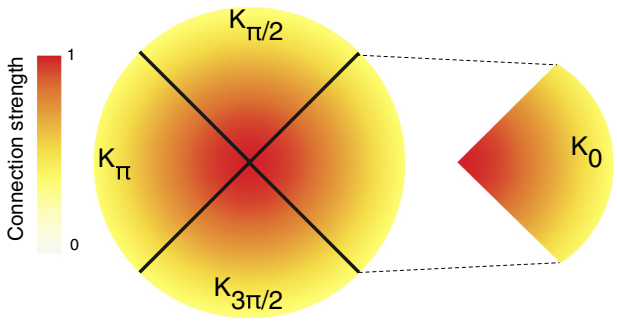

Fig. 8. Quadrants $\boldsymbol{K}_{\theta}^{s}$ of model V4 and MT Gaussian feedforward and feedback kernels.

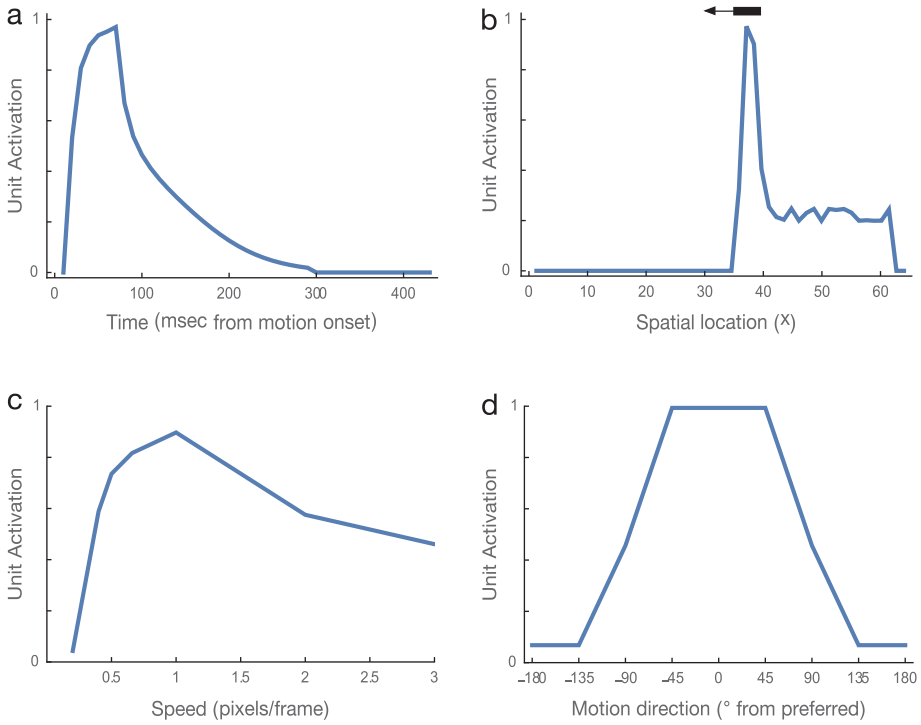

Fig. 9. Motion response properties of model V1 units (Eq. (7)) when a vertical leftward moving bar is presented. (a) The temporal response of a V1 unit tuned to the motion direction of the bar. (b) One-dimensional horizontal cross section of the V1 population response tuned to leftward motion. A horizontal cross section of the moving bar is depicted above the plot. (c) Speed tuning of a V1 unit that prefers 1 px/frame motion. (d) The tuning to motion direction of a V1 unit, where 0° indicates the preferred motion direction of the unit. Model V2 and MT units inherit their speed and direction tuning from V1.

model makes the key prediction that the border-ownership assignment in kinetically defined figures is mediated by two distinct populations of border-ownership cells (B cells): those in the magnocellular pathway that are also sensitive to motion (MB cells), and those in the parvocellular path that are also sensitive to luminance contrast (PB cells). Neurons in V1, V2, and V4, which are known to generate border-ownership signals in the context of figures defined by luminance contrast, correspond to PB cells in the model (Friedman, Zhou, & von der Heydt, 2003; Zhou, Friedman, & von der Heydt, 2000). A cornerstone property of B cells is that they are tuned to both border-ownership and oriented luminance contrast, similar to V1 complex and simple cells. If the side-of-figure selectivity is not tested, a B cell may mimic the properties of a cell tuned to oriented luminance contrast. Single cells demonstrate selectivity to oriented kinetic edges in the same areas of primate cortex that contain B cells (Chen et al., 2012; Gharat & Baker, 2012; Marcar et al., 2000; Mysore et al., 2006). We hypothesize that, similar to the relationship between oriented luminance contrast edges and border-ownership, a subpopulation of neurons tuned to kinetic edges also exhibit sensitivity to border-ownership (MB cells). The existence of B cells tuned to motion has been confirmed (von der Heydt, Qiu, & He, 2003), however additional research is needed to better understand their role in kinetic figure–ground segregation and their relationship to kinetic edge cells.

4.1. Border-ownership in context

Our model makes the specific prediction that the RFs of MB and PB cells need not be centered on the borders of a kinetically defined figure to signal border-ownership. Indeed, simulations show that the largest border-ownership signals emerge in units whose RFs are centered in the nearby vicinity of a kinetic edge. In the model, the amount of spatial offset between the border-ownership signal and the kinetic edge is controlled by the range of RF sizes in areas MT and V4, which send feedback to the MB and PB cells in V2. We simulated MT and V4 units with two and five pixel wide RFs, which gave rise to large spatial offsets between the kinetic edge and border-ownership signals. Such large offsets are unlikely to exist *in vivo*, but the model predicts nevertheless that some spatial offset is present. We selected the MT and V4 RF sizes (2 and 5 px wide) by considering that these areas have neurons with several times larger RFs than neurons in V1 (~1 px wide RFs in model V1). However, the mix and distribution of V4 and MT unit RF sizes will influence the spatial offset, as proposed by the model. For example, incorporating a greater proportion of small RF units will decrease the spatial offset. Data on the distribution of V4 and MT RF sizes that project to V2 would clarify the extent of the spatial offset, if the model prediction is correct. We are not aware of data that demonstrate how local border-ownership signals are to an edge in the context of camouflage-breaking displays. For figures defined by luminance contrast, some B cells only respond when the edge is centered in the RF (e.g. Fig. 11; Zhou, Friedman, & von der Heydt, 2000), while others are more tolerant of spatial displacements of the edge (e.g. Fig. 12; Zhou, Friedman, & von der Heydt, 2000). In either case, border-ownership signals are spread over space by at least 1° on either side of the edge. It is unclear how the proximity of a kinetic edge to the RF influences the border-ownership signal or even the response in general.

It is important to emphasize that border-ownership signals in model V2 MB and PB cells arise due to inter-areal, rather than local, competitive interactions with V4 and MT cells (Eqs. (8) and (17)). The border-ownership signals do not depend on local competition between V2 units at the same visuotopic position, but rather are 'linked' by the inter-areal connectivity. Different distributions of V4 and MT RF sizes will spatially shrink or elongate the visuotopic offset, but the pattern of border-ownership peaks will remain the same.

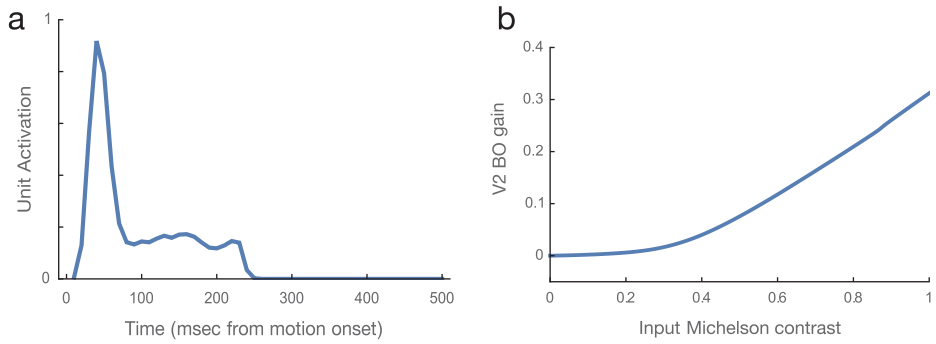

Fig. 10. Motion response properties of model V2 units (Eq. (8)). (a) The temporal response of a V2 unit tuned to the direction of a moving bar. The response is sharpened compared to that of V1 units (Fig. 9a) due to feedback from model MT (Eqs. (9) and (10)) and V4 (Eq. (11)). Note that the feedback is modulatory and does not alter the basic tuning properties of V2 units. (b) The relative increase (gain) in the largest border-ownership signal peak compared to the second largest peak as a function of the Michelson contrast between pixel values in the stationary kinetic edge display (Fig. 5a). The gain is computed according to $a - b/a + b$, where a corresponds to the peak border-ownership signal produced by the dominant V2 subpopulation (e.g. black peak in Fig. 5b) and b corresponds to the peak border-ownership signal produced by the second most active V2 subpopulation (e.g. blue peak in Fig. 5b). The magnitudes of the peaks are well differentiated when there is moderate or high contrast in the input.

Border-ownership signals developed in MB and PB cells on either side of the kinetic edge, along the axis perpendicular to the edge orientation. Recall that in the model border-ownership sensitivity is independent of the direction of motion, so the position of border-ownership peaks along the perpendicular axis arises despite whether the kinetic edge is defined by parallel, perpendicular or otherwise contrasting motion directions. Neurons that have RFs centered along the perpendicular axis are well situated to extract information about the movement of the figure attached to the kinetic edge. Viewed in the context of a moving figure, model border-ownership signals developed that lag behind and anticipate the moving edge. O'herron and von der Heydt demonstrate that when a luminance-defined figure is replaced by an ambiguous display without a clear figure–ground interpretation, border-ownership signals may persist for 500 ms or longer (O'Herron & von der Heydt, 2009). This suggests that border-ownership signals may "lag" or persist for hundreds of milliseconds when an edge of a moving figure leaves the RF of a B cell, as is consistent with the border-ownership peaks produced by our model. Persistence in the border-ownership signal when the edge is not in the RF may aid the visual system in maintaining a stable representation of a figure, despite any sudden changes in its direction of motion.

It is not clear how border-ownership signals persist when kinetic figures have moving boundaries. Single cell data indicate that in the case of a figure defined by luminance contrast that moves from one location on the retina to another, border-ownership signals "remap" (O'Herron & von der Heydt, 2013). That is, border-ownership signals transfer from B cells whose receptive fields are initially centered on the edges of the figure to those centered on the edges as the figure moves. The transfer in the border-ownership signal was shown to occur whether the figure motion resulted from a saccadic eye movement or from movement within the visual display. As shown in Fig. 6a, kinetic border-ownership signals in the proposed model track the moving kinetic edge over time. Remapping of the border-ownership signal occurs in the model due to the dynamical between V4 and MT units, which have larger receptive field sizes than B cells.

The *form* of kinetically defined figures likely plays an important role in border-ownership assignment. Royden and colleagues showed that the aspect ratio of the rectangle used to present subjects with shearing motion and texture accretion/deletion greatly impacted the figure–ground judgments of human observers (Royden, Baker, & Allman, 1988). Froyen et al. recently introduced an accretion/deletion display that contained adjacent light and dark rectangular regions of dots that moved in alternating directions (Froyen, Feldman, & Singh, 2013). Although the dots only moved horizontally, humans perceived some of the regions as figural columns rotating in depth. Human judgments of which regions appeared as figural columns depended on the symmetry, convexity, and area of the regions. The focus of the present article is on kinetic occlusion, but the interaction between form and motion should further be explored.

4.2. Comparison with existing models

The proposed model introduces a novel coding strategy for the representation of the relative ordering of nearby surfaces in depth. Many existing models segment objects within a visual scene with respect to multiple "depth planes." For example, once a figure is detected, it may be assigned to a near, intermediate, or far depth plane (Barnes & Mingolla, 2013; Berzhanskaya, Grossberg, & Mingolla, 2007; Kelly & Grossberg, 2000). Our model takes a different approach whereby the relative depth of a kinetic surface is dynamically coded with respect to the magnitude gradient of border-ownership signals produced by MB and PB cells. The population of B cells that produces the greatest border-ownership modulation signals which surface is closest in depth, the population that exhibits the second greatest modulation signals the surface that is next closest in depth, and so forth. The presence of a kinetic edge and occlusion is signaled through the B cell activity gradient. This strategy requires no dedicated system to process depth. Neural coding using activity gradients across several populations have been demonstrated in prefrontal cortex (Averbeck et al., 2002). For non-kinetically defined figures, populations of B cells may dynamically code depth in an activity gradient with respect to disparity or other attributes to which that B cells are tuned (Qiu & von der Heydt, 2005, 2007).

The proposed model extends existing approaches by clarifying how luminance contrast and motion attributes are combined by the primate visual system to afford border-ownership signals. The parvocellular pathway of the proposed model on its own is similar to a number of existing models that reproduce important characteristics of B cells through feedback from units with larger receptive fields (Craft et al., 2007; Layton, Mingolla, & Yazdanbakhsh, 2012; Mihalas et al., 2011). The proposed model is the only one we are aware of that specifically addresses border-ownership in displays containing kinetic occlusion. A number of other models, with varying degrees of biological plausibility, perform figure–ground segregation from kinetic occlusion. Computer vision approaches detect occlusion by training classifiers to learn events or junctions in the spatio-temporal structure of the

optic flow (Niyogi, 1995; Stein, Hoiem, & Hebert, 2007), or use a probabilistic framework to perform figure–ground segregation (Chou, 1995). The biological model of Barnes and Mingolla, builds on the FORMOTION class of models (Berzhanskaya, Grossberg, & Mingolla, 2007), and contains dedicated FORM and MOTION pathways (Barnes & Mingolla, 2013), analogous to the magnocellular and parvocellular pathways used in the proposed model. The model of Barnes and Mingolla uses occlusion detectors, which determine the location of motion onset and offset events. Their model also incorporates a speed-depth bias to assign faster moving textures to the foreground. Our model differs in that occlusion and kinetic edges are not explicitly computed. Sensitivity to kinetic edges and occlusion is an emergent outcome of the dynamical interactions between V2, V4, and MT.

Another distinguishing characteristic of the proposed model is that border-ownership modulation dynamically emerges nearby kinetic edges. Unlike many existing models, our model does not contain dedicated mechanisms for detecting kinetic edges. That is, border-ownership responses in the model do not depend on the a priori detection of kinetic edges. The presence of kinetic edges is read out from the B cell activity gradient. Motion discontinuities (Bayerl & Neumann, 2007; Beck & Neumann, 2010) or onsets/offsets (Barnes & Mingolla, 2013) are routinely detected through the subtraction of dot velocities between two adjacent kinetic surfaces. Many models attribute velocity subtraction to MT neurons that have receptive fields with antagonistic surrounds (Xiao et al., 1995). Our model does not perform motion differencing or assume MT neurons with antagonistic surrounds carry out this operation. MT neurons are velocity tuned and are unlikely to correctly register bilocal motion differences (Born & Bradley, 2005).

4.3. Generalization of the model

In the present article, we focused on random-dot kinematograms to capture a number of key findings surrounding how humans perceive kinetic occlusion. To highlight the dynamics of the core model mechanisms, we implemented the simplest version of the model as possible to capture the range of known human percepts. For example, in simulations we implemented V1 units tuned to 1 pixel/frame in the four cardinal directions. Note that the tuning refers to the optimal response, but units demonstrate a broader sensitivity in velocity space (Fig. 9). Therefore, the dot motion in the visual display need not precisely correlate with the unit tuning to garner an appropriate model response. Straightforward changes could be made to, for example, simulate a diamond defined by moving dots. Sensitivity to a broader range of input speeds and directions could be achieved in the Reichardt motion detection mechanism (Eq. (6)) by sampling a broader range of spatial and temporal displacements ($\vec{\Delta}$ and $\vec{\delta}$ in Eq. (6), respectively). A greater diversity of units would more realistically simulate the spectrum of response latencies observed in V1 (Nowak et al., 1995). Incorporating units tuned to a larger number of speeds and motion directions would not affect the simulation results because model V1 units compete in a shunting network (Eq. (7)). A key property of shunting networks is divisive normalization, which conserves the total activation across the network (Grossberg, 1973). Competition across velocity-tuned units in the model V1 also results in winner-take-all behavior, which ensures the activation of the most active unit approaches one, while that of the remaining units is suppressed to zero. The consequence of the normalization and winner-take-all dynamics is that the units that respond best to the velocities in the input display will eventually respond at the maximal firing rate, at the expense of the activation of other velocity-tuned units. Therefore, adding a greater diversity of velocity-tuned units does not change the qualitative dynamics of the network, as only the units tuned most similarly to the velocities detected in the input display will remain active in the long-run.

Border-ownership coding as proposed by our model extends naturally to a wider range of motion directions and axes of border-ownership. Recall that model V2 tuning to border-ownership and motion is independent, which means that a unit tuned to a particular motion direction could also code any one direction of border-ownership. For example, a MB cell tuned to leftward motion could be tuned to leftward, rightward, downward, or upward border-ownership. Border-ownership sensitivity arises in the model due to spatially and directionally offset feedback from MT and V4. Sensitivity to leftward border-ownership is mediated by feedback from units in MT and V4 whose RFs are centered to either side along the horizontal axis (specified by the quarter-sector Gaussian kernels $\boldsymbol{K}_0$ and $\boldsymbol{K}_\pi$; see Fig. 8). Sensitivity to border-ownership in a diagonal direction could be achieved through feedback from MT and V4 units offset along the same axis. For example, a MB cell would be tuned to border-ownership in the upper-right direction if it received feedback from units whose RFs were to either side along the 45° axis ($\boldsymbol{K}_{\frac{\pi}{4}}$ and $\boldsymbol{K}_{\frac{5\pi}{4}}$).

4.4. Conclusions

Predators face a challenging problem when pursuing camouflaged prey. The visual system must perform figure–ground segregation based on motion alone, since luminance variations may not be reliable. We proposed a dynamical explanation how primate cortex may solve this problem whereby border-ownership signals in magnocellular and parvocellular pathways code the relative depth of the figure's borders through an activity gradient. Our results suggest that the visual system does not need to explicitly detect a kinetic edge prior to performing figure–ground segregation, and that feedback plays an important role in detecting kinetic occlusion.

Acknowledgments

This work was supported in part by CELEST (NSF SBE-0354378 and OMA-0835976), the Office of Naval Research (ONR N00014-11-1-0535) and Air Force Office of Scientific Research (AFOSR FA9550-12-1-0436).

Appendix A. Model equations

The model is a system of Hodgkin–Huxley type ordinary differential equations that emphasizes the dynamical interactions between cells in different primate brain areas. Model units are controlled by shunting interactions, which implement a number of important dynamical properties, such as divisive normalization and boundedness (Carandini & Heeger, 2011; Grossberg, 1973; Heeger, 1992). No attempt was made to optimize curve fits between the model and psychophysical or neurophysiological data. The firing rates of model neurons range from zero to one. The model was implemented using Wolfram Mathematica 9 on a 2.66 GHz 8-core Apple Mac Pro with 64 GB of memory. Numerical integration of the model system was performed using the function NDSolve.

Equations often apply to all units in a network layer, in which case we use bold matrix notation. For example, **x** stands for the set of cells at every spatial location (i, j) in the input display.

Model equations generally resemble the following membrane equation termed a recurrent competitive field (Grossberg, 1973):

$$\frac{dx_i}{dt} = -x_i + (1 - x_i)(f(x_i) + I_i^+) - x_i\left(\sum_{k \neq i} f(x_i) + I_i^-\right) \quad (2)$$

Eq. (2) is a shunting equation that describes the activity, x, of the ith cell in a neural network layer. I_i^+ and I_i^- specify the excitatory and inhibitory inputs, respectively, that the ith cell receives. The multiplicative term that involves I_i^+ implements shunting excitation, and the term that involves I_i^- implements shunting inhibition. In Eq. (2), $f(x_i)$ is a signal function that specifies the nature of the feedback from cells in the same network layer. We often configure the signal function as the expansive nonlinearity $f(x_i) = x_i^2$, which results in a contrast-enhancing network (Grossberg, 1973).

Model units between layers connected 1-to-1 or a Gaussian kernel specifies the convergence in feedforward or feedback connections. If cells are connected 1-to-1, the cell with a receptive field centered at position (i, j) in one layer projects to another cell with a receptive field centered at position (i, j). In our implementation, each layer has the same number of units equal to the number of pixels in the input visual display. The Gaussian kernel $\boldsymbol{F}(\boldsymbol{x}; \sigma)$ defines how connections converge from one network layer onto the next when the spatial extent of a receptive field is larger than that of its input. We use the following Gaussian kernel,

$$\boldsymbol{F}(\boldsymbol{x}; \sigma, r, a) = a \times e^{-\frac{\boldsymbol{x} \cdot \boldsymbol{x}}{2\sigma^2}} \tag{3}$$

where the operator $\cdot$ corresponds to the dot product, a scales the Gaussian, and r is the kernel radius. Convolution between a matrix $\mathbf{x}$ and kernel $\mathbf{F}$ is indicated by the * operator and is always centered at each cell position (i, j).

Inputs to the model are 30 frame 64×64 pixel kinetic random dot displays. For convenience we refer to time (t) in the model as continuously spanning 0–30, and new input frames are presented to the model at integer times. At any one point in time, each pixel in the input $I(i, j)$ takes on a random value between zero and one. Across successive frames, dots were displaced by at most one pixel. Unless otherwise noted, simulation results are from the final frame of each input sequence.

In the following sections, we provide equations for units in the magnocellular and parvocellular pathways of the model.

A.1. Magnocellular pathway

The input luminance $\mathbf{I}(t)$ is transformed into signals of increments $\mathbf{J}^+(t)$ and decrements $\mathbf{J}^-(t)$, which represent the change in the input across successive frames. These signals correspond to the coding of luminance increases and decreases by the ON and OFF retinal ganglion cells.

$$\mathbf{J}^+(t) = [\mathbf{I}(\lceil t \rceil) - \mathbf{I}(\lceil t-1 \rceil)]^+ \tag{4}$$

$$\mathbf{J}^-(t) = [\mathbf{I}(\lceil t-1 \rceil) - \mathbf{I}(\lceil t \rceil)]^+ \tag{5}$$

The notation $\lceil \cdot \rceil$ refers to taking the ceiling of the operand, and $[.]^+$ denotes the half-wave rectification max(.,0).

Motion detection. We detect motion in the model using a Reichardt or correlation-based mechanism (Van Santen & Sperling, 1985). Spatial and temporal offsets in the luminance increment and decrement signals $\mathbf{J}^{\pm}(t)$ are independently compared. In the present work, we use two temporal offsets or conduction delays $\vec{\delta} = \{\delta_1, \delta_2\}$. For convenience, we set $\vec{\delta} = \{0, 1\}$ to compare signals across successive frames in the input. At every spatial position, we make two adjacent spatial comparisons $\boldsymbol{\Delta} = \{\overrightarrow{\Delta_1}, \overrightarrow{\Delta_2}\}$ in the increment and decrement signals. Each pair of offsets $\vec{\Delta}$ impacts the direction of motion that is detected. For example in a standard Cartesian system, the offsets $\vec{\Delta} = \{(0,0),(1,0)\}$ would afford a sensitivity to motion in the horizontal direction. Multiple spatial comparisons are used to improve the robustness of the motion signal.

The following equation computes the motion signal $Z(t)$ in a particular direction d:

$$\boldsymbol{Z}_{\boldsymbol{d}}^{\pm}(t) = \frac{\mathbf{J}^{\pm}_{\overrightarrow{\Delta_1}}(t-\delta_1) \cdot \mathbf{J}^{\pm}_{\overrightarrow{\Delta_2}}(t-\delta_2)}{\left\|\mathbf{J}^{\pm}_{\overrightarrow{\Delta_1}}(t-\delta_1)\right\| \cdot \left\|\mathbf{J}^{\pm}_{\overrightarrow{\Delta_2}}(t-\delta_2)\right\|} \tag{6}$$

In Eq. (6), the motion signal is computed separately for the luminance increments and decrement signals $\mathbf{J}^{\pm}(t)$. The correlation between pairs of spatio-temporal samples of $\mathbf{J}^{\pm}$ is computed in the numerator, and the result is normalized according to the product of the Euclidean norms $\|\cdot\|$. We used four unit spatial offsets $\vec{\Delta}$ to detect motion in four directions (d): leftward, rightward, upward, and downward.

Model V1 (magnocellular). We simulate V1 cells, which are tuned to one of the four motion directions described above. The motion signal for the luminance increments and decrements is combined in the input for each V1 unit. The following equation describes the dynamics of magnocellular (m) V1 cells $\boldsymbol{v}_{m,d}$ tuned to the direction of motion d:

$$\frac{1}{\tau_{V1}} \frac{d\boldsymbol{v}_{m,d}}{dt} = -\boldsymbol{v}_{m,d} + (1 - \boldsymbol{v}_{m,d}) \odot (\boldsymbol{v}_{m,d}^2 + \boldsymbol{Z}_{\boldsymbol{d}}^{+} + \boldsymbol{Z}_{\boldsymbol{d}}^{-}) - \boldsymbol{v}_{m,d} \odot \sum_{k \neq d} \boldsymbol{v}_{m,k}^2 \tag{7}$$

In Eq. (7), magnocellular V1 cells compete in a contrast-enhancing network across motion direction. The parameter τ_{V1} scales the temporal dynamics of the network. We fixed $\tau_{V1} = 5$ in all simulations. The operator $\odot$ indicates element-wise multiplication.

Fig. 9 shows response characteristics of model V1 units to a leftward moving vertical bar (Eq. (7)).

Model V2 (magnocellular). MB cells in the model are selective to motion direction d and have a side-of-figure selectivity θ. The input kinetic dot displays contained either vertically or horizontally oriented kinetic edges, so we implemented four border-ownership directions: right, up, left, and down $(\theta \in \{0, \frac{\pi}{2}, \pi, \frac{3\pi}{2}\})$.

$$\frac{1}{\tau_{V2}} \frac{d\boldsymbol{b}_{m,d,\theta}}{dt} = -\boldsymbol{b}_{m,d,\theta} + (1 - \boldsymbol{b}_{m,d,\theta}) \odot (\boldsymbol{F}_{V2} * \boldsymbol{v}_{m,d}) - \gamma_{V2_m} \times (\boldsymbol{b}_{m,d,\theta}) \odot (\boldsymbol{A}_{m,d,\theta} + \boldsymbol{B}_{m,d,\theta} + \boldsymbol{C}_{m,d,\theta}) \tag{8}$$

In Eq. (8), $\boldsymbol{F}_{V2}$ is a Gaussian kernel with $\sigma = 2$, $a = 5$, $r = 1$. The parameter γ_{V2_m} balances the ratio between the inhibitory and excitatory inputs to the cell. We fixed $\gamma_{V2_m} = 16$ and the time constant $\tau_{V2} = 10$.

The term $\boldsymbol{A}_{m,d,\theta}$ describes the feedback the MB cell receives from MT cells that have different preferred motion directions. The MB cells that have a preferred border-ownership direction toward or away from the MT cell receptive field centers are inhibited.

$$\boldsymbol{A}_{m,d,\theta} = \sum_s \sum_{k \neq d} \boldsymbol{K}^s_{MT,\theta} * \boldsymbol{m}_{k,s} + \sum_s \sum_{k \neq d} \boldsymbol{K}^s_{MT,\theta+\pi} * \boldsymbol{m}_{k,s} \tag{9}$$

The kernels $\boldsymbol{K}^s_{MT,\theta}$ and $\boldsymbol{K}^s_{V4,\theta}$ are quarter sectors of a Gaussian kernel (Fig. 8), and specify how feedback propagates from V4 and MT to V2. For example, $\boldsymbol{K}^s_{V4,0}$ is the right 90° sector of a Gaussian kernel ($\theta = 0$) and specifies the set of feedback projections from V4 to inhibit B cells in V2 from the right hand side. The kernel $\boldsymbol{K}^s_{\theta}$ results in inhibition of B cells that have the preferred border-ownership direction *toward* the V4 or MT unit's receptive field center. The kernel $\boldsymbol{K}^s_{\theta+\pi}$ results in inhibition of B cells that have the preferred border-ownership direction *away from* the V4 or MT unit's receptive field center. The superscript s indicates the receptive field size of the V4 or MT unit. We simulated V4 and MT units with two different receptive field sizes ($r = 2$ pixels, $r = 5$ pixels) to pool fine and coarse motion and luminance signals. The feedback that V2 units

receive is the summed outputs of V4 and MT units with different receptive field sizes. If the model implementation includes more than one speed sensitivity, signals from MT units with different speed tunings but the same direction tuning are summed. For all kernels $\boldsymbol{K}^s_\theta$, $c = 1$ and σ was set such that $r/\sigma = 2$.

The term $\boldsymbol{B}_{m,d,\theta}$ describes the feedback the MB cell receives from MT cells that have same preferred motion direction. The MB cells that have a preferred border-ownership direction away from the MT cells are inhibited.

$$\boldsymbol{B}_{m,d,\theta} = \sum_s \boldsymbol{K}^s_{MT,\theta+\pi} * \boldsymbol{m}_{d,s} \tag{10}$$

The term $\boldsymbol{C}_{m,d,\theta}$ describes the feedback MB cells receive from V4 cells. Each MB cell receives feedback from a V4 cell if the B cell's preferred border-ownership direction and motion sensitivity is consistent with the presence of accretion/deletion within the V4 unit's receptive field. For example, the MB cells in Fig. 3c and d signal the presence of accretion/deletion on the right side of the V4 unit's receptive field. The MB cell that has a preferred border-ownership direction toward the V4 cell's receptive field center is inhibited. The set Θ describes the MB cell border-ownership direction preferences that point toward or away from the V4 cell receptive field center in units with perpendicular motion preferences.

$$\boldsymbol{C}_{m,d,\theta} = \sum_s \sum_{q\in\Theta} \boldsymbol{K}^s_{V4,q} * \boldsymbol{n}_s \tag{11}$$

In Eq. (11), $\boldsymbol{n}_s$ indicates V4 cells with receptive field size s.

Fig. 10 depicts response properties of V2 border-ownership cells. Fig. 10a shows that the basic motion tuning of V2 units, inherited from V1, is not affected by the modulatory feedback from model MT (Eqs. (9), (10)) and V4 (Eq. (11)). Fig. 10b shows that the border-ownership responses (e.g. Fig. 5) are robust to a broad range of luminance contrasts in the visual display.

Model MT. As mentioned above, MT units have two different receptive field sizes ($r = 2$ pixels, $r = 5$ pixels) to pool fine and coarse motion signals. The following shunting equation describes the dynamics of MT cells $\boldsymbol{m}_{d,s}$ with receptive field size s and preferred motion direction d:

$$\frac{1}{\tau_{MT}}\frac{d\boldsymbol{m}_{d,s}}{dt} = -\boldsymbol{m}_{d,s} + (1 - \boldsymbol{m}_{d,s}) \odot (\boldsymbol{D}_{MT,d,s}) \tag{12}$$

In Eq. (12), we set $\tau_{MT} = 10$.

The term $\boldsymbol{D}_{MT,d,s}$ describes the feedforward input MT cells receive from model V1 units $\boldsymbol{v}_{m,d}$ with the same preferred motion direction d.

$$\boldsymbol{D}_{MT,d,s} = \sqrt{\sum_{y,z\in\Phi} (\boldsymbol{K}^s_{MT,\theta_y} * \boldsymbol{v}_{m,d}) \odot (\boldsymbol{K}^s_{MT,\theta_z} * \boldsymbol{v}_{m,d})} \tag{13}$$

In Eq. (13), Φ refers to the set of permutation pairs of Gaussian kernel sector orientations, and y and z take on specific values from the set. For example, $\Phi = \{(0, \frac{\pi}{2}), (0, \pi), \cdots\}$, in which case $y = 0$, $z = \frac{\pi}{2}$ and so forth. The term $\boldsymbol{D}_{MT,d,s}$ multiplicatively combines the activity of V1 cells that fall in different sectors of the MT unit's Gaussian-shaped receptive field. This yields nonlinear increases in the MT unit's activation proportionate to the number of quadrants of the receptive field that contain motion in the preferred direction.

A.2. Parvocellular pathway

Model V1 (parvocellular). V1 complex cells respond to luminance contrast with orientation ψ in the input luminance signal $\boldsymbol{I}$ at time t. Model units pool over input from a pair of odd-symmetric simple cell kernels $\boldsymbol{S}_\psi$

$$\boldsymbol{S}_\psi = \cos(\psi)\boldsymbol{S}_x + \sin(\psi)\boldsymbol{S}_y \tag{14}$$

where $\boldsymbol{S}_x$ and $\boldsymbol{S}_y$ are the horizontal and vertical partial derivatives of the 2D Gaussian kernel $\boldsymbol{F}$ (Freeman & Adelson, 1991; Layton, Mingolla, & Yazdanbakhsh, 2012). We obtain complex cell input $\boldsymbol{H}_\psi$ by summing over anti-phase simple cell kernels, thresholding, and rectifying the signal.

$$\boldsymbol{H}_\psi = [\boldsymbol{S}_\psi * \boldsymbol{I} + \boldsymbol{S}_{\psi+\pi} * \boldsymbol{I} - \Gamma]^+ \tag{15}$$

In our simulations, we consider horizontally and vertically oriented edges ($\psi \in \{0, \frac{\pi}{2}\}$). The complex cell input is passed through Eq. (6) with $\vec{A} = \vec{0}$ to detect the persistence in the contrast signal over time, yielding $\boldsymbol{Z}^\psi_0(t)$. The following equation describes the dynamics of parvocellular V1 cells that compete in a contrast enhancing network across preferred orientations ψ to extract the dominant edge orientation.

$$\frac{1}{\tau_{V1}}\frac{d\boldsymbol{v}_{p,\psi}}{dt} = -\boldsymbol{v}_{p,\psi} + (1 - \boldsymbol{v}_{p,\psi}) \odot (\boldsymbol{v}^2_{p,\psi} + \boldsymbol{Z}^\psi_0(t)) - \boldsymbol{v}_{p,\psi} \odot \sum_{k\neq\psi} \boldsymbol{v}^2_{p,k} \tag{16}$$

Model V2 (parvocellular). Model PB cells $\boldsymbol{b}_{p,\theta}$ are selective to oriented luminance contrast and have a side-of-figure selectivity θ.

$$\frac{1}{\tau_{V2}}\frac{d\boldsymbol{b}_{p,\theta}}{dt} = -\boldsymbol{b}_{p,\theta} + (1 - \boldsymbol{b}_{p,\theta}) \odot (\boldsymbol{L}_\theta) - \gamma_{V2_p} \times (\boldsymbol{b}_{p,\theta}) \odot (\boldsymbol{A}_{p,d,\theta} + \boldsymbol{B}_{p,\theta}) \tag{17}$$

PB cells that have leftward and rightward border-ownership preferences receive input $\boldsymbol{L}_\theta$ from V1 cells sensitive to vertically oriented contrast. PB cells that have upward and downward border-ownership preferences receive input $\boldsymbol{L}_\theta$ from V1 cells.

$$\boldsymbol{L}_\theta = \begin{cases} \boldsymbol{F}_{V2} * \boldsymbol{v}_{p,0} & \theta \in \{0, \pi\} \\ \boldsymbol{F}_{V2} * \boldsymbol{v}_{p,\pi/2} & \theta \in \{\frac{\pi}{2}, \frac{3\pi}{2}\} \end{cases} \tag{18}$$

We fixed $\gamma_{V2_p} = 16$. The term $\boldsymbol{A}_{p,d,\theta}$ describes the feedback the PB cell receives from MT cells that are sensitive to motion direction d. The PB cells that have a preferred border-ownership direction toward or away from the center of the MT cell receptive field are inhibited.

$$\boldsymbol{A}_{p,d,\theta} = \sum_s \sum_{k\neq d} \boldsymbol{K}^s_{MT,\theta} * \boldsymbol{m}_{k,s} + \sum_s \sum_{k\neq d} \boldsymbol{K}^s_{MT,\theta+\pi} * \boldsymbol{m}_{k,s} \tag{19}$$

The term $\boldsymbol{B}_{p,\theta}$ describes the feedback the PB cell receives from V4. The PB cells that have a preferred border-ownership direction away from the V4 cell receptive field center are inhibited.

$$\boldsymbol{B}_{p,\theta} = \sum_s \boldsymbol{K}^s_{V4,\theta+\pi} * \boldsymbol{n}_s \tag{20}$$

Model V4. As mentioned above, V4 units have two different receptive field sizes ($r = 2$ pixels, $r = 5$ pixels) to pool fine and coarse luminance contrast signals. The following shunting equation describes the dynamics of the V4 cell $\boldsymbol{n}_s$ with receptive field size s:

$$\frac{1}{\tau_{V4}}\frac{d\boldsymbol{n}_s}{dt} = -\boldsymbol{n}_s + (1 - \boldsymbol{n}_{d,s}) \odot (\boldsymbol{D}_{V4,s} \odot (1 + \boldsymbol{E}_{V4,s})) \tag{21}$$

In Eq. (21), $\tau_{V4} = 10$. Each V4 unit is driven by luminance contrast signals from V1 ($\boldsymbol{D}_{V4,s}$), and the response is enhanced when texture accretion/deletion is present within the receptive field ($\boldsymbol{E}_{V4,s}$). V4 units pool over orientations ψ in their feedforward inputs from parvocellular V1 cells.

$$\boldsymbol{D}_{V4,s} = \sqrt{\sum_\psi \sum_{y,z\in\Phi} (\boldsymbol{K}^s_{V4,\theta_y} * \boldsymbol{v}_{p,\psi})(\boldsymbol{K}^s_{V4,\theta_z} * \boldsymbol{v}_{p,\psi})} \tag{22}$$

Finally, the following equation describes the excitatory modulation V4 units receive from motion-sensitive magnocellular cells in V1.

$$\boldsymbol{E}_{V4,s} = \sqrt{\sum_{u\in\Psi}\sum_{y,z\in\Phi}(\boldsymbol{K}^{s}_{V4,\theta_y} * \boldsymbol{v}_{m,u}) \odot (\boldsymbol{K}^{s}_{V4,\theta_z} * \boldsymbol{v}_{m,u})} \quad (23)$$

The set Ψ defines the perpendicular motion preferences that point toward or away from the V4 cell receptive field center. The activity of V4 units $\boldsymbol{n}_s$ is enhanced, proportionate to the number of quadrants within the Gaussian receptive field that contain texture accretion/deletion.

References

Allman, J., Miezin, F., & McGuinness, E. (1985). Direction- and velocity-specific responses from beyond the classical receptive field in the middle temporal visual area (MT). *Perception, 14*(2), 105–126.

Anderson, J. C., & Martin, K. A. C. (2001). Connection from cortical area V2 to MT in macaque monkey. *The Journal of Comparative Neurology, 443*(1), 56–70. http://dx.doi.org/10.1002/cne.10100.

Angelucci, A., Levitt, J. B., Walton, E. J. S., Hupe, J.-M., Bullier, J., & Lund, J. S. (2002). Circuits for local and global signal integration in primary visual cortex. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience, 22*(19), 8633–8646.

Averbeck, B. B., Chafee, M. V., Crowe, D. A., & Georgopoulos, A. P. (2002). Parallel processing of serial movements in prefrontal cortex. *Proceedings of the National Academy of Sciences, 99*(20), 13172–13177. http://dx.doi.org/10.1073/pnas.162485599.

Barnes, T., & Mingolla, E. (2013). Neural networks. *Neural Networks, 37*, 141–164. http://dx.doi.org/10.1016/j.neunet.2012.09.011.

Bayerl, P., & Neumann, H. (2007). A fast biologically inspired algorithm for recurrent motion estimation. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 29*(2), 246–260. http://dx.doi.org/10.1109/TPAMI.2007.24.

Beck, C., & Neumann, H. (2010). Interactions of motion and form in visual cortex – A neural model. *Journal of Physiology – Paris, 104*(1), 61–70. http://dx.doi.org/10.1016/j.jphysparis.2009.11.005.

Berzhanskaya, J., Grossberg, S., & Mingolla, E. (2007). Laminar cortical dynamics of visual form and motion interactions during coherent object motion perception. *Spatial Vision, 20*(4), 337–395. http://dx.doi.org/10.1163/156856807780919000.

Born, R. T. (2000). Center-surround interactions in the middle temporal visual area of the owl monkey. *Journal of Neurophysiology, 84*(5), 2658–2669.

Born, R. T., & Bradley, D. C. (2005). Structure and function of visual area MT. *Annual Review of Neuroscience, 28*, 157–189. http://dx.doi.org/10.1146/annurev.neuro.26.041002.131052.

Carandini, M., & Heeger, D. J. (2011). Normalization as a canonical neural computation. *Nature Publishing Group*. http://dx.doi.org/10.1038/nrn3136.

Chen, M., Li, P., Zhu, S., Han, C., Xu, H., Fang, Y., et al. (2012). Motion boundary response domains in awake monkey V2. *Journal of Vision, 12*(9), 1308. http://dx.doi.org/10.1167/12.9.1308.

Chou, G. T. (1995). A model of figure–ground segregation from kinetic occlusion. *Computer Vision*.

Craft, E., Schutze, H., Niebur, E., & von der Heydt, R. (2007). A neural model of figure–ground organization. *Journal of Neurophysiology, 97*(6), 4310–4326. http://dx.doi.org/10.1152/jn.00203.2007.

Cutting, J. E. (1997). How the eye measures reality and virtual reality. *Behavior Research Methods, 29*(1), 27–36. http://dx.doi.org/10.3758/BF03200563.

Egelhaaf, M., Borst, A., & Reichardt, W. (1989). Computational structure of a biological motion-detection system as revealed by local detector analysis in the flys nervous-system. *Journal of the Optical Society of America A: Optics and Image Science, 6*(7), 1070–1087.

Freeman, W. T., & Adelson, E. H. (1991). The design and use of steerable filters. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 13*(9), 891–906.

Friedman, H. S., Zhou, H., & von der Heydt, R. (2003). The coding of uniform colour figures in monkey visual cortex. *The Journal of Physiology, 548*(2), 593–613. http://dx.doi.org/10.1113/jphysiol.2002.033555.

Froyen, V., Feldman, J., & Singh, M. (2013). Rotating columns: Relating structure-from-motion, accretion/deletion, and figure/ground. *Journal of Vision, 13*(10), 6. http://dx.doi.org/10.1167/13.10.6.

Gharat, A., & Baker, C. L. (2012). Motion-defined contour processing in the early visual cortex. *Journal of Neurophysiology, 108*(5), 1228–1243. http://dx.doi.org/10.1152/jn.00840.2011.

Gibson, J. J., Kaplan, G. A., & Reynolds, H. N. (1969). The change from visible to invisible. *Perception & Psychophysics, 5*(2), 113–116. http://dx.doi.org/10.3758/BF03210533.

Grossberg, S. (1973). Contour enhancement, short-term memory, and constancies in reverberating neural networks. *Studies in Applied Mathematics, 52*(3), 213–257.

Handa, T., Katai, S., Kuno, R., Unno, S., Inoue, M., & Mikami, A. (2008). Differential activity to shapes under shape-from-motion condition in macaque middle temporal area. *Neuroscience, 156*(4), 1118–1135. http://dx.doi.org/10.1016/j.neuroscience.2008.06.029.

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

Hubel, D. H., & Wiesel, T. N. (1968). Receptive fields and functional architecture of monkey striate cortex. *The Journal of Physiology, 195*(1), 215–243.

Hupé, J. M., James, A. C., Girard, P., Lomber, S. G., Payne, B. R., & Bullier, J. (2001). Feedback connections act on the early part of the responses in monkey visual cortex. *Journal of Neurophysiology, 85*(1), 134–145.

Kaplan, G. A., & Gibson, J. J. (1969). Kinetic disruption of optical texture: The perception of depth at an edge. *Perception & Psychophysics, 6*(4), 193–198. http://dx.doi.org/10.3758/BF03207015.

Kelly, F., & Grossberg, S. (2000). Neural dynamics of 3-D surface perception: Figure–ground separation and lightness perception. *Perception & Psychophysics, 62*(8), 1596–1618. http://dx.doi.org/10.3758/BF03212158.

Layton, O. W., Mingolla, E., & Yazdanbakhsh, A. (2012). Dynamic coding of border-ownership in visual cortex. *Journal of Vision, 12*(13), 8. http://dx.doi.org/10.1167/12.13.8.

Lund, J. S., Angelucci, A., & Bressloff, P. C. (2003). Anatomical substrates for functional columns in macaque monkey primary visual cortex. *Cerebral Cortex*.

Marcar, V. L., Raiguel, S. E., Xiao, D., & Orban, G. A. (2000). Processing of kinetically defined boundaries in areas V1 and V2 of the macaque monkey. *Journal of Neurophysiology, 84*(6), 2786–2798.

Marcar, V. L., Xiao, D. K., Raiguel, S. E., Maes, H., & Orban, G. A. (1995). Processing of kinetically defined boundaries in the cortical motion area MT of the macaque monkey. *Journal of Neurophysiology, 74*(3), 1258–1270.

Mihalas, S., Dong, Y., von der Heydt, R., Niebur, E., & Romo, R. (2011). Mechanisms of perceptual organization provide autozoom and auto-localization for attention to objects. *Proceedings of the National Academy of Sciences of the United States of America, 108*(18), 7583–7588. http://dx.doi.org/10.2307/41242224?ref=search-gateway:4e6febf973a218aa04a7bbd61ed08336.

Mysore, S. G. S., Vogels, R. R., Raiguel, S. E. S., & Orban, G. A. G. (2006). Processing of kinetic boundaries in macaque V4. *Journal of Neurophysiology, 95*(3), 1864–1880. http://dx.doi.org/10.1152/jn.00627.2005.

Mysore, S. G., Vogels, R., Raiguel, S. E., & Orban, G. A. (2008). Shape selectivity for camouflage-breaking dynamic stimuli in dorsal V4 neurons. *Cerebral Cortex, 18*(6), 1429–1443. http://dx.doi.org/10.1093/cercor/bhm176.

Nakamura, H., Gattass, R., Desimone, R., & Ungerleider, L. G. (1993). The modular organization of projections from areas V1 and V2 to areas V4 and TEO in macaques. *Journal of Neuroscience, 13*(9), 3681–3691.

Niyogi, S. A. (1995). Detecting kinetic occlusion. *Computer Vision*.

Nowak, L. G., Munk, M. H., Girard, P., & Bullier, J. (1995). Visual latencies in areas V1 and V2 of the macaque monkey. *Visual Neuroscience, 12*(2), 371–384.

O'Herron, P., & von der Heydt, R. (2009). Short-term memory for figure–ground organization in the visual cortex. *Neuron, 61*(5), 801–809. http://dx.doi.org/10.1016/j.neuron.2009.01.014.

O'Herron, P., & von der Heydt, R. (2013). Remapping of border ownership in the visual cortex. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience, 33*(5), 1964–1974. http://dx.doi.org/10.1523/JNEUROSCI.2797-12.2013.

Osorio, D., & Srinivasan, M. V. (1991). Camouflage by edge enhancement in animal coloration patterns and its implications for visual mechanisms. *Proceedings of the Royal Society B: Biological Sciences, 244*(1310), 81–85. http://dx.doi.org/10.1098/rspb.1991.0054.

Qiu, F. T., & von der Heydt, R. (2005). Figure and ground in the visual cortex: V2 combines stereoscopic cues with gestalt rules. *Neuron, 47*(1), 155–166. http://dx.doi.org/10.1016/j.neuron.2005.05.028.

Qiu, F. T., & von der Heydt, R. (2007). Neural representation of transparent overlay. *Nature Neuroscience, 10*(3), 283–284. http://dx.doi.org/10.1038/nn1853.

Regan, D., & Beverley, K. I. (1984). Figure–ground segregation by motion contrast and by luminance contrast. *Journal of the Optical Society of America A: Optics and Image Science, 1*(5), 433–442.

Ringach, D. L. (2002). Spatial structure and symmetry of simple-cell receptive fields in macaque primary visual cortex. *Journal of Neurophysiology*. http://dx.doi.org/10.1152/jn.00881.2001.

Rockland, K. S., & Knutson, T. (2000). Feedback connections from area MT of the squirrel monkey to areas V1 and V2. *The Journal of Comparative Neurology, 425*(3), 345–368. http://dx.doi.org/10.1002/1096-9861(20000925)425:3<345::AID-CNE2>3.0.CO;2-O.

Royden, C. S. C., Baker, J. F. J., & Allman, J. J. (1988). Perceptions of depth elicited by occluded and shearing motions of random dots. *Perception, 17*(3), 289–296.

Snowden, R. J., Treue, S., Erickson, R. G., & Andersen, R. A. (1991). The response of area Mt and V1 neurons to transparent motion. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience, 11*(9), 2768–2785.

Stein, A., Hoiem, D., & Hebert, M. (2007). Learning to find object boundaries using motion cues. In *IEEE international conference on computer vision. Proceedings* (pp. 1–8). http://dx.doi.org/10.1109/ICCV.2007.4408841.

Tanaka, K., Hikosaka, K., Saito, H., & Yukie, M. (1986). Analysis of local and wide-field movements in the superior temporal visual areas of the macaque monkey. *The Journal of Neuroscience, 6*(1), 134–144.

Tanigawa, H., Lu, H. D., & Roe, A. W. (2010). Functional organization for color and orientation in macaque V4. *Nature Neuroscience, 1–8*. http://dx.doi.org/10.1038/nn.2676.

Ungerleider, L. G., Galkin, T. W., Desimone, R., & Gattass, R. (2008). Cortical connections of area V4 in the macaque. *Cerebral Cortex, 18*(3), 477–499. http://dx.doi.org/10.1093/cercor/bhm061.

Van Santen, J., & Sperling, G. (1985). Elaborated Reichardt detectors. *Journal of the Optical Society of America A*.

von der Heydt, R., Qiu, F. T., & He, Z. J. (2003). *Neural mechanisms in border ownership assignment: Motion parallax and gestalt cues*. Naples, FL: Vision Sciences Society. <http://vlab.mb.jhu.edu/projects/bo-motion/results.shtml>.

Xiao, D. K., Raiguel, S., Marcar, V., Koenderink, J., & Orban, G. A. (1995). Spatial heterogeneity of inhibitory surrounds in the middle temporal visual area. *Proceedings of the National Academy of Sciences, 92*(24), 11303–11306.

Yonas, A., Craton, L. G., & Thompson, W. B. (1987). Relative motion: Kinetic information for the order of depth at an edge. *Perception & Psychophysics, 41*(1), 53–59.

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