

# Early representation of shape by onset synchronization of border-ownership-selective cells in the V1-V2 network

Yasuhiro Hatori<sup>1,2,\*</sup> and Ko Sakai<sup>1</sup>

<sup>1</sup>*Department of Computer Science, Graduate School of Systems and Information Engineering, University of Tsukuba, 1-1-1 Tennodai, Tsukuba, Ibaraki 305-8573, Japan*

<sup>2</sup>*Japan Society for the Promotion of Science, Tokyo, Japan*

\*Corresponding author: [hatori@cvs.cs.tsukuba.ac.jp](mailto:hatori@cvs.cs.tsukuba.ac.jp)

Received May 28, 2013; revised November 2, 2013; accepted January 1, 2014;  
posted February 3, 2014 (Doc. ID 191041); published March 13, 2014

Construction of surface is a crucial step toward the representation of shape through the integration of local information. Physiological studies have reported that the primary visual cortex (V1) codes the medial axis (MA) that is a skeletal structure equidistant from nearby contours, suggesting the early representation of surface in V1. Although the neural basis of surface construction has not been clarified, the onset synchronization of border ownership (BO)-selective cells is a plausible candidate for the generation of surface. We investigated computationally the representation of surface in a biophysically detailed model of primate V1-V2 networks. The simulation results showed that the simultaneous arrival of signals from BO-selective cells evoked strong responses of V1 cells located around the MA. The simulation results lead to a prediction that the perception of the direction of figure (DOF) depends on the degree of synchronous presentation of contour. We conducted a psychophysical experiment and showed that the perception of the DOF is biased toward a highly synchronized contour. These results suggest a crucial role of the onset synchronization of BO-selective cells for the construction of early representation of shape. © 2014 Optical Society of America

OCIS codes: (330.4060) Vision modeling; (330.5020) Perception psychology; (330.5510) Psychophysics; (330.7310) Vision.

<http://dx.doi.org/10.1364/JOSAA.31.000716>

## 1. INTRODUCTION

Physiological studies have reported that around 60% of neurons in V2 exhibit a selectivity to border ownership (BO) that indicates the direction of figure (DOF) along a contour (e.g., [1]). BO is essentially a local cue (e.g., [2]) that indicates the DOF with respect to the classical receptive field (CRF) of the cell. Computational analyses, together with physiological data, have indicated a crucial role of surround modulation in the determination of BO that is effective from several to 15° of visual angle [1,3,4]. The rapid response of BO-selective cells (with a latency of around 70 ms) is also consistent with the local nature of the signal, most likely the processing in early-to-intermediate-level visual areas. The integration of local BO signals is necessary to yield veridical, consistent DOFs along the entire contours within a view. The integration will lead to the construction of surfaces, which is considered to be a basis for grouping local features toward the representation of an object and its shape.

Several physiological, psychophysical, and computational studies have shown the attractiveness and plausibility of the medial axis (MA) as a cortical representation of surface. Physiological studies have showed that cells in the monkey V1 exhibit strong or modulatory responses when they are located at the MA [5–8]. A series of studies have also indicated that the activities related to the MA are observed beyond V3, up to inferior temporal (IT) cortex (e.g., [9,10]). A recent study reported that IT neurons appear to represent three-dimensional

shape by means of the MA [10]. Psychophysical studies have shown that contrast sensitivity increases at the MA of an object [11,12], suggesting that the MA contributes to the perception of shape. Computational studies have widely recognized the MA as an excellent descriptor of shape (e.g., [13–15]) with a number of advantages including applicability, compactness, and invariance to scale and view direction. A recent computational study reported that the MA helped the binding of local BO signals to establish global figure-ground segregation [16]. Although the results of that study are inconsistent with physiological findings (because the latency of the BO signal is 10 ms faster than that of the MA signal [1,5]), a combination of the MA and BO is attractive from a computational viewpoint.

Although the neural mechanisms underlying the representation of surface have not been clarified, the integration of local BO signals through the onset synchronization of BO-selective cells in V2 is a plausible candidate mechanism. Dong *et al.* reported that BO-selective cells show significant correlation from shortly after the stimulus onset up to 200 ms, with a peak at around 35 ms [17]. We hypothesize that the onset synchronization of BO-selective cells generates the MA as a way of binding local BO signals to represent surfaces. Although the origin of the synchronization is unclear, we consider that the propagation of the onset synchronization in V1 leads to the synchronization in V2, according to the physiological studies that have reported the onset synchronization

in V1 [18,19]. An alternative idea would be the simultaneous feedback from V4; however, we did not adopt this idea because of the lack of physiological evidence and the complexity of such model. The BO signals propagating from contours will meet at a location that is equidistant from the contours, which corresponds to the MA if the spikes are synchronized at the onset and propagated in areas that preserve retinotopy. If the propagation of spikes from BO-selective cells is biased toward the DOF, even in the presence of occlusion, the spikes provide signals that are sufficient for the reconstruction of surfaces by means of the MA. Onset synchronization of V1 cells along contours is not sufficient for the veridical determination of occluded surfaces and global figure-ground; however, the synchronization of BO-selective cells and their bias in propagation are capable of determining surfaces correctly.

We investigated whether the onset synchronization of BO signals induces the formation of the MA. Specifically, we examined the nature of the signals that emerge with regard to the surface representation in a biophysically detailed model of the V1-V2 network that includes V1 cells, BO-selective cells, lateral connections, and bidirectional connections between V1 and V2. We predicted that the feedback signals from the synchronized BO-selective cells originate MA in V1, and the slower signals through lateral connections in V1 enhance the MA response. The SD of the feedback connections from V2 to V1 is approximately 4 deg in the present model as consistent with physiological studies [20], and the extent of the lateral connections is approximately 2 deg [20]. The results of the simulations revealed strong responses of V1 cells along the MAs of the stimuli that extend up to 4 deg. The feedback spikes tended to reach V1 cells along the MA within a short time window, so that the temporal integration of the spikes induced a strong response from the cells. Such simultaneous arrival of the spikes tended not to occur in regions far from the MA. The lateral connections within V1 do not provide propagation sufficiently fast to reproduce the latency physiologically reported. However, the lateral signals enhance the MA response shortly after. Note that the model did not have any oscillatory or attractor mechanism that facilitates synchronization, nor any specific mechanism that groups particular cells or signals. The spatial distribution of the activity of the model V1 cells along a cross section of a square was similar to that observed physiologically in V1, indicating the plausibility of the model.

We also analyzed the responses to examine quantitatively whether the model truly represented the MA, based on the comparison between the MA obtained from the model and that computed by a mathematical method. It would be more appropriate to compare the activities of the model with those of V1 neurons physiologically recorded. However, the electrophysiological studies were limited to single cells, so that no population response to MA is available. The results of the comparison showed good agreements between the MA obtained from the model and that obtained using the mathematical method. To evaluate quantitatively whether the internal representation of the model captures accurate shape, we reconstructed object shape from the computed MA, and compared the shape of reconstruction with that of the original. Specifically, we reconstructed shape based on the responses of the model cells responding to MA and the distances of the cells from nearby contours, based on a mathematical method.

The results showed fairly correct reconstruction of shape, indicating the model's capability for the representation of surface by the MA.

The model based on the synchronization of BO signals predicted that the construction of surface (figure-ground determination) might be disturbed if contours are presented asynchronously. This prediction seems likely because this phenomenon has been described in the context of common fate, one of the Gestalt factors that determine grouping [21]. Psychophysical studies have reported that asynchronous presentation of object and background is crucial for figure-ground segregation [22,23]. However, it is uncertain whether the asynchronous presentation of object contours affects the perception of the DOF. We performed a psychophysical experiment to investigate the effect of the asynchronous presentation of contour elements in the perception of the DOF, and examined quantitatively whether the perception of the DOF is modulated by the asynchronous presentation of contours. The results showed the dependence of the DOF on the degree of asynchrony. The simulations of the model showed a tendency similar to that of the psychophysics, supporting the crucial role of onset synchronization among BO-selective cells in the formation of MA and surface.

## 2. MODEL

In the present study, we investigated the effects of onset synchronization of BO-selective cells on the construction of an early, primitive representation of shape. To examine the manner via which the signals from BO-selective cells propagate in the early- to intermediate-level visual cortices, we implemented the network model with biologically realistic spiking neurons on the NEURON simulator [24]. Each model cell has the properties of neurons in the visual cortices, including V1 cells and BO-selective cells.

Figure 1 shows schematic illustrations of the connectivity of the model [Fig. 1(a)] and processing flow [Fig. 1(b)]. The model has two layers (V1 and V2) and three connection types (feedforward, feedback, and lateral). The layers are connected recurrently with feedforward and feedback connections. Cells in the V1 layer are also connected via horizontal connections. We assigned the range and the conduction velocity of each connection type based on published physiological data: (1) the horizontal connection is short ( $<0.5$  mm) and slow (0.1 mm/ms) [25], and (2) the feedforward and feedback connections are long (10–15 mm) and fast (3 mm/ms) [26,27]. The difference in onset latency between the model cells in V1 and V2 is 10 ms, given that the conduction velocity of the feedforward and feedback connections is 3 mm/ms. This difference in onset latency is in agreement with the physiological report (9 ms) [27]. The spatial ranges of horizontal and feedback connections were given by Gaussian functions with a standard deviation (SD) of  $2.1^\circ$  and  $3.5^\circ$ , respectively. These ranges correspond to the physiological report that the horizontal connections in V1 are in the range of  $2.47^\circ \pm 0.3^\circ$  and that the feedback connections from V2 to V1 are in the range of  $3.8^\circ \pm 0.6^\circ$  [20]. These connections and velocities of the model realize the biologically detailed temporal properties of V1–V2 networks. The model consists of four functional stages: (i) contrast detection, (ii) determination of the DOF, (iii) integration of DOF signals, and (iv) competition by a winner-take-all mechanism. A description of the model V1

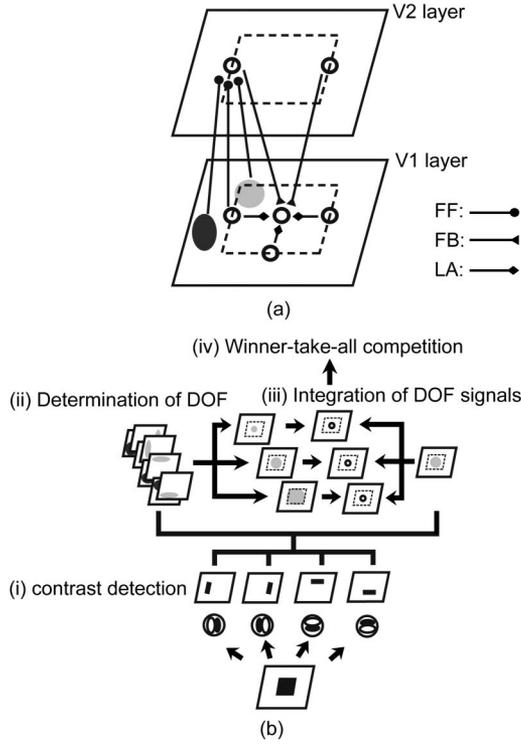


Fig. 1. Schematic illustrations of the model. (a) Model connectivity. The model has two layers (V1 and V2) and three types of connection (feedforward as indicated by FF; feedback, FB; and lateral, LA). The gray ellipses indicate the facilitatory (light gray) and suppressive (dark gray) surrounding regions of a BO-selective cell in V2, respectively, which are projected from V1 (see details in the “Determination of the DOF” section). (b) Processing flow of the model. The model is comprised of four distinct functional stages: (i) contrast detection, (ii) determination of the DOF, (iii) integration of DOF signals, and (iv) competition by winner-take-all mechanism. Luminance contrast of a stimulus is detected by four oriented Gabor filters in V1 layer (i). The DOF of every point on the object contour is computed based on the luminance contrast within CRF of BO-selective neuron and its surround (ii). The DOF signals and contrast signals are integrated via feedback connections by three distinct sizes of integration field, and via lateral connections by single integration field (iii). The responses at each retinotopic position are determined based on the maximum response among the cells (iv).

and V2 cells, with elucidation of each functional stage, is given below, and the mathematical descriptions of the essential parts of the model are given in Appendix A.

### A. Single Model Cell

The morphology of the model cell was approximated to a sphere with a diameter of 23  $\mu\text{m}$  and a resistance of 230  $\Omega \cdot \text{cm}^2$ . To calculate the membrane potential of each cell at a given time, we solved the Hodgkin–Huxley equation [28], as follows:

$$C_m \frac{dv}{dt} = g_{Na}(V - E_{Na}) + g_K(V - E_K) + g_l(V - E_l) + I, \quad (1)$$

where  $C_m$  represents the membrane capacitance;  $g_x$  and  $E_x$  represent the conductance and equilibrium potential, respectively, with the suffixes indicating the type of ion channel:  $Na$ ,  $K$ , and  $l$  represent sodium, potassium, and other ions, respectively; and  $I$  represents the input current. The parameter

Table 1. Parameters for Each Model Cell

Parameter	Value	Ref.
$C_m$	1 ( $\mu\text{F}/\text{cm}^2$ )	[29]
$E_{Na}$	50 (mV)	[29]
$E_K$	-77 (mV)	[29]
$E_l$	-54.3 (mV)	[29]
$g_{Na}$	0.04 ( $\text{S}/\text{cm}^2$ )	modified from [30]
$g_K$	0.012 ( $\text{S}/\text{cm}^2$ )	[30]
$g_l$	0.0001 ( $\text{S}/\text{cm}^2$ )	modified from [30]

values are shown in Table 1 [29,30]. The receptive field of the model cells extends over  $0.75 \times 0.75^\circ$ , without overlap.

### B. V2 Cells: Determination of the DOF

We determined the response of a BO-selective cell in V2 based on the luminance contrast surrounding the CRF. We used surround modulation for the determination of BO [3]. Surround modulation is a phenomenon in which the response of a cell is modulated by the contrast surrounding the CRF [31,32]. Sakai and Nishimura [3] reported that the selectivity to BO emerges from asymmetric surround modulation, in which facilitatory and suppressive regions are located asymmetrically with respect to the CRF. This asymmetrical configuration generates facilitative/suppressive activity of BO-selective cells depending on whether the figure is projected onto preferred DOF of the cell, leading to an assignment of the direction of BO. The activity of a BO-selective cell located at  $x_1, y_1$  and at time  $t$  was given by

$$O^2(x_1, y_1, t) = \text{input}(x_1, y_1) + c \sum_{x,y} \{E(x, y, t-d) + I(x, y, t-d)\}, \quad (2)$$

where  $x, y$  represent the retinotopic position, the first and second terms on the right-hand side of the equation represent the driving current for the CRF and surrounds, respectively (see details in Appendix A),  $c$  represents a weight for surround modulation,  $d$  represents the synaptic delay that is defined by the Euclidean distance between  $(x, y)$  and  $(x_1, y_1)$ , and  $E$  and  $I$  represent excitatory postsynaptic potential (EPSP) and inhibitory postsynaptic potential (IPSP) for surround modulation, respectively. We defined EPSP and IPSP as

$$E(x, y, t-d) = w(v-e)\{\exp(-(t-d)/\tau_{\text{decay}}^{\text{exc}}) - \exp(-(t-d)/\tau_{\text{rise}}^{\text{exc}})\}, \quad (3)$$

$$I(x, y, t-d) = w(v-e)\{\exp(-(t-d)/\tau_{\text{decay}}^{\text{inh}}) - \exp(-(t-d)/\tau_{\text{rise}}^{\text{inh}})\}, \quad (4)$$

where  $w$  is the weight of the synaptic connection defined by a Gaussian function,  $v$  and  $e$  represent the membrane potential and reversal potential, respectively, and  $\tau_{\text{rise}}^{\text{exc}}$  ( $\tau_{\text{rise}}^{\text{inh}}$ ) and  $\tau_{\text{decay}}^{\text{exc}}$  ( $\tau_{\text{decay}}^{\text{inh}}$ ) represent the time constant for the rise and decay of EPSP (IPSP), respectively. Time constants are summarized in Table 2 [29,30].

**Table 2. Time Constants of EPSP and IPSP**

Parameter	Value (ms)	Ref.
$\tau_{\text{rise}}^{\text{exc}}$	0.09	[29]
$\tau_{\text{decay}}^{\text{exc}}$	1.5	[29]
$\tau_{\text{rise}}^{\text{inh}}$	0.1	[30]
$\tau_{\text{decay}}^{\text{inh}}$	50	[30]

### C. V1 Cells: Contrast Detection and Integration of the DOF Signal

The luminance contrast of stimuli is detected in model V1 cells by computing the convolution using four oriented Gabor filters, and the detected contrast is normalized using a sigmoidal function ([33]; for details, see also Appendix A). The activity of the model V1 cells emerged from three sources of EPSPs: (1) feedback signals from BO-selective cells, (2) horizontal signals from V1 cells, and (3) ascending signals that correspond to the normalized luminance contrast. Signals from BO-selective cells were integrated via three different integration fields defined by Gaussian functions (SD of Gaussians = 0.7, 2.1, and 3.5°). Signals from V1 cells were integrated within a smaller Gaussian (SD = 2.1°) to reflect the relatively short range of horizontal connections with respect to the feedforward and feedback connections.

The spikes of BO-selective cells that were synchronized due to the stimulus onset travel from their locations on the contours (in V2) toward the DOF via feedback connections to V1. The spikes arrive (in V1) at locations that are equidistant from nearby contours within a short period because the spikes were synchronized at the contour and traveled the same distance at the same speed. The simultaneous arrivals of the spikes at the equidistant locations are expected to evoke strong activity along MA in V1 [e.g., Fig. 2(b)]. The model did not have any oscillatory mechanism that maintained the synchronization of BO-selective cells so that the synchronization of BO-selective cells was generated solely by stimulus onset. Note that the simultaneous arrival of spikes via V1 lateral connections could also form MA responses if contours were adjacent to each other. However, their contribution is not crucial because the range of the lateral connection is short. The activity of V1 cells that did not respond to the luminance contrast of contour was given by

$$O_{\sigma}^3(x_2, y_2, t) = c \sum_{x,y} [F_{\sigma}(t, D(x_2, y_2, x, y, V2)) + H(t, D(x_2, y_2, x, y, V1))], \quad (5)$$

where  $x_2$  and  $y_2$  represent the retinotopic position of the cell,  $t$  represents the time,  $c$  represents the static weight for gain control, and  $F$  and  $H$  represent the EPSPs via feedback and lateral connections, respectively. We defined  $F$  and  $H$  as

$$F_{\sigma}(t, D(x_2, y_2, x, y, V2)) = w_{\text{feedback}}^{\sigma} E(x, y, t - D(x_2, y_2, x, y, V2)), \quad (6)$$

$$H(t, D(x_2, y_2, x, y, V1)) = w_{\text{lateral}} E(x, y, t - D(x_2, y_2, x, y, V1)), \quad (7)$$

where  $w_{\text{feedback}}^{\sigma}$  and  $w_{\text{lateral}}$  represent the weight for feedback and horizontal connections, respectively, as detailed in

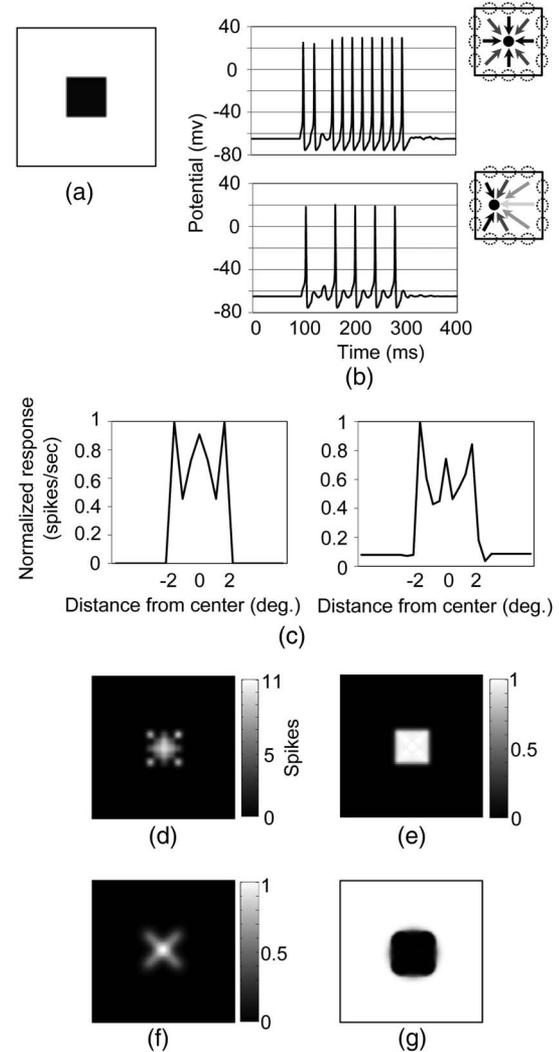


Fig. 2. Simulation results for a single square. (a) Stimulus (a black square) used in the simulation. (b) The black circles in the right panels represent the retinal positions of two example V1 cells. The left panels show the time course of the activities of the example cells. The cell located at the square center (panels on the top) showed a strong response, whereas the cell located away from the center (the bottom) showed a weak response. (c) The spatial distribution of V1 activities along the horizontal midline of the square, showing a distinct peak at the center (the left panel), which is similar to the result of a physiological study (the right; replotted from [5]). (d) The two-dimensional spatial distribution of V1 activities. The activities that respond directly to stimulus contours are not shown. Light and dark colors indicate strong and weak responses, respectively. This shows the MA computed by the model. (e) The number of equidistant contour pairs from given point is normalized, and indicated by color. It showed that strong responses potentially occur elsewhere within the square, but no responses corresponding to the MA were evoked. (f) The MA computed using a mathematical method (the 2D Medial Axis Computation package of MATLAB). The MA computed from the package was passed through Gaussian filter, and normalized. The correlation coefficient between (d) and (f) was 0.91. (g) The shape reconstructed from (d). The reconstruction error was 0.028.

Table 3. The weights decrease as the function of distances between two neurons by the factor defined by Gaussian functions, as similar to [34,35]. In the model, we set the ratio of  $w_{\text{feedback}}^{2.1}$  to  $w_{\text{lateral}}$  to 3:1.  $D$  represents the delay given by the connection type and the Euclidean distance between the cells in question, as defined by

**Table 3. Weights for Feedback and Lateral Connections<sup>a</sup>**

Connection	Value
$w_{\text{feedback}}^{0.7}$	$0.6 \times w_{\text{feedback}}^{2.1}$
$w_{\text{feedback}}^{2.1}$	0.008 or 0.0085
$w_{\text{feedback}}^{3.5}$	$1.5 \times w_{\text{feedback}}^{2.1}$
$w_{\text{lateral}}$	$0.3 \times w_{\text{feedback}}^{2.1}$

<sup>a</sup>Superscripts of  $w$  (0.7, 2.1, and 3.5) represent the SD of integration fields.

$$D(x_2, y_2, x, y, L) = \sqrt{(x_2 - x)^2 + (y_2 - y)^2 + d_{V1,L}^2/v_L}, \quad (8)$$

where  $L$  represents the signals from V1 or V2 (i.e., horizontal or feedback connection);  $d_{V1,L}$  represents the anatomical distance between the V1 layer and the  $L$  layer ( $L$  is either V1 or V2): 0 and 30 for  $d_{V1,V1}$  and  $d_{V1,V2}$ , respectively [26,27]; and  $v_L$  represents the conduction velocity of signals: 0.1 and 3 mm/ms for  $v_{V1}$  and  $v_{V2}$ , respectively. We set *onset latency* of model V1 cells to 70 ms. This onset latency is the time necessary for the signal propagation through the retina and lateral geniculate nucleus (LGN) that is not included in the model. Because physiological studies have reported that the onset latency of V1 cells ranges from 30 to 120 ms depending on the experimental procedures [36], we chose the value close to the middle.

The model V1 cells which represent MA and object boundaries were considered to be the same class of neurons because no physiological studies have reported the cell that responds solely to the MA but not to contour. Huang and Paradiso reported that V1 cells that respond to the MA are selective to object contours as well [6]. They suggested that the responses of the V1 cells were generated through the interactions among cells, whose responses were originally evoked by the contour, via feedforward, feedback, and lateral connections. Their idea is consistent with that of the present model.

#### D. V1 Cells: Winner-Take-All Competition

Three types of V1 cells exist, with different sizes of integration field, as described in the previous section. In this stage, the responses of the three types of cells were integrated using a winner-take-all method [37] to represent a final output at each retinotopic position based on the maximum response among the cells. The final output of the model at a retinotopic position  $x, y$  was given by

$$O(x, y) = \max(S_{0.7}(x, y), S_{2.1}(x, y), S_{3.5}(x, y)), \quad (9)$$

where  $S_\sigma(x, y)$  represents the number of spikes obtained from the previous stage  $O_\sigma^3(x, y, t)$ . We counted a rise in membrane potential exceeding the threshold of 20 mV as one spike.

#### E. Reconstruction of Shape

To evaluate quantitatively the accuracy of the MA computed by the model, we reconstructed a shape from the MA (see Appendix B for details of method), and calculated the reconstruction error. In essence, (1) we found V1 cells that showed the strongest response at each retinotopic location, and determined their spatial extent (SD) of integration field, and (2) we placed a Gaussian with the same SD for each location, and repeated it for all locations. This superimpose

of Gaussians constructed the shape of a surface. The reconstruction error (*Error*) represented the difference between the original, veridical image ( $I$ ), and the reconstructed image ( $RC$ ). We defined the index of the reconstruction error as

$$\text{Error} = \frac{\sum_{x,y} [I(x, y) - RC(x, y)]^2}{\sum_{x,y} [I(x, y) + RC(x, y)]^2}, \quad (10)$$

where  $x, y$  indicates the retinotopic position. The error was normalized using this equation, to range from 0 to 1.

### 3. SIMULATION RESULTS

#### A. Single Square

The present study focused on the onset synchronization of BO-selective cells as a crucial candidate for yielding the early representation of shape. We carried out simulations of the model to examine whether the onset synchronization of BO-selective cells evokes responses of V1 cells that correspond to the MA. First, we tested a simple case of a single square and examined the behavior of the model. Figure 2 shows the activities of V1 cells in response to a single square [Fig. 2(a)]. The time course of the response of an individual cell that is located at the center of the square indicated the presence of strong activity, whereas a cell located away from the center of the square exhibited a weak response, as shown in Fig. 2(b). The distribution of the activities of the cells located along the horizontal midline [Fig. 2(c)] showed a strong peak at the center, which is similar to the physiological data [5]. Note that the physiological experiment used a textured image; thus, the out-of-center regions did not reach a spontaneous level. The two-dimensional distribution of the activities [Fig. 2(d)] revealed strong responses around the locations that were equidistant from nearby contours, indicating the formation of the MA. To confirm that onset synchronization was essential for the construction of the MA, we measured whether BO-selective cells are in fact synchronized. The mean difference in spike timing among all firing BO-selective cells was 0.02 ms with SD of 0.7 ms, indicating that BO-selective cells are synchronized because of the onset synchronization of V1 cells. We also counted the pair of equidistant contour points from given point within the square [Fig. 2(e)]. The result showed that strong responses potentially occur elsewhere within the square, but no responses corresponding to the MA were evoked, indicating that onset synchronization is crucial for the formation of the MA.

To examine whether, in fact, the model yielded the MA, we computed the correlation between the model activity and the MA determined via a mathematical method [using the 2D Medial Axis Computation package of MATLAB; Fig. 2(f)]. The MA computed from the package was passed through Gaussian filter, and normalized to form the mathematical MA. The correlation coefficient between the model activity and the mathematical MA was 0.91, indicating that the model yielded a good representation of the MA. To evaluate further the accuracy of the computed MA in terms of the representation of shape, we reconstructed the shape from the computed MA and compared the reconstructed shape with the original, veridical image. The reconstructed image [Fig. 2(g)] appeared to capture the essence of the original square. The reconstruction error was 0.028, indicating that

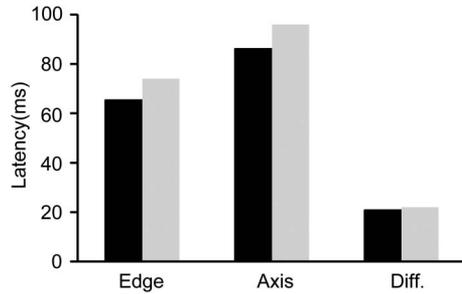


Fig. 3. Comparison of the latencies between the model (black) and physiology (gray; replotted from [6]). Edge and Axis indicate spatial positions of the cells in examination, whose CRF is located on the contours and center of the square, respectively. Diff. indicates the difference in latency between the Edge and Axis, indicating a good agreement between the model and physiology with a constant difference of about 20 ms.

the square was almost perfectly reconstructed from the MA computed using the model.

To evaluate the spatiotemporal characteristics of model V1 cells, we computed the latency of the model cell as the time needed to produce a half-height response of the first peak in post-stimulus time histogram (PSTH) (bin width = 20 ms with 10 times repetition of the simulation). Figure 3 shows the latency of model cells (Fig. 3, black) and real V1 neurons (Fig. 3, gray; [6]) with a CRF located along the edges and at the center of the object. The latency of the model and physiology at both locations differed (65.3 ms versus 74 ms for the edge; 86.1 ms versus 96 ms for the axis), whereas the relative differences between the locations for the model and physiology were quite similar (20.8 ms versus 22 ms). The difference of the absolute latency between the model and physiology depend on the selection of the onset latency of the model V1 cells

(described in the Model section). If we were chosen 10 ms-longer-value for the onset latency, the model and physiology would show perfect matches. The model and physiology showed a good agreement with regard to the relative difference in latency between the locations, supporting that the model captures the essential mechanism of the cortical network on the representation of MA.

## B. Natural Images

The simulation of the model using a single square showed the construction of an accurate MA from the synchronization of model BO-selective cells. Here, we examined whether the model is capable of yielding an MA from a variety of shapes that exist in natural scenes. We performed the simulations of the model using shapes of natural objects whose structures appeared to be different from each other. In this section, we present the simulation results of three examples with an L shape, a rounded shape, and a complex shape including curves, straight lines, and abrupt angles. Figures 4(a)–4(e) show the results obtained for an L-shaped tree branch. The response of the model [Fig. 4(c)] resembled the MA computed using the mathematical method [Fig. 4(d)]. The correlation coefficient of the two approaches was 0.65, suggesting that the model yielded a good representation of the MA. The shape reconstructed from the computed MA [Fig. 4(e)] appeared to be similar to the original shape. The reconstruction error was 0.17, indicating a good reconstruction of the shape. The reconstructed shape expanded toward the outside of the original contours because the shape was constructed solely from the MA, without taking into account the contours. Although the model could have yielded a better shape if it had used contour and/or BO signals, we did not take into account such information because our focus was on the MA emerging from

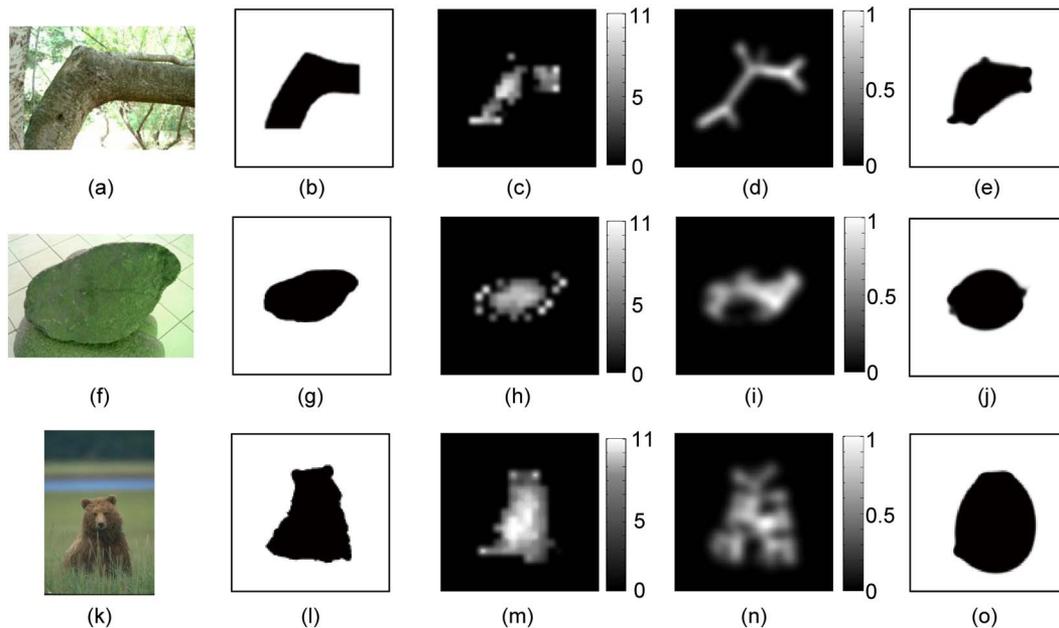


Fig. 4. Simulation results obtained for an L-shaped tree branch (a)–(e), a rounded stone (f)–(j), and a bear ((k)–(o); the natural image of a bear was taken from the Berkeley Segmentation Dataset [38]). The conventions used were the same as those described in Fig. 2. (a), (f), (k) Natural images of the L-shaped tree branch, the rounded stone, and the bear. (b), (g), (l) The binary stimuli used in the simulations. (c), (h), (m) The two-dimensional, spatial distribution of V1 activities, representing the MA. (d), (i), (n) The MA computed using a mathematical method. (e), (j), (o) The reconstructed images from (c), (h), (m), respectively. (a)–(e) The correlation between (c) and (d) was 0.65. The reconstruction error was 0.17. (f)–(j) The correlation between (h) and (i) was 0.76 and the reconstruction error was 0.14. (k)–(o) The correlation of (m) and (n) was 0.78, and the reconstruction error was 0.15.

the onset synchronization. Figures 4(f)–4(j) show the simulation results obtained for a rounded stone. Determining the MA of a rounded object is difficult because a circular shape tends to evoke responses only around its center. The correlation coefficient between the MA computed using the model [Fig. 4(h)] and that computed using the mathematical method [Fig. 4(i)] was 0.76, indicating a good representation of the MA. The reconstructed shape [Fig. 4(j)] appeared to resemble the original stone, with a reconstruction error of 0.14, indicating a fairly good representation of the shape. Figures 4(k)–4(o) show the third example of the simulations, using an image of a bear. The shape of the bear (from the Berkeley Segmentation Dataset; [38]) was challenging in that its contour included curves, straight lines, and abrupt angles. The computed MA [Fig. 4(m)] appeared as a blob, rather than as axes. Because the shape was complex and fairly rounded, the signals from BO-selective cells tended to collide elsewhere. Note that the MA computed using the mathematical method [Fig. 4(n)] was also fairly complex and did not appear as axes. The correlation coefficient between the computed MA and the mathematical MA was 0.78, indicating that the model yielded a good representation of the MA. The reconstructed shape [Fig. 4(o)] lacked the details of the bear, such as the ears and shoulders; however, the reconstruction error was 0.15, indicating a fairly good representation of the overall shape. The present model did not have multiple size scales. Had the model included multiple scales, it would have detected detailed structure as well as overall shape.

As reported in this section, we performed the simulations of the model using a variety of stimuli, including natural images, to investigate the hypothesis that the onset synchronization of BO-selective cells triggers the early representation of shape in VI. As expected from the hypothesis, the model cells in VI that were equidistant from contours showed strong responses, indicating the construction of the MA in VI. The mean correlation between the MAs obtained from the model and those obtained using the mathematical method was  $>0.7$ , indicating a good construction of the MA for a variety of shapes. A square yielded the best accuracy (0.91) among all stimuli, whereas rounded contours seemed to yield lower accuracy, which appears reasonable given that the MA is easily detected from a square but not from a circle. A square also yielded the smallest reconstruction error (0.03) among all stimuli. The other shapes yielded fairly small errors (up to 0.3), indicating a fair representation of shape. The results of the correlation and reconstruction analyses are summarized in Table 4 (see also Appendix C). Given its limited function, such as the use of a single size scale and the absence of contour signal, the model yielded a surprisingly accurate representation of the MA. This result is consistent with the hypothesis that onset synchronization of BO-selective cells plays an important role in the representation of the MA.

### C. Multiple Objects

Natural scenes generally include multiple objects. Here, we examined whether the model was capable of yielding an MA in the presence of multiple objects. The assumption of our model that appears to be most characteristic is the directionality of BO-signal propagation, i.e., the BO signals propagate toward the direction of BO. This directionality enables the model to determine a correct MA in the presence of

**Table 4. Correlation with Mathematical MA and the Reconstruction Error**

Shape	Correlation Coefficient	Error of Reconstruction
Square	0.91	0.028
L-shaped tree branch	0.65	0.17
Rounded stone	0.76	0.14
Bear	0.78	0.15
Two separated squares	0.89	0.028
Overlapping square (occluding)	0.75	0.082
Overlapped square (occluded)	0.69	0.038
Rectangle	0.67	0.14
Triangle	0.60	0.32
U-shaped object	0.73	0.46

occlusion and multiple objects. Given this directionality, BO signals do not propagate toward the outside of the object; thus, a spurious MA does not appear where no object exists. If an object occludes another, the MA of the occluding object emerges correctly, whereas that of the occluded object may not. However, because the occluding edge is not an obstacle to the propagation of BO signals from the occluded object, a reasonable extension of occluded shape may appear. These phenomena appear natural in general. To test such capability of the model, we performed simulations of the model using two types of stimuli: two squares separated from each other [Fig. 5(a)] and two overlapping squares [Fig. 6(a)]. The results of the simulation using two separate squares are shown in Fig. 5. Strong responses were observed inside the squares (figure), and no response was observed between the two squares (background), as shown in Fig. 5(b). The correlation between the computed MA [Fig. 5(b)] and the mathematical MA [Fig. 5(c)] was 0.89, indicating a good correlation between the two approaches. The error of reconstruction [Fig. 5(d)] was 0.028; this indicates that the computed MA represented the stimulus almost perfectly. The results obtained for the two overlapping squares are shown in Fig. 6. Strong responses were observed within the squares. In addition, we examined separately the computed MA for each square.

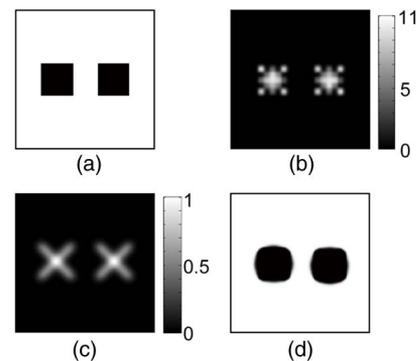


Fig. 5. Simulation results obtained for two separated squares. The conventions used were the same as those described in Fig. 2. (a) The stimulus. (b) Strong responses were observed inside the squares, whereas no response was observed between the two squares. (c) The MA computed using a mathematical method. (d) The reconstructed images. The correlation between (b) and (c) was 0.89 and the reconstruction error was 0.028.

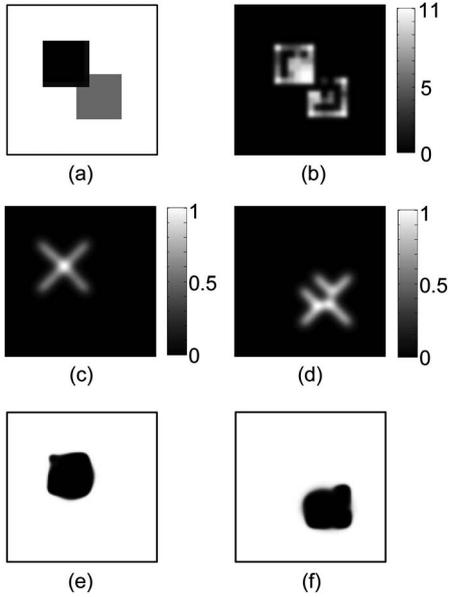


Fig. 6. Simulation results obtained for two overlapping squares. The conventions used were the same as those described in Fig. 2. (a) The stimulus used for the simulation. (b) The two-dimensional, spatial distribution of V1 activities. Strong responses were observed inside the squares. (c) and (d) MAs computed using a mathematical method for occluding (c) and occluded squares (d). The correlations between the mathematical MA and the model MA were 0.75 and 0.69 for the occluding and occluded squares, respectively. (e) and (f) The reconstructed shapes from each model MA of occluding and occluded squares, respectively. Those MAs were plotted together in (b). The reconstruction errors were 0.082 and 0.038 for (e) and (f), respectively.

Because the model does not include a mechanism for grouping the responses, we examined separately the responses of the cells that were located within each square. The MAs obtained using the mathematical method for occluding and occluded squares (on the left and right, respectively) are shown in Figs. 6(c) and 6(d), respectively. The correlation between the mathematical MA and the corresponding computed MA was 0.75 and 0.69 for the occluding and occluded squares, respectively, indicating that the model showed good representation of the MA. The reconstructed shape from each MA is shown in Figs. 6(e) and 6(f). The reconstruction errors were 0.082 and 0.039, respectively, indicating almost perfect reconstruction by the model. These results show that a simple mechanism of directional propagation of BO signals, together with synchronization, is capable of yielding a fairly correct MA, even in the presence of occlusion and multiple objects.

**D. Bias Based on the Degree of Synchronization**

Our model was capable of producing an early representation of shape in the form of the MA. Given that the onset synchronization of BO-selective cells generates the MA, the degree of synchronization appears to be a crucial factor for the representation of shape. Therefore, we thought it would be interesting to examine the presentation of the contours of stimuli in a somewhat asynchronous manner, to decrease the degree of synchronization among BO-selective cells. We carried out simulations of the model using ambiguous figures in which two figures share the border in-between, as shown in Fig. 7(a). To manipulate the degree of synchronization, onset timing of contours of two figures is differentiated. The onset of

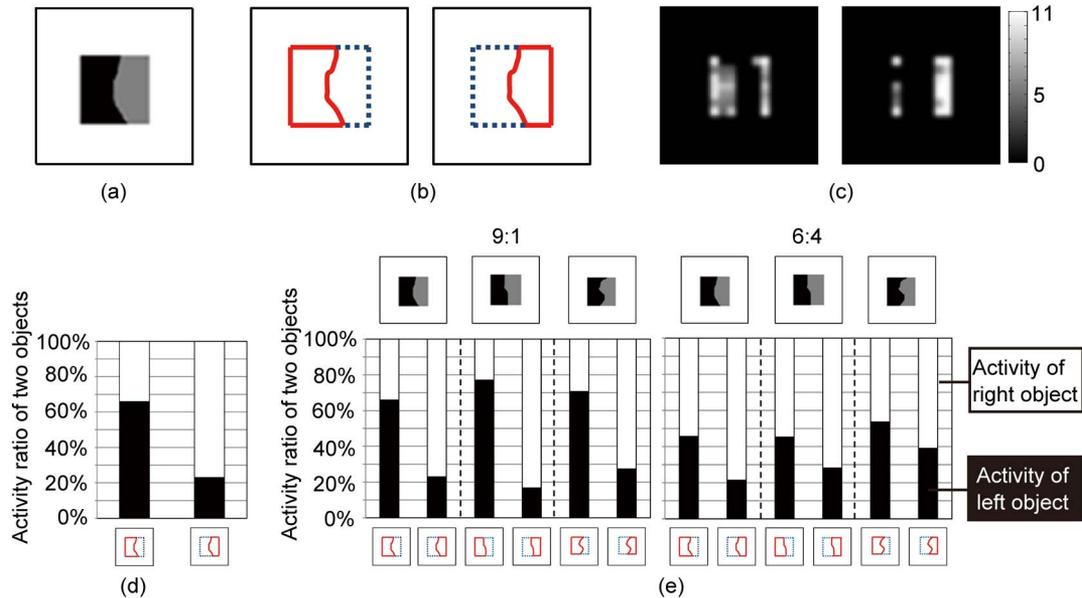


Fig. 7. Simulation results for ambiguous figures. (a) An example stimulus used for the simulation. The stimuli were small patches of natural images from the Berkeley Segmentation Dataset [38]. The two regions divided by a border were filled with dark and light grays. (b) Schematic illustrations of the degree of synchronization. A portion (90% or 60%) of BO-selective cells responded to the border of an ambiguous figure were synchronized with those on the peripheral contours of an either side of the border [the left or right; highly synchronized side is denoted by red (solid) lines], and the rest (10% or 40%) of BO-selective cells were synchronized with those on the opposite side [denoted by blue (dotted) lines]. The simulations were conducted with 12 conditions (stimulus types = 3; highly synchronized side=the left or right; synchronization ratio = 9:1 or 6:4). (c) The spatial distributions of V1 activities when BO-selective cells were mostly (90%) synchronized with the left and right sides of the stimulus are shown in the left and right panels, respectively. The activities appear to be biased toward the direction of stronger synchronization. (d) A quantitative analysis of the bias shown in (c). The total activities within the left (black) and right (white) regions were plotted for the two synchronization conditions [red (solid) icons indicate highly synchronized side]. A bias toward the direction of stronger synchronization was observed. (e) The simulation results for the three stimuli that are shown at the top. The degrees of synchronization were 9:1 and 6:4 for the left and right panels, respectively. The bias was observed in all stimuli used in the simulations, albeit to a lesser degree in the 6:4 condition.

a particular portion (90% or 60%) of the border is equated to one figure (e.g., on the left), and the rest of portion to the other figure (on the right). The portion of BO-selective cells on the border synchronized to the cells on the contour of one figure (left), and the rest of BO-selective cells to those on the other figure (right). It was expected that a lesser degree of synchronization would yield a small difference in response between the objects. We constructed three stimuli from selected patches of natural scenes (from the Berkeley Segmentation Dataset; [38]) with a contour that passed through the center of the patches and appeared ambiguous. The results of simulations using ambiguous figures are shown in Fig. 7. A portion of model BO-selective cells that responded to the central border was synchronized to one side (either left or right), and the remaining model BO-selective cells were synchronized to the alternative side [see Fig. 7(b)]. When the model cells were synchronized mostly (90%) to the left, the responses were biased toward the left of the border, leading to the strong representation of shape on the left side. Similarly, when synchronized mostly to the right, the responses were biased toward the right [Fig. 7(c), left panel]. To examine this bias quantitatively, we summed the responses within the left and right regions for the left- and right-synchronized cases. As seen in Fig. 7(d), a bias toward the synchronized direction was observed clearly. We performed the same tests using two other ambiguous figures. The results of this analysis [Fig. 7(e), left panel] revealed a tendency that was similar to that observed for the first figure. We tested a lesser degree of synchronization (60%) using the same test. The results of this test [Fig. 7(e), right panel] also revealed a bias toward the direction of synchronization, albeit to a lesser degree. This bias of response toward the direction of synchronization, with dependence on the degree of the synchronization, supports the hypothesis that synchronization plays a crucial role in the construction of the MA.

#### 4. PSYCHOPHYSICAL EXPERIMENT: PERCEPTUAL BIAS TOWARD SYNCHRONIZED CONTOURS

The simulation study described in the previous section showed that the onset synchronization modulates the activity of cells responding to the MA. The activity within a more synchronized region was stronger than that within a less synchronized region. This result led to the prediction that the degree of synchronization affects the perception of DOF such that a side with stronger synchronization tends to be perceived as the DOF. Psychophysical studies have suggested that onset asynchrony between object and background is required to segregate figure from background [22,23]. However, it has not been reported whether the degree of synchronization of object contour affects the perception of DOF. Therefore, we developed a psychophysical paradigm for investigating the effect of asynchronous presentation of contour elements in the determination of DOF. We performed psychophysical experiment to examine quantitatively whether the perception of DOF is in fact modulated by the synchrony of contours. We expected that the onset synchronization of BO-selective cells would be degraded if contour segments were shown asynchronously, so that the construction of the MA and the surface would also be weakened. We used the ambiguous figures that were used in the simulations;

however, here they were constructed using blinking dots (dynamically changing luminance) rather than a solid line. We controlled the degree of synchronization by changing the number of dots that blinked in each phase.

#### A. Methods

Figure 8 illustrates the experimental procedure. Stimuli were presented on a liquid crystal display (Mitsubishi Diamond-crysta RDT 197S; response time, 5 ms) using a refresh rate of 70 Hz. A red fixation aid (visual angle,  $0.2 \times 0.2^\circ$ ) was presented at the center of the display for 1500 ms, together with a random mask. The fixation aid was adjusted to the eye level of each participant. After the fixation, a test stimulus (visual angle,  $6.3 \times 6.3^\circ$ ) was presented on a mid-gray background ( $81.85 \text{ cd/m}^2$ ) for 860 ms. Test stimuli consisted of blinking dots that alternated between mid-gray and either black or white ( $315.8$  and  $0.316 \text{ cd/m}^2$ , respectively), as illustrated in Fig. 9. The blinking dots were placed on a border between two regions (border dots), the surrounding contours of a square (outline dots), and elsewhere (noise dots). The size of the outline square was  $4.9 \times 4.9^\circ$ . The size of the dots and the spaces between them was  $0.03 \times 0.03^\circ$ . The dots along the border and outline were aligned in two lines [see Fig. 9(b)]. The outline dots formed a square of  $4 \times 4^\circ$ . At any moment, only half of the dots were displayed, to avoid the appearance of a solid contour [Fig. 9(b)]. All dots blinked (changed their luminance) at the same frequency (7 Hz), but possibly with distinct phases. The degree of synchronization was controlled using the number of dots that shared the same phase of blinking. For instance, at 90% synchronization, nine out of 10 dots blinked with the same timing. Outline dots located on either the left or the right of the border were more synchronized (60% or 90%) with the border dots than with those located on the other side (40% or 10%). The blinking phase of each dot was chosen randomly at every trial. The

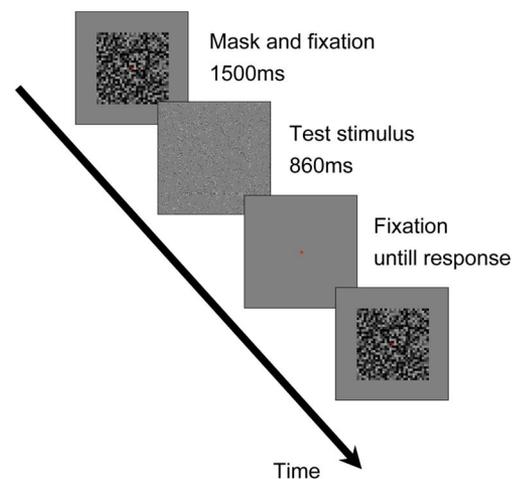


Fig. 8. Schematic diagram of the sequence used for stimulus presentation. A fixation point (a red small dot) was presented at the center of a stimulus for 1500 ms with a random mask. The participants were instructed to remember the position of the fixation point. The test stimulus was presented on a gray background for 860 ms. Participants were asked to judge the direction of figure at the fixation point using two alternative forced choices without the feedback of correct answer. The fixation point on gray background was presented until the participants responded.

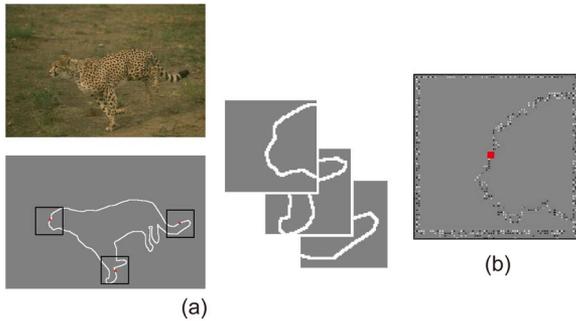


Fig. 9. Configurations of test stimulus. (a) An original image from the Berkeley Segmentation Dataset (the top-left panel; [38]) and the contour of an object (cheetah) drawn by human participants (white line in the bottom-left panel). Small patches (the right panel) were extracted from the contour (denoted by black squares in the lower-left panel). (b) An example of the stimulus dots that were aligned in two lines along the contour that passed through the center of the stimulus and those along the outline square. Noise dots are not shown here for presentation purpose.

noise dots were placed randomly (not overlapped with the border and outline). The luminance of the noise dots was changed at the probability of 50% every 43 ms. To keep the mean luminance identical at any moment, we set one half of the dots as white and the other half as black. The presentation order of the stimuli and the conditions were randomized. Twenty-five trials were repeated for every stimulus and condition. Participants were instructed to remember the position of the fixation aid until the test stimulus disappeared and were asked to indicate the direction of the figure at the fixation point using two-alternative forced choice (left or right) task without feedback regarding the correct answer. Five participants (three men and two women, all in their twenties) who were naïve to this experiment performed the experiment. All participants had normal or corrected-to-normal vision. The experiment was approved by the ethics committee in Graduate School of Systems and Information Engineering, University of Tsukuba.

## B. Results

We performed the experiments using two conditions in which the ratio of synchronization was 6:4 and 9:1 between the two regions, in accordance with the corresponding simulations. The perceived DOF for each condition is shown in Fig. 10. The participants more often perceived a figure in the direction of the region surrounded by dots with a higher synchronization ratio compared with the other direction, which exhibited lower synchronization (pairwise  $t$  test,  $P < 0.01$  for both conditions). This result indicates that synchronization of contours is a crucial factor for the perception of surface. As the ratio of synchronization increased from 6:4 to 9:1, this bias increased slightly. A two-way ANOVA with factors of synchronization and participants showed significance for the two main factors ( $P < 0.01$ ). Although interaction was also significant ( $P < 0.01$ ), all subjects showed the same direction of bias. This result indicates that the perception of surface depends quantitatively on the degree of synchronization. Although the magnitude of the bias appeared smaller than that obtained in the simulations, we observed the same tendency of bias in the psychophysics and simulations. This result is consistent with the hypothesis that onset synchronization

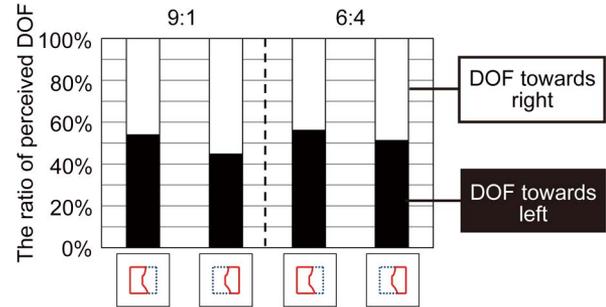


Fig. 10. Results of the psychophysical experiments. The graphs on the left and right show the results obtained for 9:1 and 6:4 synchronization ratios, respectively. The icons placed at the bottom of the graphs show the configuration of stimuli, with red (solid) lines indicating the region of higher synchronization. Black and white bars indicate the ratio of the perceived DOF. In both the 9:1 and 6:4 conditions, the participants more often perceived a figure in the direction of higher synchronization ( $P < 0.01$ ). The magnitude of perceptual bias toward a higher synchronization direction was larger in the 9:1 case than in the 6:4 case ( $P < 0.001$ ).

of BO-selective cells plays a crucial role in the representation of MA and surface.

## 5. DISCUSSION

We investigated computationally whether the onset synchronization of BO-selective cells yields early representation of surface by means of the MA. We constructed a biophysically detailed model of V1-V2 networks that included V1 cells, BO-selective cells, lateral connections, and bidirectional connections between V1 and V2. The simultaneous arrival of signals via feedback connections from BO-selective cells due to onset synchronization evokes strong responses of V1 cells located equidistantly from the contours, generating the representation of the MA. The synchronous firing of BO-selective cells in V2 leads to the simultaneous arrival of the signals via efferent connections to V1 cells that are equidistant from the contours, generating the representation of the MA in V1 with a short latency. Our simulation results showed that the representation of the MA agreed with the results of physiological studies regarding the distribution of activity [5] and latency [6]. Further analysis showed that the mean correlation between MAs obtained from computation using our model and from the mathematical method was 0.74, indicating that the model yields a fairly good representation of the MA. To evaluate the accuracy of the computed MA in terms of the representation of shape, we reconstructed shapes from the computed MA and compared the reconstructed shapes with the original shapes. The reconstructed shapes appeared to be similar to the original shapes. Quantitative analysis showed that the mean reconstruction error was 0.16, indicating a good representation of shape. These results support the hypothesis that the onset synchronization of BO-selective cells plays a crucial role in the construction of the early, primitive representation of surface and shape. The model based on the synchronization of BO signals predicted that the construction of surface (figure-ground segregation) might be disturbed if contours are presented asynchronously. We performed psychophysical experiments to examine quantitatively whether the perception of DOF is modulated by the asynchrony of contours. The results of this analysis showed the dependence of DOF on the degree of synchronization. The simulations

of the model showed the tendency similar to that observed in the psychophysical experiment, supporting the hypothesis that onset synchronization of BO-selective cells plays a crucial role in the representation of MA.

It has been proposed that lateral connections within V1 account for the emergence of the MA [15,39]. Although these models showed the activities related to the MA, they failed to reproduce the short latency of V1 cells. Lee *et al.* reported the latency of V1 cells that responded to the MA was 80–90 ms after stimulus onset [5]. The conduction velocity of the lateral connection is slow (0.1 mm/ms); thus, the lateral connection is not sufficiently fast to establish a latency of 80–90 ms after stimulus onset. The conduction delay through lateral connections between the edge and axis is 60–80 ms, given that (1) the cortical distance between the edge and center of a figure is approximately 6–8 mm (as estimated from the eccentricity of the recording site (3°–4°; [5]), (2) the cortical magnification factor at the eccentricity is 4.16 mm/° [40], and (3) the stimulus size between the edge and the axis is 2° [5]. Because intercortical connections are much faster than lateral connections, a latency of 80–90 ms can be achieved by feedforward/feedback connections, as indicated by our results.

Our simulation results showed that the model was capable of producing the MA of stimulus that extends up to 4°, based on the feedback from V2 to V1 (Fig. 2). This observation is consistent with physiological study: the SD of feedback connections from V2 to V1 is 4° on average [20]. However, Huang and Paradiso have reported physiologically that V1 shows strong response to the disc center even when the diameter of the disc is larger than 10° [6]. This result suggests that V1 receives feedback signals from cortical regions not limited to V2 for the generation of MA. For instance, the extent of feedback from V3 to V1 is twice as large as that from V2 [20]. For the generation of MA to large stimuli, feedback signals from higher cortical regions might need to be taken into account.

The assignment of BO is crucial for the segmentation of figure from ground and the construction of surface representation. Although neural mechanisms underlying the BO selectivity have not been clarified, we have chosen to utilize asymmetric surround modulation for the generation of BO signal [3]. The surround modulation has been reported in a number of visual cortices. Our model captured the essence of the surround modulation reported in V1 cells that has been investigated most precisely [31]. The neural mechanisms underlying the surround modulation have not been clarified; however, its short latency suggests either feedforward or fast feedback. Our model used the feedforward mechanism to establish the surround modulation. Because the latency of the surround modulation is similar for both cases, if we used the feedback mechanism, the simulation results of forming MA from the synchronization of BO-selective cells would not be altered. A model based on the feedback has been reported by Craft *et al.* [41]. They proposed Grouping cell in V4 whose circular surround structure pooled the afferent signals from certain BO-selective cells, and feedback to the BO cells to reproduce the selectivity. Although the idea of feedback appears to provide useful suggestion, the Grouping cell has not been reported physiologically. A more recent, computational study has suggested that competition at V4 level (with another type of imaginary cell) would further refine the BO

selectivity [42]. Further modeling study, as well as physiological study on the neural mechanisms underlying the BO selectivity, is expected to advance the understanding of surface representation.

In the present study, we evaluated the computed MA in comparison with that obtained by the mathematical method. It would be more appropriate if we could compare the computed MA with physiological quantity. Although electrophysiological studies have shown neural activities corresponding to the MA in several cortical regions (e.g., [5,10]), they were limited to single-cell responses. We need population activity to know the true cortical representation of MA that is not limited to a single point on the axis. Our model showed good agreements with the physiology in terms of single-cell responses. A comparison of the model with the population activity is desired. It is greatly expected to perform simultaneous recording of multiple cells with regard to the representation of the MA.

Our psychophysical experiment showed the perceptual bias toward a more synchronized side in the perception of DOF, indicating a good agreement with the simulation results. This phenomenon can be explained as the result of binding contour components by common fate that is known as a Gestalt factor of grouping [21–23]. The participants tended to perceive figure if its contour components blinked more coherently (changed commonly the brightness of the dots) than the other figure. Although the neural mechanism underlying common fate has not been clarified, our model may provide insight into the mechanism of common fate. Specifically, the synchronization of feedback signals from higher cortical regions could be an excellent indicator for assessing coherency of the components.

We hypothesized that the simultaneous onset of stimulus contours generates the synchronized response of BO-selective cells. The model did not include any oscillatory or attractor mechanism that maintains synchronization because little is known about oscillation of BO-selective cells. A recent physiological study has reported that BO-selective cells show synchronized activity when they are stimulated by one figure with attention (e.g., A. Martin and R. von der Heydt, program no. 724.04, Annual Meeting of Society for Neuroscience 2012), and their results could also be considered as an indication of oscillation among the cells. Although the neural mechanism underlying the synchronization has not been reported, the stimulus onset, oscillation and attraction are plausible candidates for producing the synchronized activities of BO-selective cells. Simultaneous onset of a stimulus makes the synchronization of BO-selective cells at their first spikes, and oscillatory mechanisms sustain the synchronization of spikes afterward. The onset synchronization and the oscillatory mechanisms work on different time scales in general. If, in fact, oscillatory mechanisms sustain the synchronization of BO-selective cells, the model is expected to include a mechanism for oscillation in addition to the present network mechanisms [43,44].

Our model showed synchronization as long as the duration of simulation. A reason could be fixed values of parameters (Table 1) and a single type of surround modulation for the model BO-selective cells. Recent physiological studies (e.g., [45,46]) have reported that synchronization of BO-selective cells lasts 1000 ms after stimulus onset mostly in 20 Hz band.

The behavior of the model BO-selective cells is consistent with this finding rather than the previous study reporting onset synchronization [17]. Further analysis is needed to investigate the effects of the variability of the parameters and the surround modulation of the model cells.

An alternative candidate for generating the synchronization among BO-selective cells is feedback signals from higher cortical regions such as V4. Synchronization by feedback signals may provide the MA similar to that reported in this study. An important advantage of the present model is that the model reproduces the MA response in V1 without an assumption of the feedback from higher cortical areas. For instance, an assumption of Grouping cell in V4 may be attractive computationally [41,42], however, such cell has not been reported physiologically. Accumulating evidences have suggested that simultaneous stimulation of cells in fact affects the degree of synchronization [18,19]. Zhou *et al.* have reported that the synchronization of V1 cells decreases as the continuity of contour decreases [19], supporting that simultaneous stimulation causes the synchronization among BO-selective cells. We consider that the onset synchronization is a plausible candidate for generating the synchronization of BO-selective neurons.

Given that the response to the MA is generated in V1, it is still an open question that what cortical regions receive the MA responses and construct the neural representation of shape. We developed the model of V1-V2 networks, and showed that the model constructs the MA responses in V1. The obtained MA responses in V1 appear to proceed to higher cortical regions, and generate the cortical representation of shape. In fact, IT neurons are reported to encode the shape of an object with the configuration of several axial components (MA) and surfaces [10]. These results suggest that V1 provides the local, primitive MA to the higher cortical regions.

## APPENDIX A: CONTRAST DETECTION

A mathematical description of contrast detection is shown here. This computation mimicked the characteristics of a simple cell in the V1 layer. Luminance contrast was detected by convolution of four oriented Gabor filters and stimuli. The detected contrast at retinal position  $(x_0, y_0)$  is given by the following equation:

$$T_\theta(x_0, y_0) = \begin{cases} (I_0 * G_\theta)(x, y) & (\text{if } T_\theta > 0) \\ 0 & (\text{otherwise}) \end{cases}, \quad (\text{A1})$$

where  $*$  represents convolution and  $I_0$  represents the element of input stimulus at  $(x_0, y_0)$ . To mimic the nonlinear property of actual V1 neurons, the contrast detected was passed through a sigmoidal function.

## APPENDIX B: RECONSTRUCTION ALGORITHM

An algorithm for reconstruction of a shape is described here. The reconstruction of a shape was achieved using the following steps. First, we defined the cells that responded to the MA:

$$M(x, y) = \begin{cases} 1 & (\text{if } t_{\text{resp}} \geq t_{\text{threshold}}) \\ 0 & (\text{otherwise}) \end{cases}, \quad (\text{B1})$$

where  $t_{\text{resp}}$  is the latency of a V1 cell and  $t_{\text{threshold}}$  represents the threshold latency of V1 cells that responded to the MA. We

**Table 5. Weights for Reconstruction<sup>a</sup>**

SD	Weight
$w_{0.7}$	0.6
$w_{2.1}$	1.0
$w_{3.5}$	1.5

<sup>a</sup>Subscripts of  $w$  (0.7, 2.1, and 3.5) represent the SD of integration fields.

used  $t_{\text{threshold}} = 77$  ms. The threshold was determined by the onset latency of model V1 cells (70 ms) and the latency of V1 cells that respond to MA (86 ms; Fig. 3). Second, we found the V1 cell that showed the strongest response at each spatial position  $(x, y)$ , and obtained the number of spike and SD of the integration field of the cell:

$$N(x, y) = \max_\sigma(S_{0.7}(x, y), S_{2.1}(x, y), S_{3.5}(x, y)), \quad (\text{B2})$$

$$\sigma(x, y) = \begin{cases} \arg \max_\sigma(S_{0.7}(x, y), S_{2.1}(x, y), S_{3.5}(x, y)) & (\text{if } M(x, y) = 1) \\ 0 & (\text{otherwise}) \end{cases}, \quad (\text{B3})$$

where  $\sigma$  represents the SD of the integration field of V1 cells at that spatial position. Third, we superimposed the weighted ( $w_\sigma$ ; Table 5) Gaussian functions with a parameter defined by the center position  $(x_1, y_1)$  and SD ( $\sigma(x_1, y_1)$ ):

$$T(x, y) = \sum_{x_1, y_1} N(x_1, y_1) \times \text{gauss}_{x_1, y_1}(x, y), \quad (\text{B4})$$

$$\text{gauss}_{x_1, y_1}(x, y) = \frac{w_\sigma(x_1, y_1)}{2 \pi G(x_1, y_1)} \exp\left(-\frac{(x - x_1)^2 + (y - y_1)^2}{\sigma(x_1, y_1)^2}\right). \quad (\text{B5})$$

Finally, we passed  $T$  through the sigmoidal function and obtained the reconstructed image ( $RC$ ), as follows:

$$RC(x, y) = \frac{1}{1 + \exp(-(T(x, y) - \text{threshold})\text{slope})}, \quad (\text{B6})$$

where the *threshold* and *slope* represent the origin and steepness of the sigmoidal function, respectively. We used the following settings:  $\text{slope} = 300$  and  $\text{threshold} = 0.3 \times \text{MAX}$  (MAX means the maximum value of the  $RC$ ), with the exception of the rounded stone, for which  $\text{threshold} = 0.4 \times \text{MAX}$ , and the L-shaped tree branch and U-shaped object, for which  $\text{threshold} = 0.5 \times \text{MAX}$ .

## APPENDIX C: SIMULATION RESULTS FOR OTHER STIMULI

In this section, we show the simulation results obtained for three additional stimuli (a rectangle, an equilateral triangle, and a U-shaped object). All results are shown in Fig. 11.

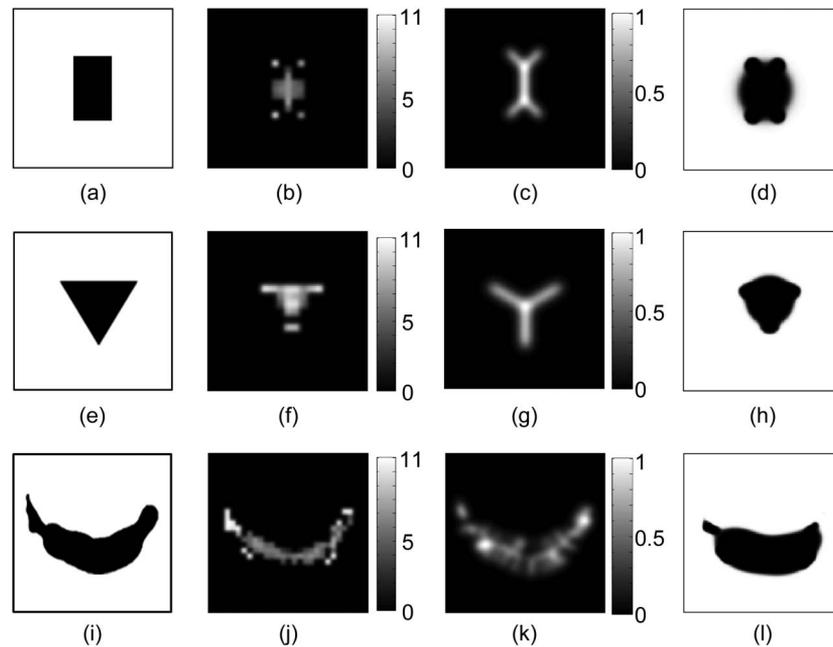


Fig. 11. Simulation results for three additional stimuli. The conventions used were the same as those described in Fig. 2. (a), (e), (i) The binary stimuli used for the simulations. (b), (f), (j) The spatial distributions of V1 activities. (c), (g), (k) MA computed from a mathematical method. (d), (h), (l) The reconstructed shapes from the MA computed by the model. (The top row) The simulation results for the rectangle. The correlation coefficient of (b) and (c) was 0.67 and the reconstruction error of (d) was 0.14. (The middle row) The simulation results for the triangle. The correlation coefficient of (f) and (g) was 0.60 and the reconstruction error of (h) was 0.32. (The bottom row) The correlation coefficient of (j) and (k) was 0.73 and the reconstruction error of (l) was 0.46.

## ACKNOWLEDGMENTS

We thank M. Sumi for data collection in the psychophysical experiment. This work was supported by grant-in aids from Japan Society for the Promotion of Science (KAKENHI 22300090, 24.3368), and the Ministry of Education, Culture, Sports, Science and Technology of Japan (23135503, 25135704 (Shitsukan)).

## REFERENCES

- H. Zhou, H. S. Friedman, and R. von der Heydt, "Coding of border ownership in monkey visual cortex," *J. Neurosci.* **20**, 6594–6611 (2000).
- S. H. Kim and J. Feldman, "Globally inconsistent figure/ground relations induced by a negative part," *J. Vis.* **9**(10):8, 1–13 (2009).
- K. Sakai and H. Nishimura, "Surrounding suppression and facilitation in the determination of border ownership," *J. Cogn. Neurosci.* **18**, 562–579 (2006).
- N. R. Zhang and R. von der Heydt, "Analysis of the context integration mechanisms underlying figure-ground organization in the visual cortex," *J. Neurosci.* **30**, 6482–6496 (2010).
- T. S. Lee, D. Mumford, R. Romero, and V. A. F. Lamme, "The role of the primary visual cortex in higher level vision," *Vis. Res.* **38**, 2429–2454 (1998).
- X. Huang and M. A. Paradiso, "V1 response timing and surface filling-in," *J. Neurophysiol.* **100**, 539–547 (2008).
- V. A. F. Lamme, "The neurophysiology of figure-ground segregation in primary visual cortex," *J. Neurosci.* **15**, 1605–1615 (1995).
- K. Zipser, V. A. F. Lamme, and P. H. Schiller, "Contextual modulation in primary visual cortex," *J. Neurosci.* **16**, 7376–7389 (1996).
- M. D. Lescroart and I. Biederman, "Cortical representation of medial axis structure," *Cereb. Cortex* **23**, 629–637 (2013).
- C. C. Hung, E. T. Carlson, and C. E. Connor, "Medial axis shape coding in macaque inferotemporal cortex," *Neuron* **74**, 1099–1113 (2012).
- I. Kovacs and B. Julesz, "Perceptual sensitivity maps within globally defined visual shapes," *Nature* **370**, 644–646 (1994).
- I. Kovacs, A. Feher, and B. Julesz, "Medial-point description of shape: a representation for action coding and its psychophysical correlates," *Vis. Res.* **38**, 2323–2333 (1998).
- Y. Hatori and K. Sakai, "Robust detection of medial-axis by onset synchronization of border-ownership selective cells and shape reconstruction from its medial-axis," *Lect. Notes Comput. Sci.* **5506**, 301–309 (2009).
- D. Marr and H. K. Nishihara, "Representation and recognition of the spatial organization of three-dimensional shapes," *Proc. R. Soc. Lond. B, Biol. Sci.* **200**, 269–294 (1978).
- B. B. Kimia, "On the role of medial geometry in human vision," *J. Physiol. Paris* **97**, 155–190 (2003).
- V. Froyen, J. Feldman, and M. Singh, "A bayesian framework for figure-ground interpretation," *Adv. Neural Inf. Process Syst.* **23**, 631–639 (2010).
- Y. Dong, S. Mihalas, F. Qiu, R. von der Heydt, and E. Niebur, "Synchrony and the binding problem in macaque visual cortex," *J. Vis.* **8**(7):30, 1–16 (2008).
- J. M. Samonds and A. B. Bonds, "Gamma oscillation maintains stimulus structure-dependent synchronization in cat visual cortex," *J. Neurophysiol.* **93**, 223–236 (2005).
- Z. Zhou, M. R. Bernard, and A. B. Bonds, "Deconstruction of spatial integrity in visual stimulus detected by modulation of synchronized activity in cat visual cortex," *J. Neurosci.* **28**, 3759–3768 (2008).
- A. Angelucci, J. B. Levitt, E. J. S. Walton, J. M. Hupe, J. Bullier, and J. S. Lund, "Circuits for local and global signal integration in primary visual cortex," *J. Neurosci.* **22**, 8633–8646 (2002).
- A. B. Sekuler and P. J. Bennett, "Generalized common fate: grouping by common luminance changes," *Psychol. Sci.* **12**, 437–444 (2001).
- P. J. Hancock, L. Walton, G. Mitchell, Y. Plenderleith, and W. A. Phillips, "Segregation by onset asynchrony," *J. Vis.* **8**(7):21, 1–21 (2008).
- M. Usher and N. Donnelly, "Visual synchrony affects binding and segmentation in perception," *Nature* **394**, 179–182 (1998).
- M. L. Hines and N. T. Carnevale, "The NEURON simulation environment," *Neural Comput.* **9**, 1179–1209 (1997).
- V. Bringuier, F. Chavane, L. Glaeser, and Y. Fregnac, "Horizontal propagation of visual activity in the synaptic integration field of area 17 neurons," *Science* **283**, 695–699 (1999).

26. P. Girard, J. M. Hupe, and J. Bullier, "Feedforward and feedback connections between areas V1 and V2 of the Monkey have the similar rapid conduction velocities," *J. Neurophysiol.* **85**, 1328–1331 (2001).
27. L. G. Nowak, M. H. Munk, P. Girard, and J. Bullier, "Visual latencies in areas V1 and V2 of the macaque monkey," *Vis. Neurosci.* **12**, 371–384 (1995).
28. A. L. Hodgkin and A. F. Huxley, "A quantitative description of membrane current and its application to conduction and excitation in nerve," *J. Physiol.* **117**, 500–544 (1952).
29. W. Gerstner and W. Kistler, *Spiking Neuron Models: Single Neurons, Populations, Plasticity* (Cambridge University, 2002).
30. K. A. Archie and B. W. Mel, "A model for intradendritic computation of binocular disparity," *Nat. Neurosci.* **3**, 54–63 (2000).
31. H. E. Jones, W. Wang, and A. M. Sillito, "Spatial organization and magnitude of orientation contrast interactions in primate V1," *J. Neurophysiol.* **88**, 2796–2808 (2002).
32. T. S. Meese, R. J. Summers, D. J. Holmes, and S. A. Wallis, "Contextual modulation involves suppression and facilitation from the center and the surround," *J. Vis.* **7**(4):7, 1–21 (2007).
33. M. Carandini, D. J. Heeger, and J. A. Movshon, "Linearity and normalization in simple cells of the macaque primary visual cortex," *J. Neurosci.* **17**, 8621–8644 (1997).
34. G. Deco and E. T. Rolls, "A neurodynamical cortical model of visual attention and invariant object recognition," *Vision Res.* **44**, 621–642 (2004).
35. T. Poggio and F. Girosi, "Regularization algorithm for learning that are equivalent to multilayer networks," *Science* **247**, 978–982 (1990).
36. L. G. Nowak and J. Bullier, "The timing of information transfer in the visual system," in *Cerebral Cortex*, K. S. Rockland, J. H. Kaas, and A. Peters, eds. (Plenum, 1997), Vol. **12**, pp. 205–233.
37. E. T. Rolls and G. Deco, *Computational Neuroscience of Vision* (Oxford University, 2002).
38. C. C. Fowlkes, D. R. Martin, and J. Malik, "Local figure-ground cues are valid for natural images," *J. Vis.* **7**(8):2, 1–9 (2007).
39. L. Zhaoping, "V1 mechanisms and some figure-ground and border effects," *J. Physiol. Paris* **97**, 503–515 (2003).
40. A. Cowey and E. T. Rolls, "Human cortical magnification factor and its relation to visual acuity," *Exp. Brain Res.* **21**, 447–454 (1974).
41. E. Craft, H. Schutze, E. Niebur, and R. von der Heydt, "A neural model of figure-ground organization," *J. Neurophysiol.* **97**, 4310–4326 (2007).
42. O. W. Layton, E. Mingolla, and A. Yazdanbakhsh, "Dynamic coding of border-ownership in visual cortex," *J. Vis.* **12**(13):8, 1–21 (2012).
43. L. F. Abbott, "A network of oscillators," *J. Phys. A* **23**, 3835–3859 (1990).
44. X. J. Wang and G. Buzsaki, "Gamma oscillation by synaptic inhibition in a hippocampal interneuronal network model," *J. Neurosci.* **16**, 6402–6413 (1996).
45. A. Martin and R. von der Heydt, "Contour binding and selective attention increase coherence between neural signals in visual cortex," *Perception* **40**, 49 (2011).
46. A. Martin and R. von der Heydt, "Binding and selective attention increase coherence between distant sites in early visual cortex," *J. Vis.* **11**(11), 179 (2011).