A border-ownership model based on computational electromagnetism

doi: 10.1016/j.neunet.2017.12.004

Title

A border ownership model based on computational electromagnetism

Author name and Affiliations

Zaem Zainal Abidin[a], Shunji Satoh[b]

Laboratory for Human Informatics,
Graduate School of Information Systems,
The University of Electro-Communications.
Chofugaoka 1-5-1, Chofu, Tokyo 182-8585, Japan.

[a]zaem@hi.is.uec.ac.jp
[b]shunji@uec.ac.jp

Corresponding author

Shunji Satoh

Abstract

The mathematical relation between a vector electric field and its corresponding scalar potential field is useful to formulate computational problems of lower/middle-order visual processing, specifically related to the assignment of borders to the side of the object: so-called border ownership (BO). BO coding is a key process for extracting the objects from the background, allowing one to organize a cluttered scene. We propose that the problem is solvable simultaneously by application of a theorem of electromagnetism, i.e., "conservative vector fields have zero rotation, or "curl." We hypothesize that (i) the BO signal is definable as a vector electric field with arrowheads pointing to the inner side of perceived objects, and (ii) its corresponding scalar field carries information related to perceived order in depth of occluding/occluded objects. A simple model was developed based on this computational theory. Model results qualitatively agree with object-side selectivity of BO-coding neurons, and with perceptions of object order. The model update rule can be reproduced as a plausible neural network that presents new interpretations of existing physiological results. Results of this study also suggest that T-junction detectors are unnecessary to calculate depth order.

Keywords: Neural computation, Middle vision, Depth perception, Perceptual organization, Receptive field

1. Introduction

Humans can perceive objects and their relative order from a cluttered scene filled with objects that are mutually occlusive. In the real world, when one object overlaps another, the border one sees between the two objects is attributed to the occluding object. Assigning borders to the perceived object is believed to be a key process of perception; it facilitates distinction of objects from the background (Nakayama, Shimojo, & Silverman, 1989). This process is called border ownership (BO) coding. One can consider an ambiguous figure that can be perceived as either a vase or two faces (**Figure 1a**). In the perception of a vase, the border is assigned to the inner white area, which is perceived as closer to the viewer (**Figure 1b**). Similarly, in the perception of two faces, the border is assigned to the outer black area, which is perceived as closer to the viewer (**Figure 1c**). This example illustrates how BO coding and the perception of depth order are closely related.

Neurophysiological experiments using simple shaded stimuli on areas V1, V2, and V4 in the visual cortex of monkeys revealed the existence of neurons which respond more strongly to an edge when the "owner" of the edge is located on a specific side (Zhou, Friedman, & von der Heydt, 2000). Spatial selectivity for the owner's location was found to emerge in neurons of area V2 in the visual cortex. Border ownership (BO) is believed to be coded by calculating the difference in response of these sets of neurons with opposite object-side preference towards a stimulus.

Various models have been proposed to code BO as neural networks in area V2 (Craft, Schütze, Niebur, & von der Heydt, 2007; Li, 2005; Sakai, Nishimura, Shimizu, & Kondo, 2012). The model explained by Li (2005) requires more than 20 free parameters to function well. Craft and his colleagues suggest hypothetical "grouping cells," of which the receptive fields have an annular distribution (Craft et al., 2007). The model produced by Sakai et al. relies on randomly generated feedforward neural connections (Sakai et al., 2012). These studies reproduce the responses of BO-related neurons, but they are fundamentally flawed in their approach to elucidating visual systems from computational viewpoints. It is quite difficult to elucidate what sort of mathematical problem those models are intended to solve.

No mathematical formulation of the BO problem has been proposed to date. However, many problems of early vision can be formulated using regularization theory (Poggio, Torre, & Koch, 1985). Neural network models to solve these problems are drawn from the theory. Similarly, if the BO coding problem is expressed in a mathematically well-defined fashion, then one might be able to deduce a neural network model for the mathematical problem, and in turn understand the characteristics of the neural connections involved. Consequently, we might also, for example, reduce the number of parameters in Li's model.

A fundamental question must be considered: what is the visual system trying to achieve through BO coding? As shown in **Figure 1**, BO coding enables the visual cortex to calculate a rough reproduction of the outer world information, which includes depth order in addition to observable

two-dimensional (2D) information such as surface brightness and edges. A further study of neurons in V2 revealed the existence of neurons that respond selectively to the edges of objects with actual stereoscopic depth information (Qiu & von der Heydt, 2005). Some of these neurons show selectivity to both 2D figures (owners) as well as three-dimensional (3D) figures placed on a certain side of their receptive field. This result reinforces the idea that area V2 might be involved in the process of a rough reproduction of a 3D image, and that BO coding and depth coding are interdependent processes.

As described in this paper, we attempt to overcome the flaws of current studies by formulating BO coding as a well-defined mathematical problem and subsequently solving a middle-level visual task: estimation of the depth order of objects from 2D images. We hope that formulation and solutions related to BO coding and order estimation presented in this paper will serve as the basis for a comprehensive model that can aid in the elucidation of low/middle-level visual processing. The remainder of this paper is organized as follows: Section 2 presents discussion of the key ideas underpinning our theory and proposes a neural network model for BO coding. Section 3 presents results of numerical simulations of our model. Section 4 emphasizes some important findings derived from simulation results, including the deduced neural connections. Section 5 is the conclusion.

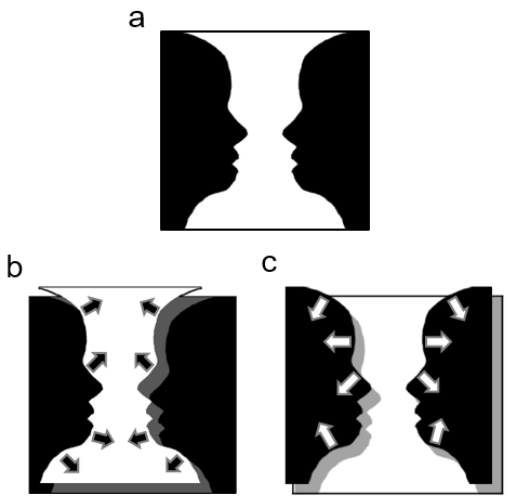

Figure 1. Image presenting the phenomenon of figure-ground organization. (a) From the 2D image on the left, humans can interpret the image as a vase or as two faces. (b) When the image is perceived as a vase, the border is perceived as belonging to the white region, which is regarded as a region closer to the viewer. (c) Similarly, when the image is perceived as two faces, the border is perceived as belonging to the black region, which is regarded as a region closer to the viewer. Arrows indicate the direction of the object the region owns. This tendency demonstrates BO coding as a key process in perception.

1

2. Method

2.1 Computational theory

The relation between depth order and BO must be reviewed before formulating the BO assignment problem. In an inset of **Figure 2**, a light gray rectangle overlapping a dark gray rectangle, over a dark background is visible. This perception of "occluding" or "occluded" objects can be represented using a scalar field signifying depth order. We define the depth order as $\phi(x, y)$. A region with a larger depth order of ϕ coincides with its perception as closer to the viewer. The occluding rectangle has a larger value ($\phi = 2$) than the occluded rectangle ($\phi = 1$). The gradient of depth order $\phi(x, y)$, calculated using spatial differentiation, results in a two-dimension vector field presented in **Figure 2**, which we define as $\boldsymbol{E}(x, y) = (E_x(x, y), E_y(x, y))^T$. The arrowheads of these vectors face the object. The gradient of a depth-order scalar field resembles BO signals.

A core concept in electromagnetism, a mainstream theory in theoretical physics, is the model which describes the following relation: the gradient of a scalar electric potential $\phi(x, y)$ is equal to its corresponding electric field $\boldsymbol{E}(x, y)$. By analogy, likening BO signals to the electric vector field, and likening order information of occluding/occluded objects to a scalar electric potential, one might infer that the theorem in electromagnetism shown below merely describes the relation between depth order and BO signals.

$$\boldsymbol{E}(x, y) = \boldsymbol{\nabla}\phi(x, y). \tag{1}$$

Therein, $\boldsymbol{\nabla}$ is the gradient operator $\boldsymbol{\nabla} \stackrel{\text{def}}{=} (\partial/\partial x, \partial/\partial y)$.

We then proceed to solve the equation. At the stimulus onset, both vector field $\boldsymbol{E}(x, y)$ and scalar field $\phi(x, y)$ are unknown values. In other words, one must provide initial values and update equations for these values so that the relation in **Equation (1)** is satisfied. For this study, we approach the problem by investigating the requirements of vector field $\boldsymbol{E}(x, y)$ for **Equation (1)** to hold true. Consequently, scalar field $\phi(x, y)$ will be treated as a variable that is dependent on $\boldsymbol{E}(x, y)$.

For **Equation (1)** to hold true, $\boldsymbol{E}(x, y)$ must be a "conservative field." By definition, the result of line integration of $\boldsymbol{E}(x, y)$ between any two points in a conservative field is the same irrespective of the path taken. The integration is expected be equal to $\phi(x, y)$. A mathematical theorem regarding conservative fields holds that the "curl," or rotation, of a vector field $\boldsymbol{E}(x, y)$ must be zero at all spatial points (x, y).

A short explanation of "curl" operation can elucidate this point. Calculating the "curl" at a specific point in a vector field is similar to observing the rotation of a windmill at the respective cross marks in **Figure 3.** Wind blowing a windmill is expressed as $\boldsymbol{E}$. If one observes the clockwise (counter-clockwise) rotation of the windmill, then we define the "curl" value as positive (negative).

Two examples in **Figures 3a, 3b** have a non-zero "curl" at the cross mark because a windmill would rotate in either a counter-clockwise (positive curl) or a clockwise direction (negative curl). However, the "curl" at the cross mark in **Figures 3c, 3d** is zero. Mathematically, the "curl" of a two-dimensional vector field $\boldsymbol{E}(x, y)$ is defined as the difference of partial differentiation on its components E_x and E_y, as shown in **Equation (2)** below.

$$\mathrm{curl}(\boldsymbol{E}) = \frac{\partial}{\partial x} E_y - \frac{\partial}{\partial y} E_x. \tag{2}$$

Using the theorem of zero "curl," we construct an energy function for minimization.

$$J[\boldsymbol{E}] = \frac{1}{2} \iint (\mathrm{curl}(\boldsymbol{E}))^2 \, dxdy \rightarrow \min. \tag{3}$$

Integration is conducted for the region of the image. Applying the steepest descent method on energy function $J[\boldsymbol{E}]$, we obtain an update rule for a two-dimensional vector field changing in time $\boldsymbol{E}(x, y, t)$ as well as individual update rules for E_x and E_y.

$$\frac{\partial \boldsymbol{E}(x,y,t)}{\partial t} = \frac{\partial}{\partial t}\begin{pmatrix} E_x \\ E_y \end{pmatrix} \propto \begin{pmatrix} \frac{\partial^2}{\partial y^2} E_x - \frac{\partial^2}{\partial x \partial y} E_y \\ -\frac{\partial^2}{\partial x \partial y} E_x + \frac{\partial^2}{\partial x^2} E_y \end{pmatrix} = \begin{pmatrix} \frac{\partial^2}{\partial y^2} & -\frac{\partial^2}{\partial x \partial y} \\ -\frac{\partial^2}{\partial x \partial y} & \frac{\partial^2}{\partial x^2} \end{pmatrix} \begin{pmatrix} E_x \\ E_y \end{pmatrix}. \tag{4}$$

It is noteworthy that t represents the time step required for $\boldsymbol{E}(x, y, t)$ to reach its steady state $\boldsymbol{E}(x, y)$ by the steepest descent method. One can express the update rule using only $\mathrm{curl}(\boldsymbol{E})$ and the perpendicular of the gradient operator $\boldsymbol{\nabla}^{\perp} \stackrel{\text{def}}{=} (-\partial/\partial y, \partial/\partial x)$ as demonstrated in **Equation (5)** below.

$$\frac{\partial \boldsymbol{E}(x,y,t)}{\partial t} \propto \begin{pmatrix} -\frac{\partial}{\partial y}\mathrm{curl}(\boldsymbol{E}) \\ \frac{\partial}{\partial x}\mathrm{curl}(\boldsymbol{E}) \end{pmatrix} = \boldsymbol{\nabla}^{\perp}\mathrm{curl}(\boldsymbol{E}). \tag{5}$$

2.2 Model description for numerical simulation

To simulate **Equation (4)** to 2D image data distributed in discretized (x, y), differential operators are replaced with the Gaussian derivative (Lindeberg, 1994).

$$G_\sigma(x,y)=\frac{1}{2\pi\sigma^2}\exp\left(-\frac{x^2+y^2}{2\sigma^2}\right). \tag{6}$$

$$\frac{\partial \boldsymbol{E}(x,y,t)}{\partial t}=\frac{\partial}{\partial t}\begin{pmatrix}E_x\\E_y\end{pmatrix}\propto\begin{pmatrix}\frac{\partial^2}{\partial y^2}G_\sigma * E_x-\frac{\partial^2}{dxdy}G_\sigma * E_y\\-\frac{\partial^2}{\partial x\partial y}G_\sigma * E_x+\frac{\partial^2}{\partial y^2}G_\sigma * E_y\end{pmatrix}. \tag{7}$$

Therein, G_σ is the two-dimensional Gaussian function with standard deviation σ. The operator $*$ is the convolution operator. The standard deviation was set as $\sigma=1$. Here, x and y are discretized by unit length 1. Update rules for both horizontal vectors E_x and vertical vectors E_y were applied on borders, defined by $B(x,y)\in\{0,1\}$, where 1 denotes the border and 0 otherwise.

We define the initial vector field at $t=0$ as $\boldsymbol{E}_{\text{init}}(x,y)$, whereas the vector field after sufficiently minimizing energy function $J[\boldsymbol{E}]$ is defined as $\boldsymbol{E}_{\text{end}}(x,y)$. In our simulations, vector updating was conducted for 1000 iterations. Initial vectors $\boldsymbol{E}_{\text{init}}(x,y)$ were set as vectors extending towards the inner side of L-junctions, in line with Gestalt observations that the concave side of a contour is likely to be perceived as the inner side of an object. Initial vectors were calculated based on the partial derivatives of border $B(x,y)$ at L-junctions using central difference approximation,

$$\boldsymbol{E}_{\text{init}}(x,y)=\begin{pmatrix}\frac{\partial}{\partial x}B(x,y)\\\frac{\partial}{\partial y}B(x,y)\end{pmatrix}\approx\frac{1}{2}\begin{pmatrix}B(x+1,\mathrm{y})-B(x-1,\mathrm{y})\\B(x,\mathrm{y}+1)-B(x,\mathrm{y}-1)\end{pmatrix}. \tag{8}$$

Zero vectors are assigned at every other point on the border, including T-junctions.

At a spatial position (x_0,y_0), the depth order value $\phi(x_0,y_0,t)$ was calculated using line integration of $\boldsymbol{E}(x,y,t)$. Consider an open path C starting from an arbitrary position and ending at (x_0,y_0), $C(s)=(x(s),y(s))$, where s signifies the arc-length parameter. When the tangent of $C(s)$ is defined as $\hat{\boldsymbol{t}}(s)=dC(s)/ds$, line integral of vectors $\boldsymbol{E}(s)$ is calculated using $\sum_s \boldsymbol{E}(s)\cdot\hat{\boldsymbol{t}}(s)$. For simplicity, we limit integral paths to horizontal and vertical lines passing through (x_0,y_0). Consequently, we obtain

$$\phi(x_0,y_0,t)=\sum_x E_x(x,y_0,t)\,\mathrm{sign}(x_0-x)+\sum_y E_y(x_0,y,t)\,\mathrm{sign}(y_0-y). \tag{9}$$

Therein, $\mathrm{sign}(x)=1$ for $x>0$ and $\mathrm{sign}(x)=-1$ for $x<0$. Summation was operated on the image domain excluding its own position, (x_0,y_0). The first component in **Equation (9)** describes integration on the horizontal direction, whereas the second describes integration on the

vertical path. Depth order and BO signals were rescaled according to perceived depth order.

2.3 Modeling V2 responses

To compare model outputs with actual neuronal outputs qualitatively, we must address how BO signals affect the response of a V2 neuron. For simplicity, our explanation will address only horizontal BO signals. The output of a neuron with selectivity for objects located on the right side (on the left side) of its receptive field is designated as $\mathrm{V2}_{\rightarrow}$ ($\mathrm{V2}_{\leftarrow}$). These outputs can be modeled as the sum of two components: the response to the border in the receptive field (Border in **Equation (10)** and **Equation (11)**), and information encoding ownership from outside the receptive field ($\mathrm{Owner}_{\rightarrow}$ and $\mathrm{Owner}_{\leftarrow}$),

$$\mathrm{V2}_{\rightarrow} = \mathrm{Border} + \mathrm{Owner}_{\rightarrow}, \tag{10}$$

$$\mathrm{V2}_{\leftarrow} = \mathrm{Border} + \mathrm{Owner}_{\leftarrow}. \tag{11}$$

Therein, $\mathrm{Owner}_{\rightarrow}$ and $\mathrm{Owner}_{\leftarrow}$ are either positive values or zero; Border takes a positive value when an object's border is placed in the receptive fields. BO signal E_x can then be defined as the difference in outputs of the two neurons,

$$E_x = \mathrm{V2}_{\rightarrow} - \mathrm{V2}_{\leftarrow} = \mathrm{Owner}_{\rightarrow} - \mathrm{Owner}_{\leftarrow} \tag{12}$$

In **Figure 4**, an object is located on the left side of the receptive field. Consequently, information related to ownership on the left side is positive: $\mathrm{Owner}_{\leftarrow} > 0$. In contrast, no information related to ownership exists for the object on the right side: $\mathrm{Owner}_{\rightarrow} = 0$. This results in $E_x = -\mathrm{Owner}_{\leftarrow} < 0$, or a BO signal signifying that an object is located on the left. Similarly, BO signals in the vertical direction E_y are calculated. The model proposed by Craft et al. (2007) uses a similar approach to reproduce V2 responses by considering border signals and ownership signals.

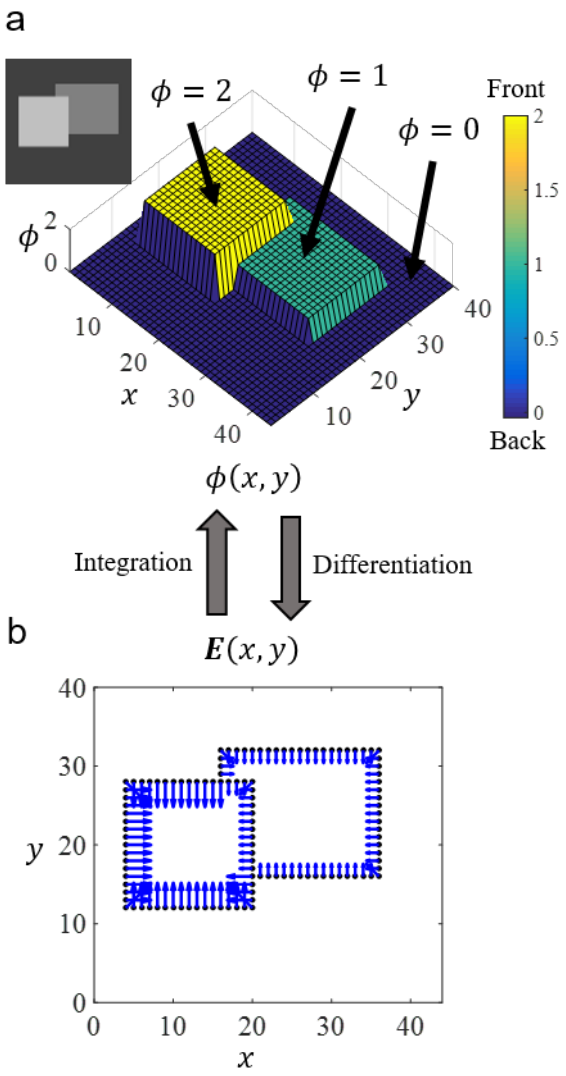

Figure 2. Standard occlusion problem of two overlapping rectangles. (a) Perceived depth order of the image in the inset. The rectangle on the left has a higher depth order $(\boldsymbol{\phi = 2})$ than that of the occluded rectangle on the right $(\boldsymbol{\phi = 1})$. (b) Gradient of perceived depth order produces vectors $\boldsymbol{E(x, y)}$ with arrowheads pointing to the "owner" of edges. These vectors resemble BO signals, suggesting that the depth order and BO signals share an integration–differentiation relation similar to that of electric potential and its corresponding electric field.

1

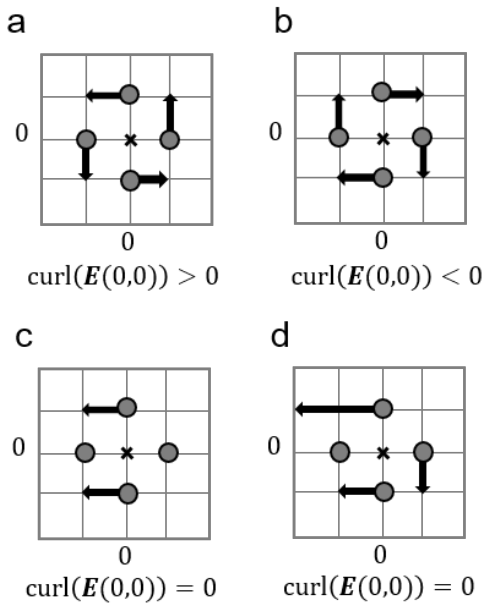

Figure 3. Rotation of two-dimensional vector field $\boldsymbol{E}(\boldsymbol{x},\boldsymbol{y})$ calculated using the "curl" operator at the cross mark positioned at $(\mathbf{0},\mathbf{0})$. Placing a windmill at the cross mark in (a) will cause it to rotate counter-clockwise because of wind vector $\boldsymbol{E}$, which coincides with a positive "curl" value. Similarly, (b) coincides with a negative "curl" value and a clockwise rotation of a windmill placed at the cross mark. Examples presented in (c) and (d) produce a zero "curl" value, and a windmill will not rotate if placed at the cross mark. Two-dimensional rotation "curl" is zero at all spatial points for a conservative field.

1

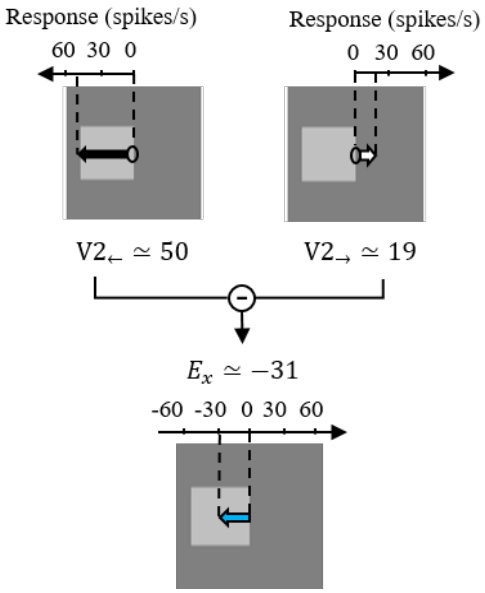

Figure 4. Data were adapted based on results of the experiment conducted by Zhou et al. (2000) to demonstrate how the difference in response of a pair of object-side selective neurons ($\mathbf{V2_{\rightarrow} \simeq 50}$ and $\mathbf{V2_{\leftarrow} \simeq 19\ \ spikes/s}$) produces a BO signal for objects in the horizontal direction, $\boldsymbol{E_x = V2_{\rightarrow} - V2_{\leftarrow} \simeq -31}$.The receptive fields of the two neurons are presented as ovals. Arrow directions represent the object-side preference of the neuron, whereas lengths correspond to their response towards the stimulus. In this case, $\mathbf{V2_{\leftarrow} > V2_{\rightarrow}}$ because a rectangular object colored light gray is located on the left side of the receptive field.

1

2

3. Results

First, numerical simulations were conducted for three stimuli used in neurophysiological experiments by Zhou et al. (2000). These include a simple square (**Figure 5a**), a C-shaped figure (**Figure 5b**) and overlapping rectangles (**Figure 5c**). Initial vectors for all three stimuli were assigned to the concave side of the contour, which does not necessarily point to the perceived object, as shown in the C-shaped figure (**Figure 5b)**. Calculated BO signals, or the vector field after energy minimization $\boldsymbol{E}_{\text{end}}(x, y)$ agreed qualitatively with neurophysiological results for a neuron that responds strongly if a figure (owner) locates on the right of its receptive field. This area corresponds to the red circle in each diagram. Vectors that initially were zeros or which faced outwards at the concave region of the C-shaped figure were reversed to face the inside of the object. In addition, the depth order $\phi(x, y)$ for all three stimuli agrees with perception, especially in the case of overlapping rectangles (**Figure 5c**). A larger value of depth order ϕ was obtained for the occluding rectangle in comparison to the occluded rectangle. The graph in **Figure 6** presents a comparison of the relative response of the model at $\boldsymbol{E}_{\text{end}}(20,21)$ for the three stimuli in **Figure 5**. **Equation (12)** was used to calculate BO signals from actual neural responses in Zhou et al. (2000). Our model captures the dip in response to a C-shaped figure.

To evaluate the model's robustness towards both local concavity and occlusion, an image of a C-shaped figure overlapping a square was examined (**Figure 7**). Immediately after the update rule was applied, the BO vector field and corresponding depth order evolve based on local concavity cues. As time progresses, information related to the global figure shape and order prevails over local cues. After energy minimization, the depth order which agrees with human perception of a C-shaped figure occluding a rectangle was obtained. Despite the absence of a specific T-junction detector in the model, BO signals were assigned successfully towards the occluding figure.

An interesting characteristic of our model is that it can also calculate BO signals for "ambiguous" borders. **Figure 8** presents two examples of figures that might be perceived with an "ambiguous" border that is attributed to a left rectangle, a right one, or both. **Figure 8a** can be perceived as two interlocking L-shaped figures, whereas **Figure 8b** can be perceived as two adjoined rectangles. Zero vectors were assigned at the ambiguous border along the center of the figures in **Figures 8a, 8b**, with similar corresponding depth values for regions on both sides of the ambiguous border.

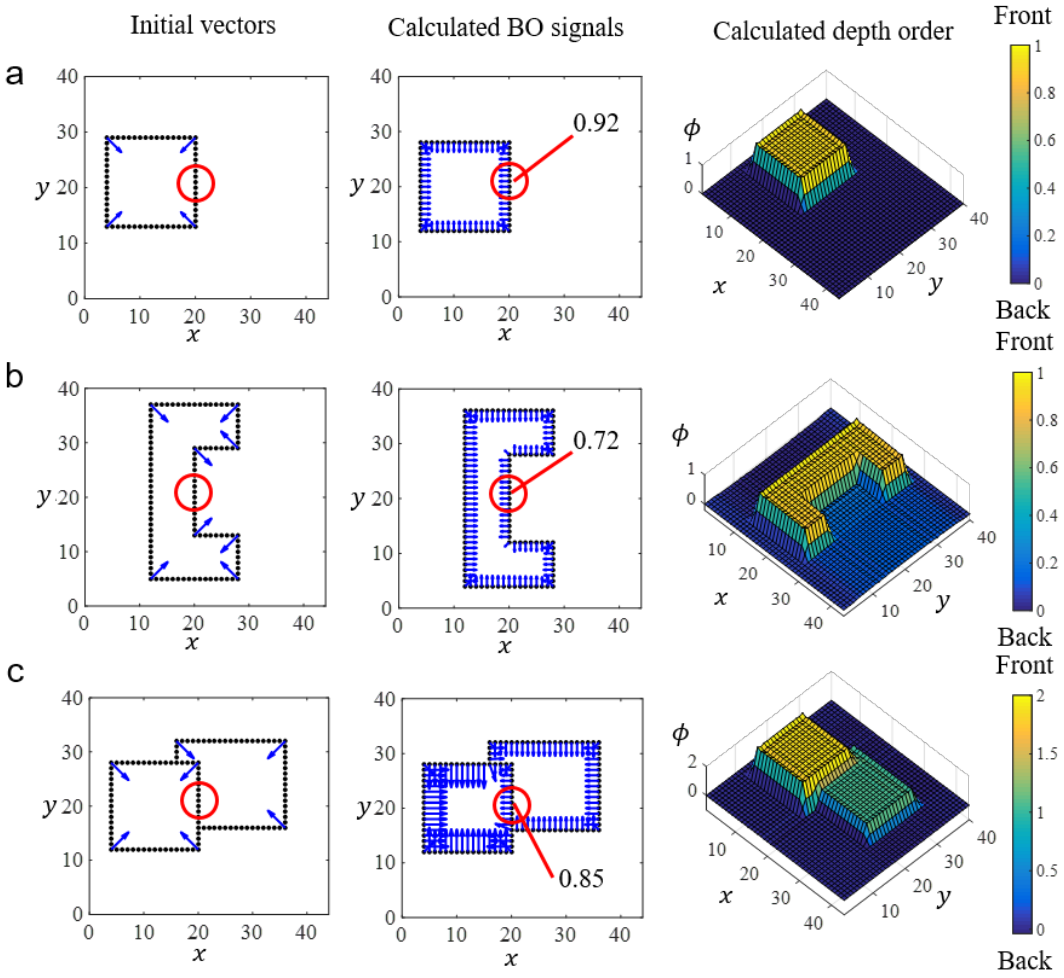

Figure 5. Calculated BO signals and depth order for three stimuli (a, square; b, C-shaped figure; c, overlapping figure) similar to those used for neurophysiological experiments. Initial vector field $\boldsymbol{E_{init}(x,y)}$ consists of vectors facing the concave side of the L-junctions. Vector field after minimization $\boldsymbol{E_{end}(x,y)}$ and its corresponding depth order $\boldsymbol{\phi(x,y)}$ are calculated simultaneously by the model. For each stimulus, the vectors in the red circle centered around $\boldsymbol{(20,21)}$ face leftward, where the occluding object exists. This leftward facing agrees qualitatively with results from neurophysiological experiments for a neuron with selectivity for an object on the left of its receptive field. Vector magnitudes of 0.92, 0.72, and 0.85 are shown for $\boldsymbol{|E_{end}(20,21)|}$ in all three examples.

1

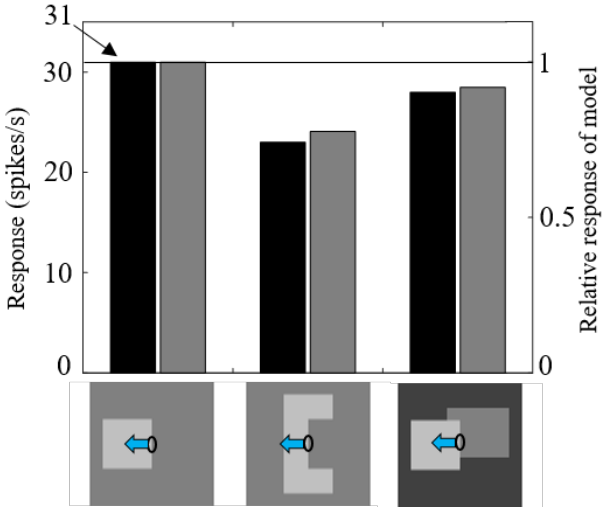

Figure 6. Relative response of our model in comparison to BO signals calculated using experimental data reported by Zhou et al. (2000). Responses are from the same neuron in **Figure 4.** The neuron's receptive field is presented as an oval, whereas an arrow represents the direction of the BO signal of a V2 neuron for stimuli of three types. Black bars correspond to BO signals calculated from actual data. Gray bars correspond to the relative response of the model.

1

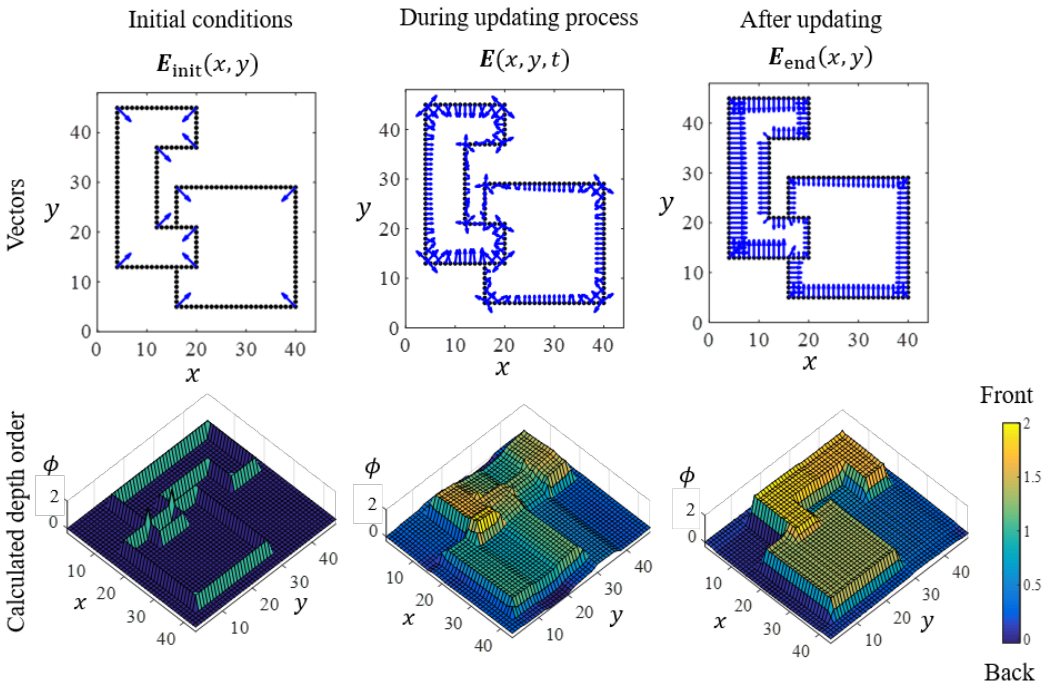

Figure 7. Calculated BO signals (vectors) and depth order over time for an overlapping C-shaped figure. During the updating process, vectors around the concave area face away from the object because of local cues. Correct vector directions were assigned for the overlapping C-shaped figure after energy was minimized sufficiently.

1

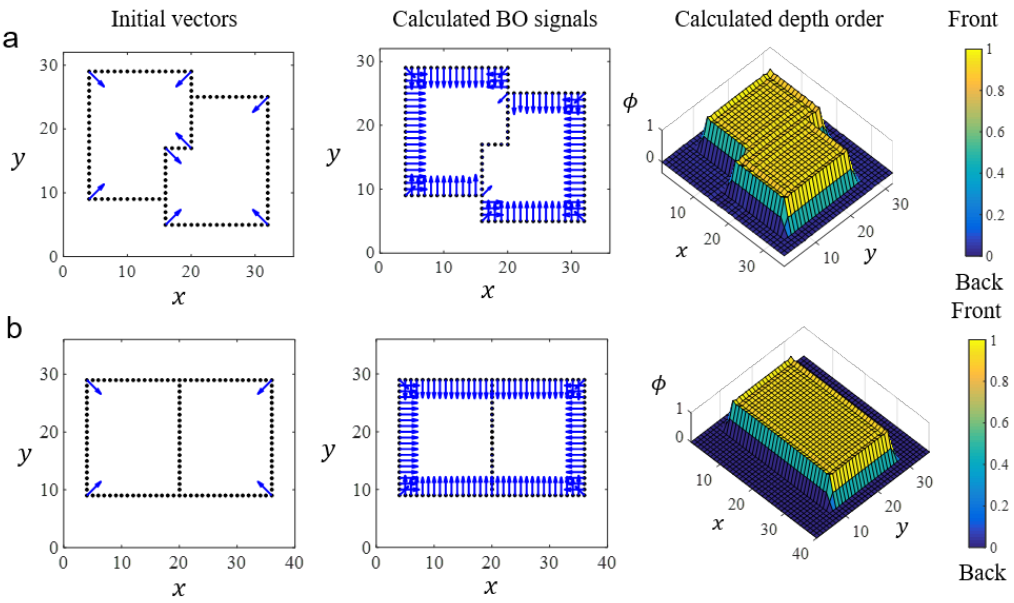

Figure 8. Calculated BO signals and depth order for figures with "ambiguous" borders for (a) interlocking L-shaped blocks and (b) adjoined square blocks. The resulting vector field reflects BO ambiguity, with zero vectors along the center and both sides of equally valued depth order. The T-junctions in both figures do not correspond to a junction of three surfaces of different depth order.

2

4. Discussion

The proposed model simultaneously generated BO signals and depth order that match human perception. As **Figure 7** shows, the resultant BO signals of the model, $\boldsymbol{E}_{\mathrm{end}}(x, y)$, is tolerant of edge arrangement (concavity or convexity) and occlusion, although some initial BO signals $\boldsymbol{E}_{\mathrm{init}}(x, y)$ are not correct, as shown in the concave L-junctions. We cannot determine whether an L-junction is concave or convex, but our model distinguishes it automatically by spatial propagation of BO signals among adjacent BO-coding neurons.

The model can produce depth order for overlapping objects even though it uses no detectors of T-junctions, which have been used in other studies as a specific and useful clue to detect boundaries between occluding and occluded objects (Thielscher & Neumann, 2008). T-junctions might be informative, but a local T-junction is insufficient to determine a unique solution of BO assignment and order estimation, as exemplified in **Figure 9**, which portrays four possibilities of surface order and corresponding BO signals represented as vectors of different lengths. The rotations (curl) at the T-junction are zero for all vector combinations shown. **Figures 9c, 9d** shows that T-junctions do not necessarily correspond to the intersection of three surfaces with different depth order. Our model uses global contextual information to determine appropriate BO signals and surface depth order at T-junctions, as presented on the right side of **Figure 9**.

In normal circumstances, 3D depth information in a visual scene can be estimated from, for example, binocular disparity. When such 3D depth information is available, depth order $\phi(x, y)$ is directly observable. The visual system might learn the relationship between depth order and edge alignments in a 3D environment. This a priori knowledge learned in a 3D environment might be useful to assign initial BO signals of 2D images as used for this study.

4.1 Deduced neural mechanisms

We observed that the neural connections deduced from **Equation (7)** were of a similar structure to that of Li's model, suggesting that our proposed model can serve as a theoretical base for describing horizontal connections within area V2 (Li, 2005). When $E_x(x, y) > 0$, the value of E_x can be regarded as the bias signal (Owner$_{\rightarrow}$ in **Equation (10)**) of a BO-coding neuron that shows selective response if the edge owner is located on the right of its receptive field (x, y). For simplicity, we refer respectively to E_x and E_y as horizontally and vertically selective neurons. Specific examination of $\partial E_x(x, y)/\partial t$ of **Equation (7)** reveals two components: (i) $\partial^2 G/\partial y^2$ represents lateral connections to adjacent horizontally selective neurons E_x and self-inhibition; and (ii) $-\partial^2 G/\partial x \partial y$ represents connections to vertically selective neurons E_y. Those lateral connections between BO-coding neurons are presented in **Figure 10**. Connections with horizontally selective neurons, $\partial^2 G/\partial y^2$, is equivalent with diffusion process of E_x along the y-direction, whereas connections with vertically selective neurons, $-\partial^2 G/\partial x \partial y$, are diagonal. Notably, the

second component $\partial^2/\partial x\partial y$ carries a larger value at the corners of an object. Some neurons in V2 reportedly respond to acute angles (Ito & Goda, 2011; Ito & Komatsu, 2004). Their responses might be summed to produce the necessary signals for BO coding. Only one parameter is necessary to formulate the lateral connections in our models: σ.

The argument against lateral connections as the processing medium is that response latency increases with object size (Sugihara, Qiu, & von der Heydt, 2011). As **Figure 7** shows, BO signals around the L-junctions are assigned faster than the center of contours between two L-junctions, in contradiction with the experimentally obtained results. Multi-scale processing and top-down feedback might help to overcome this shortcoming. Increasing the kernel size of the Gaussian derivative in **Equation (7)** by adjusting standard deviation σ will generate an update rule that encompasses a larger spatial area. Feedback connections carry spatially global information faster than lateral connections do (Girard, Hupé, & Bullier, 2001).

Multi-scale processing might aid in amodal completion and solving of problems involving subjective contours, to which neurons in V2 reportedly respond (Day & Kasperczyk, 1983; Peterhans, Esther & von der Heydt, 1989).

4.2 Comparisons with existing models

Results obtained from our study can serve as a theoretical basis for other models. Unlike previous studies, our model formulates the BO assignment task as a well-defined mathematical problem. As a result, the only parameter in the update rule in **Equation (7)** is standard deviation σ for the Gaussian derivative kernels. In contrast, Li's model requires the adjustment of 23 parameters to code BO assignment through lateral connections with surrounding neurons. Facilitative connections derived from our model and those explicitly defined by Li were found to have similar qualitative properties (**Figure 11**).

The present study is not the first to demonstrate how the integration of BO signals might be used to calculate perceived depth order. Kogo's model computes BO based on the geometric alignment of edges and T-junctions as occlusion cues, then proceeds to integrate these BO signals for surface completion (Kogo, Strecha, Van Gool, & Wagemans, 2010).

4.3 Limitations of the model

The deduced neural network described earlier is a generic one that does not account for the diverse properties of neurons. For example, some neurons reportedly respond selectively to square figures, but not C-shaped figures (Zhou et al., 2000).

The focus of this study was limited to updating BO signals, represented as a two-dimensional vector field $\boldsymbol{E}(x, y)$. Formulating an update rule for depth order $\phi(x, y)$, or the surface reconstruction process, might aid in reproducing subjective contours. For example, the level-set

method used in models, which reproduces subjective contours, might be adapted to achieve reproduction of subjective contours (Sarti, Malladi, & Sethian, 2002; Zhu & Chan, 2007). Some neurons in V2 reportedly respond to these subjective contours (Peterhans, Esther & von der Heydt, 1989). Moreover, neurons in V4 reportedly respond to the illusory surfaces which are perceived along with subjective contours (Cox et al., 2013). Interactive processing between contour coding neurons and surface coding neurons through recurrent processing might produce the desired responses.

As described herein, we introduced an example of perceptual alternation as shown in **Figure 1**. This stochastic change in perception cannot be reproduced by our model because the final result converges to a single value. Our model's output depends on how the initial vectors are set. For example, if we reverse initial vectors of a C-shaped figure, then our model will reproduce the perception of a C-shaped hole, as shown in **Figure 12**. The possibility exists that our model can be modified to account for perceptual alternation.

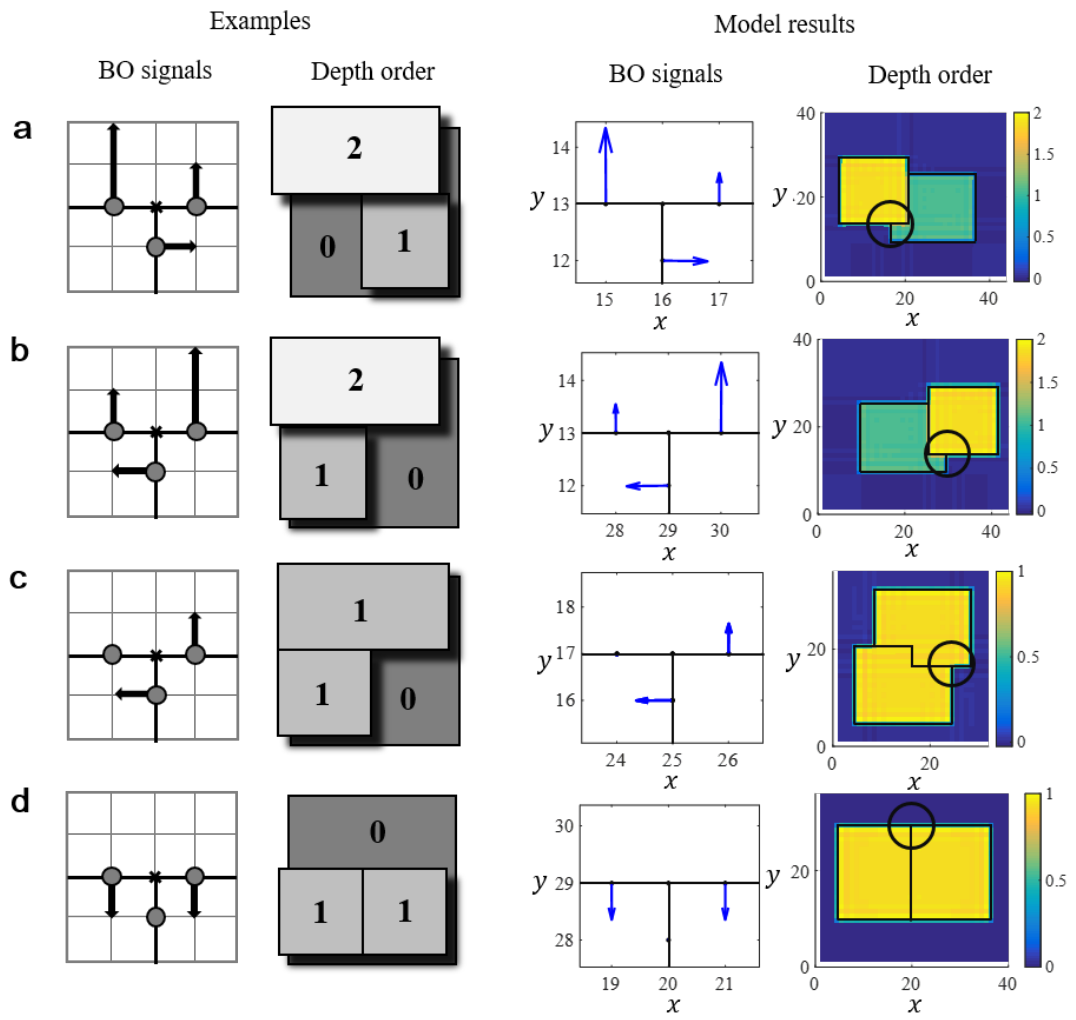

Figure 9. Four possible combinations of BO vectors around T-junctions and the corresponding depth order of surfaces for each combination. Short vectors are of unit length one; long vectors are of unit length two. Vector rotation at the T-junction is zero for any combination. Integers represent the corresponding depth order value, where a larger value coincides with a surface perceived as closer to the viewer. A vector value of 0 is visible at ambiguous borders in (c) and (d). These examples are accompanied by model results for vectors around the T-junctions in the circled region of calculated depth order. (a) and (b) coincide with the T-junction in **Figure 5c**, (c) coincides with the T-junction in **Figure 8a**, and (d) coincides with the T-junction in **Figure 8b**. Model results were rotated accordingly.

1

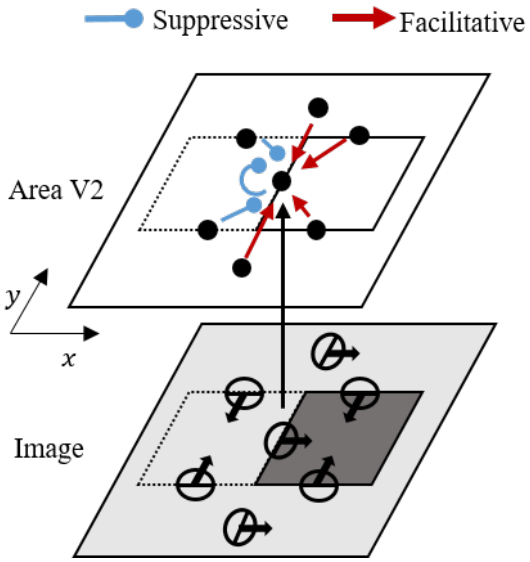

Figure 10. Illustration of the proposed neural network for a neuron with selectivity for an object on the right of a vertical border. Receptive fields of neurons are represented by ovals on the image; arrows represent their side-of-object selectivity. The neuron in question has a receptive field covering the left border of the square in the image. Neurons in area V2, shown by black dots with corresponding receptive fields on the image, are connected by facilitative and suppressive neural connections depending on the side-of-object selectivity and spatial positions of surrounding neurons. These connections were derived from the update rule in **Equation (7)**.

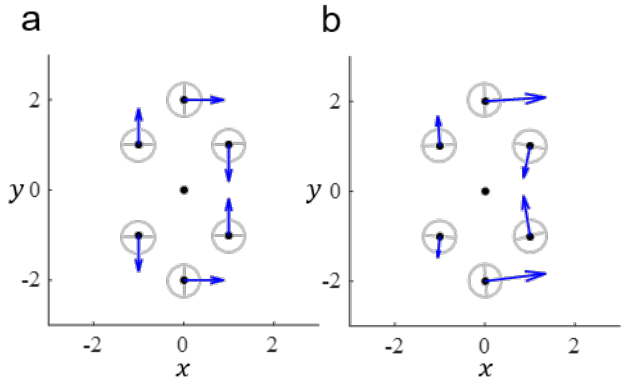

Figure 11. Facilitative connections from six surrounding neurons for a neuron at $(\mathbf{0}, \mathbf{0})$ with preference towards an object on the right side of its receptive field. Connection strengths were derived from weights of (a) our model $(\boldsymbol{\sigma} = \mathbf{1})$ and (b) Li's model (Li, 2005). Vector arrowhead directions indicate the object-side preference of the neuron; their lengths correspond to the connection strength.

2

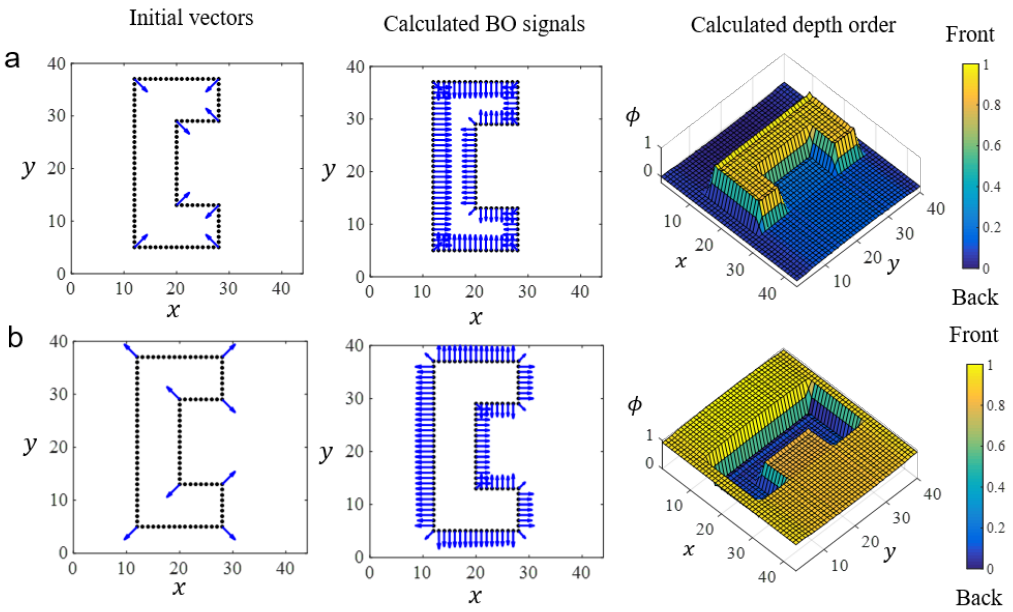

Figure 12. Reversing the direction of initial vectors for a C-shaped figure results in a reversal of calculated BO signals and corresponding depth order. Our model can reproduce both the perception of (a) a C-shaped object and (b) a C-shaped hole in a deterministic manner. This example demonstrates how our model might serve as a basis for reproducing perceptual alternation.

1

5. Conclusion

We demonstrated the importance of formulating the problem in a well-defined manner, by application of a theorem in electromagnetism to solve the problem of BO coding in the visual system. The model can reproduce human perception of depth order qualitatively for various figures. Nevertheless, the model is incapable of producing results of perceptual switching, as shown in **Figure 1**, because initial vectors are given in a deterministic manner. Future work shall include further development of the model to include interactive processing between $\boldsymbol{E}(x, y)$ and $\phi(x, y)$, as well as quantitative reproduction of the firing rates of BO-coding neurons.

References

Cox, M. A., Schmid, M. C., Peters, A. J., Saunders, R. C., Leopold, D. A., & Maier, A. (2013). Receptive field focus of visual area V4 neurons determines responses to illusory surfaces. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(42), 17095–17100.

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

Day, R. H., & Kasperczyk, R. T. (1983). Amodal completion as a basis for illusory contours. *Perception & Psychophysics*, *33*(4), 355–364.

Girard, P., Hupé, J. M., & Bullier, J. (2001). Feedforward and Feedback Connections Between Areas V1 and V2 of the Monkey Have Similar Rapid Conduction Velocities. *Journal of Neurophysiology*, *85*(3), 1328–1331.

Ito, M., & Goda, N. (2011). Mechanisms underlying the representation of angles embedded within contour stimuli in area V2 of macaque monkeys. *European Journal of Neuroscience*, *33*(1), 130–142.

Ito, M., & Komatsu, H. (2004). Representation of angles embedded within contour stimuli in area V2 of macaque monkeys. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *24*(13), 3313–3324.

Kogo, N., Strecha, C., Van Gool, L., & Wagemans, J. (2010). Surface construction by a 2-D differentiation-integration process: a neurocomputational model for perceived border ownership, depth, and lightness in Kanizsa figures. *Psychological Review*, *117*(2), 406–439.

Li, Z. (2005). Border ownership from intracortical interactions in visual area V2. *Neuron*, *47*(1), 143–153.

Lindeberg, T. (1994). Scale-space theory: a basic tool for analyzing structures at different scales. *Journal of Applied Statistics*, *21*(1), 225–270.

Nakayama, K., Shimojo, S., & Silverman, G. H. (1989). Stereoscopic depth: Its relation to image segmentation, grouping, and the recognition of occluded objects. *Perception*, *18*, 55–68.

Peterhans, E., & von der Heydt, R. (1989). Mechanisms of Contour Perception Contours Bridging Gaps in Monkey Visual Cortex . *Journal of Neuroscience*, *9*(5), 1749–1763.

Poggio, T., Torre, V., & Koch, C. (1985). Computational vision and regularization theory. *Nature*, *317*, 314–319.

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.

Sakai, K., Nishimura, H., Shimizu, R., & Kondo, K. (2012). Consistent and robust determination of

border ownership based on asymmetric surrounding contrast. *Neural Networks*, *33*, 257–274.

Sarti, A., Malladi, R., & Sethian, J. A. (2002). Subjective surfaces: A geometric model for boundary completion. *International Journal of Computer Vision*, *46*(3), 201–221.

Sugihara, T., Qiu, F. T., & von der Heydt, R. (2011). The speed of context integration in the visual cortex. *Journal of Neurophysiology*, *106*(1), 374–385.

Thielscher, A., & Neumann, H. (2008). Globally consistent depth sorting of overlapping 2D surfaces in a model using local recurrent interactions. *Biological Cybernetics*, *98*(4), 305–337.

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

Zhu, W., & Chan, T. (2007). A variational model for capturing illusory contours using curvature. *Journal of Mathematical Imaging and Vision*, *27*(1), 29–40.