A feedback model of figure-ground assignment

Dražen Domijan
Mia Šetić

University of Rijeka, Rijeka, Croatia

A computational model is proposed in order to explain how bottom-up and top-down signals are combined into a unified perception of figure and background. The model is based on the interaction between the ventral and the dorsal stream. The dorsal stream computes saliency based on boundary signals provided by the simple and the complex cortical cells. Output from the dorsal stream is projected to the surface network which serves as a blackboard on which the surface representation is formed. The surface network is a recurrent network which segregates different surfaces by assigning different firing rates to them. The figure is labeled by the maximal firing rate. Computer simulations showed that the model correctly assigns figural status to the surface with a smaller size, a greater contrast, convexity, surroundedness, horizontal–vertical orientation and a higher spatial frequency content. The simple gradient of activity in the dorsal stream enables the simulation of the new principles of the lower region and the top–bottom polarity. The model also explains how the exogenous attention and the endogenous attention may reverse the figural assignment. Due to the local excitation in the surface network, neural activity at the cued region will spread over the whole surface representation. Therefore, the model implements the object-based attentional selection.

Keywords: attention, boundary, figure-ground, neural model, saliency


Introduction

Figure-ground organization is an important early step in visual processing. It separates structured input to which processing efforts should be devoted (figure) from less structured background. Gestalt psychologists identified several variables which influence figure-ground assignment including size, contrast, surroundedness, convexity, symmetry, parallelism and horizontal–vertical axes (Palmer, 1999). For instance, smaller surfaces tend to be perceived as figures as well as surfaces with larger contrast. When there is a single surface on the display it will be perceived as a figure since it is completely surrounded with background. Symmetrical surfaces and surfaces with parallel boundaries also tend to be seen as figures. Horizontal–vertical axes refer to a tendency to assign figural status to the surface which is aligned along vertical or horizontal axes in comparison with the oblique alignment.

Recently, two new factors affecting figure-ground organization were revealed. Those are the lower region and the top–bottom polarity. Lower region refers to the tendency to assign a figural status to the surface in the lower part of the visual field (Vecera, Vogel, & Woodman, 2002). On the other hand, the top–bottom polarity refers to the tendency to assign a figural status to the surface with a wide base and a narrow top (Hullemann & Humphreys, 2004). Hullemann and Humphreys argued that the top–bottom polarity could not be reduced to the principle of the lower region because the tendency to assign figural status to the surface with the wide bottom remains even when the figure is presented in the top portion of the visual display. Also, they tried to make sure that no other principles contributed to the figural assignment in their displays, such as contrast or size. Fowlkes, Martin, and Malik (2007) provide evidence that some of the principles of figure-ground organization may have ecological validity due to the fact that they influence figure-ground assignment in complex real-world images. They used a large collection of natural images which were segmented into figure and ground regions by human observers. Results showed that observers had a tendency to assign figural status according to the principle of size, convexity and lower region. Other principles were not tested but it is reasonable to assume that they would also generalize to real scenes.

Top-down processes like familiarity and attention may also influence figure-ground assignment. Peterson (1994; Peterson & Gibson, 1994a, 1994b) showed that figural status in an ambiguous figure is assigned to the surface that is recognized as a part of the familiar object. Baylis and Driver (1995; Driver & Baylis, 1996) showed that endogenous (or voluntary) attention may change the figural status of the surface. On the other hand, endogenous (or automatic) attention was not able to reverse the figure and the background. Vecera, Flevaris, and Filapek (2004) clarified this issue by suggesting that cueing in the experimental paradigm of Baylis and Driver (1995) was not effective because they presented the cue near the surface which should be perceived as a figure. When the cue is presented on the surface, participants readily assign a figural status to the cued surface. All these findings are in

doi: 10.1167/8.7.10

Received September 30, 2007; published May 30, 2008

ISSN 1534-7362 © ARVO
contrast with the standard model of the visual processing explicitly formulated by Pylyshyn (1999) in the concept of cognitive impenetrability. According to this view, the task of visual perception is to deliver an accurate representation of the external world on which the higher level cognitive processes could operate. On the other hand, top-down processes should not influence visual perception because they may interfere with the central task of visual perception. However, it seems that top-down processes actually enhance the ability of the visual system to deliver accurate information rather than diminishing it.

Neurophysiological investigations of figure-ground assignment reveal enhanced firing rate in the primary visual cortex of monkeys (Lamme, 1995; Lee, 2003; Zipser, Lamme, & Schiller, 1996). Activity enhancement was observed in the region which corresponds to the perceived figure in texture segregation. Analysis of timing of neural responses revealed that shortly after the stimulus presentation only boundaries of the figure are detected. Later, neural activity in the interior of the surface is also enhanced (Lamme, Rodriguez-Rodriguez, & Spekreijse, 1999). Therefore, the figure is represented in the primary visual cortex by different firing rate compared to the background. Although, it should be mentioned that investigation in different laboratory failed to find such interior enhancement (Rossi, Desimone, & Ungerleider, 2001). Based on these findings, Lamme and Roelfsema (2000) proposed a model of visual processing which includes two distinct modes of operation. The fast feedforward mode detects elementary features like oriented boundaries, colors, and motions. The slower feedback mode enables binding of non-local information into the representation of global shapes or objects. Recently, Zhou, Friedman, and von der Heydt (2000; Qiu, & von der Heydt, 2005) discovered a special group of neurons which detect border ownership in the V2 cortex. The figure is distinguished from the background by different responses to the same boundary. If the figure is on one side of the boundary, a certain neuron will fire, but if the figure is on the other side of the same boundary the same neuron will be silenced and another neuron will show an enhanced firing rate. These boundary signals are available as early as 30 ms after the input presentation, placing an important constraint for the models of border ownership (Craft, Schütze, Niebur, & von der Heydt, 2007; Jehee, Lamme, & Roelfsema, 2007).

There are many models proposed to explain the properties of figure-ground organization. Several models focused on the boundary assignment and modeling of results by Zhou et al. study. Zhaoping (2005) used the recurrent excitatory–inhibitory network to show that a correct boundary assignment can be achieved if proper connection patterns are set up in the network. This approach was criticized by Craft et al. (2007) and by Jehee et al. (2007) who pointed out that recurrent or horizontal connections are not sufficient to account for the fast response of the boundary assignment neurons in the V2. Craft et al. (2007) and Jehee et al. (2007) used feedback projections from higher level visual centers in order to disambiguate local boundary responses and provide a coherent figural assignment. Sakai and Nishimura (2006) considered in which way border ownership might arise from various configurations of surround facilitation and suppression in the visual cortex. Vecera and O’Reilly (1998) showed how boundary assignment might be influenced by the object recognition. Their model successfully simulates the findings that familiar shapes are more likely to become a figure. However, their model was based on a biophysically unrealistic computational mechanism and it often produces spurious states that do not correspond with none of the two possible perceptual interpretations. Roelfsema, Lamme, Spekreijse, and Bosch (2002) simulated the interior enhancement observed in monkey primary visual cortex. Their model correctly reproduces different timings of boundary signals and interior signals. However, their model was restricted to the texture-segregation task and they did not consider more complex image configurations. Grossberg (1994, 1997) proposed a complex multi-scale neural architecture for visual perception which solves the figure-ground assignment as part of the larger issue of interaction of depth and surface perception. In the model, different network layers correspond to different depth planes. The surface that appears on a near depth plane is perceived as a figure. Despite its success in explaining a wide range of behavioral and neurophysiological data, it is not clear how this model explains the lower region and the top–bottom polarity as a source of figural assignment. Also, it does not incorporate the difference between endogenous and exogenous orientation of attention and their influence on figure-ground organization. The same criticism is true for all other previously mentioned models.

The aim of the present paper is to introduce the neural model of visual processing that can account for classical and recently discovered principles of figure-ground assignment. The model implements the notion of object-based biased competition and explains how top-down processes may influence figure-ground assignment (Vecera, 2000). The model incorporates several neurophysiological findings such as the interaction between the ventral and the dorsal visual processing stream (Bullier, 2001), gradient synaptic weights and neural activity enhancement corresponding to the figural status (Lamme, 1995; Lee, 2003). These components help to explain how bottom-up and top-down signals can be combined into a unified percep of figure and background.

Methods

We propose architecture for visual processing which can account for classical and new principles of figural assignment. The model is based on the feedback interaction
between the ventral and the dorsal processing stream (Bullier, 2001). In particular, it is suggested that the dorsal processing stream computes saliency of the presented surfaces and feeds this information back into the construction of the surface representation. The model is depicted in Figure 1. It consists of five network layers.

The first layer is an input whose activity is scaled in the range between 0, indicating a black surface, and 2, indicating a white surface. The second layer simulates the properties of the lateral geniculate nucleus (LGN) where cells with on-center off-surround and off-center on-surround receptive fields detect anisotropic luminance.

Figure 1. A neural architecture for figure-ground assignment. The input is first processed by on-center and off-center cells in the LGN. At the level of the primary visual cortex, simple and complex cortical cells detect the surface's boundaries. They are depicted with horizontal and vertical ellipses representing their different orientation selectivity. Here, the separation between the ventral and the dorsal stream is implemented by two different spatial resolutions. Simple cells in the ventral stream detect sharp boundaries because the ratio between the lengths of the axes of their receptive field is larger. On the other hand, simple cells in the dorsal stream have a lower axes ratio so they provide a blurred version of the boundaries. The parietal cortex contains the blur cells with a large receptive field depicted as circles and corner cells sensitive to co-occurrence of horizontal and vertical boundaries. The blur cells detect density and strength of boundary signals while the corner cells detect important geometrical relationships between boundary segments such as T-junctions and corners. In this way the parietal cortex computes a simplified version of saliency. Boundary signals from the ventral stream and saliency signals from the parietal cortex are combined in the surface network. In the surface network different surfaces are labeled with different firing rates. Activity labeling is illustrated by squares with different shades of gray. The black and the gray surfaces constitute a background and the white surface is a figure.
discontinuities in the input image and thus suppress uniform luminance regions.

The third layer computes oriented luminance discontinuities (edges or boundaries) by simulating the properties of simple and complex cells in the primary visual cortex. At this network stage we introduce the segregation into two processing streams. Two separate networks compute oriented boundary responses in different ways corresponding to the properties of the ventral and the dorsal stream. One network simulates the ventral stream by producing a sharp response to the boundary (thin ellipses in Figure 1). Another network simulates dorsal stream by producing a blurred response to the same boundary (thick ellipses in Figure 1). The sharpness of the response is controlled by the shape of the receptive field of the simple cells. The shape is described with two parameters which control the elongation of the receptive field in the direction of the preferred orientation and in the direction perpendicular to the preferred orientation. In Appendix A we provide specific values for these parameters. The output of the ventral boundary network is used to construct the surface representation at the fifth layer by providing a stopping signal for spreading excitatory activation. The dorsal boundary network feeds its output to the fourth layer. The fourth layer simulates the computation of saliency in the parietal cortex (Itti & Koch, 2000, 2001; Treue, 2003). The response of the parietal cortex to the input from the dorsal boundary network contains four components bound together. Those are: blur cells, corner cells, activity gradient and random fluctuation. Blur cells compute strength or density of the boundary signals in their receptive fields. This is achieved by convolution of the input from the dorsal boundary network with Gaussian kernels. Therefore, blur cells are sensitive to the amount of boundary responses irrespective of their orientation. On the other hand, corner cells detect co-occurrence of horizontal and vertical boundary responses (Anzai, Peng, & Van Essen, 2007). They are particularly sensitive to modal or amodal corners and T-junctions. By amodal corner we mean the corner that is not visible due to the occlusion, but whose existence is implicated by the visible boundaries. The receptive field of the corner cell is constructed by multiplicative interaction between outputs of the convolution of horizontal and vertical kernels with the input from the dorsal boundary network. Corner cells might be considered as an example of a more general class of cells tuned to complex shapes such as cross, square, circle or star found in V1, V2, and V4. Some of these cells are tuned to particular angles, to the orientation of angles or to the combination of angles and their orientations (Hedégé & Van Essen, 2000; Pasupathy & Connor, 1999; Shevelev, 1998). Equivalent selectivity to complex shapes was also found in the posterior parietal cortex (Sereno & Maunsell, 1998). Here, we considered only selectivity for cross pattern or for co-occurrence of horizontal and vertical boundary segments (Anzai et al., 2007).

The activity of the parietal network is weighted differently in the upper and the lower part of the spatial map before it can influence the next processing step. Different weighting is achieved by using a small gradient in the intensity of neural activity in the parietal cortex. We assumed that the gradient arises from the tonic input whose connections to the neurons in the parietal cortex have different strength for different network locations. The strength of synaptic weights increases from the top locations toward the bottom locations. This is the crucial ingredient in explaining the lower region and the top–bottom polarity. It will allow the surfaces lying in the lower spatial positions to achieve a competitive advantage over the surfaces in the upper location in the visual field. Such gradient is consistent with psychophysical reports showing better performance in the lower visual field for a number of visual tasks (Levine & McAnany, 2005; McAnany & Levine, 2007; Previc, 1990; Rubin, Nakayama, & Shapley, 1996).

Furthermore, we introduce a small amount of random fluctuation in the activity of the neurons in the parietal cortex. Reason for this is the fact that in the case of two ambiguous surfaces (i.e., surfaces with equal chance to be selected as figure) one surface is selected by chance. The same randomness also helps explain why sometimes participants select a background instead of a figure (Vecera et al. 2002). In other words, figure-ground assignment is not entirely a deterministic process. The described principles of figure-ground assignment only point to the most likely choice. On some occasions participants choose a surface that is not favored by the bottom-up or the top-down cue. It should be noted that assigning figural status to the background is not the same thing as the spurious state in the model of Vecera and O’Reilly (1998). Their model produces network solutions which do not correspond with any particular surfaces present in the input. This is not possible in the present model, which will always represent complete surfaces. Although described components of the activity in the parietal cortex represents a highly rudimentary saliency computation it will be sufficient to model many aspects of figure-ground organization.

Finally, the outputs from the ventral boundary layer and the parietal cortex are combined in a recurrent network which provides the surface representation (Domijan, 2004). The recurrent surface network implements object-based lateral inhibition and enables only one connected group of cells to attain the maximal rate of activity. This group of cells corresponds to the surface of the perceived figure. The surface network is a model of visual segmentation based on a rate code. It includes a recurrent network of mutually connected excitatory and inhibitory neurons. However, the standard model of a neuron is extended with dendritic inhibition. The mechanism of dendritic inhibition assumes that dendrites operate as independent computational devices with their own input–output relationships (Häusser & Mel, 2003; London &
Dendritic inhibition allows separation of the activity amplitude for different groups of neurons. Therefore, different objects are represented by different neuron activity rate. On the other hand, neurons that represent the same object have the same activity level. Due to the fact that the surface network is a recurrent network, it is able to store current activity pattern into working memory in accordance with neurophysiological findings about persistent neural activity in the primary visual cortex after the stimulus has been removed from the visual field (Supér, Spekreijse, & Lamme, 2001).

The surface network operates by spreading activation among excitatory neurons until it hits an inhibitory obstacle provided by a separate network which is sensitive to the object’s boundary. This is the ventral boundary network which simulates the properties of simple and complex neurons observed in the monkey’s primary visual cortex. The interaction between the ventral boundary network and the output from the parietal cortex achieves a proper representation of the object in the surface network because all neurons which participate in the representation of the same object obtain the same level of activity. Therefore, all spatial locations occupied by the object will be bound together in the visual representation by the same level of firing rate.

The surface network constitutes an intermediate visual representation where local image features are bound into global shapes that correspond to the objects in the visual field (Albright & Stoner, 2002; Nakayama, He, & Shimojo, 1995; Neumann & Mingolla, 2003). The surface network is in many ways similar to the filling-in stage in a model proposed by Grossberg and Todorović (1988). The filling-in network integrates boundary signals with surface properties such as color and brightness. An important difference between the models is that activity in the filling-in stage corresponds to the perceived brightness of the surface. On the other hand, in the surface network, the firing rate is used to signal abstract property of relative importance (or salience) of the given surface. The most important surface is the figure and it will be represented with the maximal firing rate. All other surfaces are regarded as less important and they will be represented by a certain activity level lower than the maximum. The surface network is intended to model firing rate enhancement observed in neurophysiological studies corresponding to homogenous areas in visual input (Kinoshita & Komatsu, 2001; Lamme, 1995; Lee, 2003; MacEvoy, Kim, & Paradiso, 1998; Rossi, Rittenhouse, & Paradiso, 1996; Zipser et al., 1996). However, the model has shortcomings due to the fact that it requires massive all-to-all lateral inhibition in order to achieve sharp separation of activity amplitudes corresponding to different surfaces.

Results

In order to test the capabilities of the proposed neural architecture we performed a series of computer simulations. Network equations and the parameter set are given in the Appendix A. The results of computer simulations are presented in the following format. Columns in the figures correspond with different network layers. The leftmost column represents the input image. The next is a `boundary’ column which shows the output of the computation in the ventral boundary layer in the primary visual cortex. Computation in the LGN is omitted in the results because it does not significantly contribute to the explanation of figure-ground organization. The column labeled ‘parietal’ shows activity in the parietal cortex responsible for saliency computation. The rightmost column is the surface network. In all images, brighter pixels denote higher neural activity and darker pixels denote lower level of neural activity. In the surface network the white area corresponds to the figure because neurons representing this area attained the strongest activity level. It can be said that the figural area won in the competition with other surfaces. However, the background area is not completely erased from the surface representation. It just attained a lower level of activity. Therefore, it is possible to direct attention to the background surface as it will be shown below. The figural assignment should not be conceptualized as equal to the attentional selection. Attention may operate between the figure and the ground and both must be represented in the surface network.

Classical and recent principles of figure-ground organization

Computer simulations presented in Figure 2 showed that the model correctly assigns figural status to the surrounded region, the convex region, the surface with smaller size and larger contrast. The key component to achieve this is a difference in activity across the parietal cortex. Stronger activity in the parietal cortex is observed in the area with greater density of edges responses from the boundary network. Greater edge density makes this area more salient. When the activity from the parietal cortex is fed to the surface network it brings competitive advantage to area with greater density of boundary responses. It is interesting to note that surroundedness and convexity are not mutually exclusive principles. Palmer (1999) noted that surrounded surfaces tend to be convex and small. Jehee et al. (2007) explicitly treated convexity and surroundedness as equivalents. However, we will treat them separately because there are important differences in the input images. In the case of surroundedness there are.
just two surfaces (Figure 2A). On the other hand, in the case of convexity we have three surfaces and two of them stand out as figures (Figure 2B). However, both black and white surface are figures with respect to the gray background. Moreover, the principle of convexity forces us to ascribe figural status only to the black surface and not to the white surface. The white surface becomes background as well as the gray surface. Now, we can analyze network responses to these inputs. The crucial activity differences which will decide the figural assignment are observed in the parietal network. In Figure 2A we can see a strong response at the center of the black surface. This is a consequence of the fact that blur cells in the parietal cortex have an isotropic receptive field and they combine activity from boundary cells with horizontal and vertical orientation selectivity. Therefore, the strongest activity will be in the central area where both horizontal and vertical boundaries are detected. The central area of the image will provide competitive advantage for the black surface in the surface network.

In Figure 2B, the network response in the parietal network at the central area of the image is also pronounced. But this only explains how the gray surface became the background. The difference between black and white surfaces is explained with the fact that the central region where the boundary signals are the strongest belongs to the convex surface. In the central region there are three black–white edges which will produce strong activity in the parietal cortex. On the other hand, parietal activity in the bottom part of the concave surface is weaker because it is surrounded with two black–white edges and one gray–white edge. Gray–white edge has lower contrast and it reduces the saliency of the concave surface. Therefore, the presented model does not compute convexity directly with some sophisticated algorithm (Fowlkes et al., 2007). Rather, convexity is a consequence of the more general sensitivity to the figures with blob-like shapes. It should be noted that the ability to select a convex surface as a figure depends on the control of other factors. For instance, both surfaces have the same contrast. If the contrast of the white surface is increased, it will become a figure. In this way, the model makes a testable prediction that contrast will take over convexity if these two factors compete. It should be mentioned that this analysis does not preclude the possibility that there exists a more complex mechanism for detecting convexity as it will be discussed with regard to symmetry and parallelism. However, it is interesting to note that sensitivity to convex surfaces may arise indirectly from the saliency computation over boundary signals.

Figure 2. Computer simulations illustrating the model ability to explain the classical principles of figural assignment. Columns represent different network stages. Neural activity in a two-dimensional network is depicted as shades of gray, where brighter pixels indicate higher activity level at corresponding spatial location. At the surface network, the figure is always designated as a white surface. For the display we selected a $30 \times 30$ central square of the network dimensions in order to reduce redundancy in the gray background surface. (A) Surroundedness. (B) Convexity. (C) Size. (D) Contrast.
Figure 2C shows that the smaller surface will become a figure. Again, the crucial factor which decides the winner in the surface network is the activity in the parietal cortex. It is interesting to note that the gray background surface appeared lighter than similar surfaces in other simulations in this set. The reason for this is lateral inhibition between cells in the surface network. Lateral inhibition is assumed to operate over the whole network. Therefore, the surface which loses in the competition for figural status will receive the amount of inhibition proportional to the number of winning cells or cells coding the figural region. In this simulation, the number of winning cells is smaller compared to other input configurations. The cells coding the larger surface will settle on the activity level which is closer to the maximal level when compared to other simulations. Figure 2D shows that the surface with greater contrast is designated as a figure. We alter the contrast by increasing the luminance of the gray background region from 1 to 1.5 making it closer to the white surface whose luminance was set to 2. The luminance of the black surface was set to 0. Greater contrast implies stronger response of the boundary cells. Stronger activity in the boundary network will be reflected in the stronger response of the boundary cells. Therefore, the surface with greater contrast became more salient and it won the competition in the surface network. The proposed network treats the contrast and the size as equivalent perceptual attributes. There is no special mechanism for computing size so the parietal cortex does not make a distinction between size and contrast. Reducing the size of the surface has the same effect as increasing the contrast. This is consistent with the findings of Huang and Pashler (2005) who used the visual search paradigm in order to show how size and contrast are combined into the saliency measure. However, in empirical work on the perceptual organization there is no explicit treatment of the relative strength of certain factors of perceptual organization. Our model predicts that, with respect to the size and the contrast, their likelihood of becoming a figure should be proportional to the likelihood that they appear as a target in the visual search. We note that the faint gradient of activity is seen in the parietal network in Figure 1. It is shown as a slight increase of brightness starting from top to bottom horizontal positions. The gradient will allow the network to favor surfaces with a wider area and surfaces in the lower parts of the space. However, the gradient of activity does not contribute much to the explanation of classical principles of figural assignment, but it is important to note that it does not conflict with them either. The same is true for enhances response at the corners. They arise from the activity of the corner cells which will be important component in explaining occlusion and other more complex input configurations.

Figure 3 shows how the model simulates influence of the horizontal–vertical alignment and the spatial frequency content on figure-ground organization. The principle of horizontal–vertical orientation states that the surface aligned in the vertical and horizontal orientations will be perceived as a figure when compared with oblique alignment. It may be considered as a special case of the oblique effect, that is, reduced acuity for oblique lines and gratings relative to horizontal and vertical orientations (Appelle, 1972). In Figure 3A, input consists of two non-adjacent squares where the right square is rotated for 90° around its center of mass. The model favored the left square due to its enhanced sensitivity to corners in the parietal cortex. It should be noted that ventral boundary computation actually favors the oblique square since the complex cells responses are stronger on the right side. However, computation in the parietal cortex overrides this initial advantage due to the stronger response of corner cells to the left square compared to the right square. This works only if both squares are positioned on the same horizontal line. If the right square is placed slightly bellow its current position it will become a figure. Therefore, the model prediction is that horizontal–vertical orientation is not a strong cue for figural assignment and it is easily overridden by the lower region. It might be argued that such comparison is not fair due to the fact that corner cells are only sensitive to horizontal and vertical boundary segments. However, even if we include corner cells with other orientations it is always possible to differentially weighted contribution of horizontal–vertical and oblique corner cells to the output of the parietal cortex. Moreover, horizontal–vertical corner cells should receive stronger emphasis in the parietal cortex in accordance with the oblique effect.

Klymenko and Weisstein (1986) showed that the surface with higher spatial frequency content tends to be perceived as a figure. The surface with higher spatial frequency will produce a response with greater density in blur cells in the parietal cortex. In that way, the spatial frequency effect could be considered as a special case of the size principle, that is, a tendency to ascribe figural status to smaller surfaces. This is shown in Figure 3B. Here, two surfaces are presented as input. Surface on the left side has higher spatial frequency content compared with the right surface. We did not use a full range of amplitudes in the creation of the texture because they will produce weak boundary response which will impair the surface representation in the surface network. Instead, background was set to 0 and the input amplitude on the surfaces varied in the interval [1 2]. In this way, strong boundary response is observed along the whole surface. For the same reason, we separate two surfaces with a small gap.

Due to its smaller size, the gap between surfaces actually won the competition for figural status. However, the important point is that the textured surface on the left side receives higher activity amplitude indicating that it receives precedence when compared with surface on the right side. The left surface might gain figural status when we remove the representation of the gap in the surface network by providing selective inhibition only to cells.
coding the gap area. Such inhibition might arise from top-down or attentional signals which enforces the change in figural assignment. The problem with the gap would not arise if the surfaces would be sufficiently separated but that would not correspond with stimuli used in behavioral experiments. It is interesting to note that the gap area has the same luminance as the background area. Nevertheless, neural activity in the surface network at locations corresponding to the gap does not leak outside into the background area despite the lack of horizontal boundary signals at the ends of the gap. Leakage is prevented by inhibition from the ventral boundary network to the surface network. The inhibition extends from the location of particular boundary segment toward its neighborhood in the same direction as indicated by the preferred orientation of the boundary cell in the ventral boundary network. That is, a boundary signal does not prevent leakage in the surface network only at its particular location but it also sends inhibitory signals to locations in its vicinity. The same mechanism prevents leakage at corners and T-junctions where boundary signals are weak.

When we look at the response of the parietal cortex to surfaces with different spatial frequency content, we see strong boundary response not just at the edges of the surfaces but also in the interior of the surfaces. Such distributed boundary responses to the small luminance increments and decrements are named boundary webs by Grossberg and Mingolla (1987). Here, we showed that sensitivity of the parietal cortex to boundary webs enables the network to prefer surface with higher spatial frequency content. The exact shape of the luminance profile is less important so the model would respond in a similar way for sinusoidal gratings, square-wave gratings or linear luminance staircases. Figure 3C showed that reversing the contrast of the texture of surfaces does not alter the order of their assignment. The same is true when we compare textured surface with the homogenous surface with the same average contrast and size (Figure 3D). Although Klymenko and Weisstein (1986) did not explicitly test configuration shown in Figure 4D it makes ecological sense to choose the surface with high spatial frequency content as a figure because it may contain more information when compared to the homogenous surface. Recently, Palmer and Ghose (2008) showed that gradients of light intensity arising from surface curvature in three-dimensional space produce strong cue for figure-ground

Figure 3. Computer simulations illustrating the model ability to explain the horizontal–vertical advantage and the high spatial frequency advantage in figural assignment. (A) Simulation of the horizontal–vertical principle. Figural status is assigned to the square on the left side due to the fact that its boundaries are aligned in the horizontal and vertical axes. (B) Simulation showing advantage of high-spatial frequency content. A small area between the surfaces becomes a figure, but the surface on the left receives higher activity label compared to the right surface. (C) High spatial frequency advantage holds irrespective of the phase of luminance modulations. (D) High spatial frequency advantage holds in comparison with uniform surfaces.
organization. Luminance gradients will produce a distributed response of oriented simple and complex cells or boundary webs in a similar vein as it was shown in Figures 3B, 3C, and 3D. In the present model, the parietal cortex will detect high density of boundary signals at luminance gradients and strong activity at the parietal cortex will enable the surface with luminance gradients to become a figure. Therefore, the explanation why are luminance gradients effective cue for figure-ground organization is reduced to model’s sensitivity to boundary webs (Grossberg & Mingolla, 1987).

Finally, we should mention that we did not simulate symmetry and parallelism as factors of figure-ground assignment because they require a more complex non-local saliency computation. We suggest that incorporating more sophisticated spatial filters in the model’ parietal cortex should yield sensitivity to symmetry and parallelism. In line with this conclusion, several models of symmetry perception showed that at least bilateral symmetry perception could be explained using spatial filters (Dakin & Watt, 1994; Huang & Pashler, 2002).

Computer simulations presented in Figure 4 showed that the model correctly assigns figural status to the surface at the bottom location and to the surface with a heavy bottom part. Furthermore, it is shown that moving the configuration in the bottom part of the visual display does not change the preference for the lower region. In this simulation we present the full 50 × 50 network. Also, we showed that the contrast reversal does not change the figural assignment in the top–bottom polarity. This is true for the lower region also but this was not shown. Hulleman and Humphreys (2004) argued that the top–bottom polarity is a distinct principle of figure-ground organization from the lower region. However, we simulate both factors using the common neural mechanism. The crucial factor which contributes to the explanation of these principles is a gradient of activity in the parietal cortex. We assumed that the gradient emanates from the tonic input which is differently weighted for different spatial locations. The gradient is made in such a way that the bottom parts of the representation are stronger if compared to the top parts. Different synaptic weighting provides competitive advantage of the bottom parts of the image and induces the tendency to assign a figure to bottom areas. This is true even when both surfaces are positioned in the lower part of the visual field. In that case the lower surface will still receive stronger activity compared to the top surface and it will win the competition. The gradient of activity is not an absolute measure of the preference. It just provides a relative advantage to the

Figure 4. Computer simulations illustrating the model ability to explain new principles of figural assignment. (A) Simulation of the lower region. (B) The lower region holds even in the case where the surfaces are moved in the bottom part of the visual display. (C) Simulation of the top–bottom polarity. (D) The top–bottom polarity does not depend on the contrast of the surfaces. For the display in (C) and (D) we selected a 30 × 30 central square of the network dimensions.
lower surfaces in competition for figural status in the surface network. We do not simulate how this gradient of activity may arise from learning. In the model, all connection weights are fixed. In these simulations we used the same parameter set as in the simulations presented in the Figure 2. Therefore, the model explains the classical and the new cues of figural assignment in a unified way.

Occlusion, context, and closure

Classical and recently discovered principles of figure-ground organization do not exhaust all interesting input configurations. An important constraint for principles of figure-ground organization is that they consider only two-dimensional input configurations. However, in the three-dimensional world objects may occlude each other. Occlusion is signaled by the presence of the T-junctions which are a powerful cue for signaling depth and figural status (Grossberg, 1994, 1997). How the proposed network deals with the occlusion is illustrated in Figure 5. In Figure 5A we presented two overlapping squares where the occluded square had higher contrast when compared to the occluding square. In this case, the principle of size and contrast will favor the occluded surface. However, the model correctly selects the occluding surface due to its sensitivity to corners. When we inspect the response of the parietal cortex, we see strong activity not only at the locations of the T-junctions but also at the location of the missing (or amodal) corner of the occluded surface. The missing corner is not visible due to the occluded surface but its presence is implicated by the boundaries of the occluded surface. Due to the fact that the corner cells have large receptive fields they will show strong response to the missing corner. Furthermore, strong activity of corner cells at T-junctions and at the missing corner will jointly contribute to the competitive advantage of the occluding surface over the occluded surface in the competition for figural status. It may be argued that in Figure 5A the occluding surface receives support from the principle of the lower region which gives higher emphasis to boundaries positioned in a lower part of the visual field. But, in Figure 5B three principles of the figure-ground organization (size, contrast and lower region) are working in favor of the occluded surface. Therefore, it might be considered the worst case scenario for the model. Still, the corner cells in the parietal cortex override all other influences and enable the occluding surface to become a figure. Enlarging the occluded surface in order to cover an even wider area in the lower space will not make a difference (Figure 5C). On the other hand, reducing the size of the occluded surface beyond a certain limit reduces the activity of corner cells and enables the occluded surface to become a figure.

Figure 5. Computer simulations illustrating the model’s ability to assign figural status to the occluding surface. (A) The occluded surface is positioned in the upper hemifield. (B) The occluded surface is position in the lower hemifield. (C) The occluded surface is enlarged. In all three cases figural status is assigned to the occluding surface in the surface network.
Figure 6 illustrates the model’s sensitivity to geometrical properties of the input. In general, the model will exhibit preference for smaller surfaces which is ecologically valid (Fowlkes et al., 2007), but it may alter its preference in specific geometrical circumstances. Figure 6A showed an interesting example of a small black surface placed into the gray ring. What is really seen here is a black surface which is placed on top of a larger gray surface (Gilchrist, 1994). Although the visible gray ring is actually thinner when compared to the width of the black square, the network showed preference for the black surface. The reason for such a choice is strong response of corner cells which override simple density computation by blur cells in the parietal cortex. However, in specific circumstances where a small surface is abutting the larger surface, the surface network changes its preference and allows the larger surface to become a figure (Figure 6B). This is due to the T-junction sensitivity exhibited by corner cells. Perceptually, T-junctions signal occlusion of the small surface by a larger surface. When the smaller surface is detached from the larger surface it will again win the competition for figural status (Figure 6C). Figures 6A and 6C also shows that smaller size overrides the lower region as a principle of figural assignment. That is, the smaller surface is assigned figural status despite the fact that the lower boundary of the larger surface (or ring) is positioned lower in visual space where it is more heavily weighted due to the activity gradient in the parietal cortex. It is not clear how models, based on the border ownership, would respond to these input configurations.

Contextual interactions between abutting small and large surfaces might partially explain the results of Vecera and Palmer (2006). They investigated whether the lower-region preference might be rotated in space so as to influence left-right figural decisions. In their experiment, the left-right ambiguous figure-ground display was attached to receding planes (walls) which induces depth perception either by linear perspective or by texture gradients. Participants reported seeing as ‘figure’ surface that was attached to the receding depth plane. Vecera and Palmer (2006) provide an ecological analysis and justification why such biases should exist. Here, we showed that the same result might be explained using local analysis of edge responses at the border of the ambiguous display and the texture gradient (or linear perspective). Figure 7 illustrates how the ambiguous figure-ground display might be biased by small surfaces placed near it. Small surfaces approximate texture gradients in the display of Vecera and Palmer (2006). The texture gradient that induces depth in horizontal direction must have elements placed differently on the left and on the right side of the ambiguous figure. The near side of the gradient should have texture...
elements placed lower in space. On the other hand, texture elements on the far side must be positioned upper in space. In Figure 7A the small surface on the right is placed lower in space which forces assignment of figural status to the right large surface. In Figure 7B the position of small surfaces is reversed which results in change in the perceived figure. In Figure 7C both small surfaces are placed up relative to the horizontal border of the ambiguous display. Stimulus configuration in Figure 7C would approximately correspond to the control condition in which there is no bias to favor one surface over another. Here, figural assignment is performed merely by chance. In this way, the model simulates psychophysical findings of Vecera and Palmer (2006).

Figure 7. A further illustration of the model’s sensitivity to contextual interactions. (A) When small surfaces are attached to large surfaces, figural status of the large surfaces will be determined by the position of the small surfaces in the vertical dimension. The small surface on the right is lower in space and the large surface on the right side becomes a figure. (B) The small surface on the left side is lower in space and it forces large surface on the left side to become a figure. (C) The contextual cue provided by small surfaces is valid only if they are positioned low enough in space. If small surfaces are positioned too high in space, they will not be able to influence the figural assignment of the large surfaces. Here, the figure is determined by chance. In this way, the model simulates psychophysical findings of Vecera and Palmer (2006).

Figure 8 illustrates how the model handles connectedness and closure. They are particularly effective forces in organizing perception. For instance, connectedness and closure may override proximity (Elder & Zucker, 1993; Palmer & Rock, 1994). In Figure 8A three evenly spaced vertical lines are shown. The network assigns figural status to the line in the middle of the display due to the fact that it receives the strongest support from the parietal cortex. The left and the right lines receive less support and they are labeled with a lower activity level. It is interesting to note that lines on the left and right sides are represented by different activity amplitudes. Despite the fact that they are identical, they are treated as different surfaces by the network. According to the present
approach, even thin lines are actually perceived as surfaces. This point was further elaborated in Grossberg (1994). Figure 8B shows what happened when two vertical lines became connected by horizontal lines in order to form an outline of the square. In this case, the left and the middle vertical lines become a part of the same surface (outline) and receive the same activity amplitude in the surface network. In this way, the model implements sensitivity to element connectedness (Palmer & Rock, 1994). Also, the interior of the square becomes a separate surface. Moreover, the interior of the square becomes a figure. On the other hand, the representation of the line on the right side is completely erased in the surface network due to strong inhibition from the square and its outline.

Figure 8B illustrates the fact that the interior of the shape as well as the outline of the same shape might be used for object recognition because both are represented as separate surfaces. Figure 8C shows that the interior of the square remains to be seen as a figure even if we introduce a small gap in the horizontal lines of the square. Neural activity in the surface network at locations corresponding to the interior of the square does not leak outside into the background area despite the lack of boundary signals at gaps. An explanation for this surface network property is provided in the analysis of Figures 3B, 3C, and 3D where we also observed similar phenomenon. Shortly, leakage is prevented by extended inhibition provided by the ventral boundary network. However, this accounts only for small gaps. At larger gaps, boundary signals will not be able to prevent leakage and some other more global boundary processes, such as sensitivity to illusory contours, are needed (Grossberg & Mingolla, 1985). Of course, the presented analysis does not fully account for what is seen in the case of Figure 8C. For instance, line ends created by the gaps produce perception of the illusory rectangle covering the central part of the occluded square. Furthermore, the outline of the square appears to be amodally completed behind the illusory figure. Our simplified version of the boundary computation does not incorporate sensitivity to illusory contours and it cannot account for such percept. The present model might be augmented by incorporating a new network layer which will contain cells sensitive to collinear but non-adjacent boundary signals, such as bipole cells. The output of the bipole cells should contribute to both processing streams. First, it should prevent leakage of neural activity in the surface network at locations of illusory contours. Secondly, it should influence saliency computation in order to give competitive advantage to the illusory figures.
Attention

Figure 9 illustrates how exogenous attention may bias figural assignment in the model. We implemented exogenous attention with the assumption that the spatial cue is presented for a short period of time, so it does not leave a trace on boundary computation in the ventral stream but its impact is sensed in the parietal network, which will bias the competition in the surface network. In other words, the cue will not be represented in the final surface representation because it is removed from the input and there is no boundary signal left which can contain activity propagation in the surface network in the region where it was presented. However, its influence will be seen in the surface network indirectly because even brief exposure to the cue can bias the network to prefer one surface over the other. Consider, for instance, the top row where the cue is presented on the left side from the two identical surfaces (i.e., black and white surfaces in the input). By ‘identical surfaces’ we mean surfaces identical in shape and contrast which makes them equal candidates for figural status. Due to the fact that the cue is presented away from both surfaces it will bias the network to select the gray background as a figure (Figure 9A). This surface is not informative and it should be inhibited in order to focus on the remaining two surfaces in the middle of the display. After the representation of the gray background in the surface network is removed by inhibition, the remaining surfaces have an equal chance of becoming a figure (Figure 9B). On the other hand, when the cue is presented on one of the two surfaces in the middle it will bring the competitive advantage to that surface and it will become a figure (Figure 9C). In this way we explain why Baylis and Driver (1995; Driver & Baylis, 1996) did not find evidence of exogenous attention influencing figural assignment. In accordance with the behavioral results of Vecera et al. (2004), the spatial cue can influence figure-ground assignment only when it is presented on the surface, but not merely near the surface which is cued to become a figure.

Figure 10 illustrates how endogenous attention may reverse the figure and the background. The input is an ambiguous figure which does not possess any cue to favor the left or the right surface. In this case, the figural status is randomly assigned to one of the surfaces. Here, we do not show activity in all network layers, rather, we focus on the surface network and how neural activity in the surface network evolves with time. Therefore, in Figure 10 the columns represent different points in time with the leftmost column representing the start of the simulation (T1) and the rightmost column corresponds to the final network state (T4). After the surface network settled into the stable solution indicating the figural status of the surfaces (Figure 10A), we applied a spatially restricted but
strong signal in the center of the background surface (Figure 10B) which corresponds to the internal (or voluntary) decision to look at the background. The endogenous signal is actually a single pixel, although other shapes (like the Gaussian spread) of the top-down influences will produce the same result. The location where the endogenous signal is applied is chosen randomly. Any other location on the background surface will lead to the same result. The important point is that the endogenous signal must be strong in order to change activity distribution in the network. More precisely, strong signal means stronger than maximal level of activity displayed by the cells coding the figural region. Only if this condition is met, the top-down signal will be able to bias the surface network into changing the figural status of surfaces. The reason for this is dendritic inhibition which protects cells from lateral inhibition. When the top-down signal allows a small number of neurons (one neuron in this simulation) representing the background surface to attain a stronger activity level (compared to the activity level of the figural region) these neurons will be protected by their own dendritic inhibition, so the cells in the figural region cannot inhibit them. On the other hand, the cells in the background region start to inhibit cells representing the figure. At the same time, local excitatory signals enable the spreading of the top-down signal on the whole background surface. As more and more background cells become active they induce stronger inhibition of the figure. At the end of this process, the previous background surface is now a figure and a previous figural surface is now background (Figure 10B). This simulation also illustrates how the model implements object-based selection (Driver, Davis, Russell, Turatto, & Freeman, 2001) and elemental visual operators such as tracing and region filling or labeling (Roelfsema, 2006). Excitatory horizontal connection in the surface network enables activity spreading in a way that the whole surface is selected. Evidence that such activity spreading may exist in the primary visual cortex was provided by Roelfsema, Lamme, and Spekreijse (1998).
Furthermore, we showed how a spontaneous reversal may occur in the model. Figure 10C illustrates what happens when the excitatory neurotransmitter mediating the self-excitation habituates. The surface network contains self-excitatory connections which enable the network to store the activity pattern in the working memory after the input ceases. The input from the parietal cortex is considered to be transient and the surface network needs such mechanism in order to sustain the surface representation. Habituation in the self-excitatory pathway occurs only for cells which are maximally excited, that is, the cells coding the figural region. As can be seen, the figure gradually loses its competitive advantage due to the passive decay which drives a cell’s activity to a resting state. Precisely speaking, the activity in the figural regions drops while the activity in the background region remains at the same level (T1 and T2). Only after the activity in the figural area drops below the activity level in the background area it is possible for the cells’ activity in the background area to start growing until they reach the upper bound. At the moment when the figure and the background have the same activity level (T3) there is no interference between surfaces because boundary signals prevent any interaction despite the fact that they are indistinguishable at that moment. Therefore, the switch between surfaces occurs quickly and smoothly. After a while, neurotransmission at the self-excitatory circuit is restored and previously figural region is now becoming a background. The same process of habituation may now occur at the newly established figure leading to another reversal. In this way, it is possible to simulate bi-stable perception of ambiguous figures. Together, Figures 9 and 10 showed how exogenous and endogenous attention change the figural status of the surfaces. They illustrate how the model provides biologically plausible implementation of the idea that top-down signals related to the attention have an active role in figure-ground assignment.

Discussion

Computer simulations showed that the proposed model is able to account for a wide range of classical and new effects on figure-ground process. The effects of surround- edness, convexity, size, contrast, horizontal–vertical orientation, and spatial frequency are explained by the properties of the neurons in the parietal network. The model’s parietal cortex contains two types of spatial maps: blur cells and corner cells. The blur cells have large concentric receptive fields which detect the total amount of activity or energy. A stronger response of the blur cells is observed in the area where a greater number of boundary signals exist or where the contrast of boundary signals is higher. On the other hand, corner cells are sensitive to coincidence of horizontal and vertical boundary segments. A strong response of corner cells is observed at modal or amodal corners and T-junctions. They enable the model to handle more complex input configurations such as occlusions, contextual interactions between figure and background texture, and the effect of element connectedness and closure on figural assignment (Elder & Zucker, 1993; Palmer & Rock, 1994; Vecera & Palmer, 2006). Together, blur cells and corner cells contribute to the model’s sensitivity to the boundary web, that is, spatially distributed boundary responses to luminance gradients (Grossberg & Mingolla, 1987). Luminance gradients are shown to be particularly effective in determining figure-ground organization (Palmer & Ghose, 2008).

Recently discovered principles of the lower region and the top–bottom polarity are explained by the gradient of activity in the parietal cortex, which is projected via feedback connections to the surface network. The gradient is designed so as to favor surfaces in the lower part of the visual field. Other things being equal, the gradient will bring competitive advantage to the surface that is located at a lower position or to the surface with a wider lower part. Existence of such a gradient is supported by psychophysical investigations which showed different sensitivities of the human visual system in the lower and the upper part of the visual field (Levine & McAnany, 2005; McAnany & Levine, 2007; Previc, 1990; Rubin et al., 1996). A recent brain imaging study of perceptual filling-in revealed stronger neural activity in the parietal cortex when the stimulus is presented in the lower visual field consistent with our model (Mendolla, Conner, Sharma, Baehekar, & Lemieux, 2006). Moreover, behavioral measure of the filling-in showed asymmetry in the lower and the upper visual field. More specifically, filling-in was reported to be easier and takes less time in the lower visual field. Based on these findings, Mendolla et al. (2006) suggested that preferential representation of the lower visual field in the parietal cortex induces bias in attentional control mechanisms which influence perceptual tasks. Neuropsychological evidence support this conclusion by showing that patients with visuo-spatial neglect exhibit more pronounced impairment in the lower than in the upper visual field (Ládavas, Carletti, & Gori, 1994). Single-unit recordings in monkeys also indicate asymmetry in neural activity related to upper and lower visual fields (Maunsell & Newsome, 1987).

We modeled differences between upper and lower hemifields as a smooth increase in the spatial representa- tion in the parietal cortex from the top to the bottom of visual space. Similar gradients were already used in the modeling literature. For instance, Groh (2001) used a simple linear function between synaptic weights and locations in a model of transformation from the sensory place code to the motor rate code. A direct anatomical test of this assumption is provided by Moschovakis et al. (1998) who studied the cat’s oculomotor system. They discovered a graded strength of anatomical projections.

Downloaded from jov.arvojournals.org on 05/07/2019
from distinct superior colliculus cells to burst generators for evoking saccadic movements. Models of visuo-spatial neglect also used the gradient of synaptic weights which runs from the left to the right on the spatial map in the parietal cortex (Monaghan & Shillcock, 2004; Pouget & Sejnowski, 2001). Here, we propose that the same type of computational mechanism could exist in the vertical dimension too.

Our model explains how exogenous and endogenous attention may change figural assignment. Again, the crucial component for explaining these effects is the parietal network which may mediate both exogenous and endogenous attention. The model explains why the Baylis and Driver (1995) did not find evidence that cueing can change the figural assignment in an ambiguous figure display. Consistent with the findings of Vecera et al. (2004), the model assigns figural status to the cued surface only when the cue is presented on the surface and not merely near the target surface. When the cue is presented on the gray background, this background will be selected as a figural region in the first place. After removal of the background from the representation in the surface network, remaining surfaces have equal chance to become the figure. This is due to the fact that there is no other factor that would favor one surface over the other. The model also demonstrates how voluntary or endogenous attention may reverse the figural assignment. Previous models of figure-ground organization which consider attentional influences were focused on top-down (or endogenous) attention only (Craft et al., 2007; Kienker, Sejnowski, Hinton, & Schumacher, 1986; Li, 2005). They did not consider distinction between endogenous and exogenous attention and how external attentional cues might influence figural assignment (Vecera et al., 2004).

Vecera (2000) proposed an interactive account of the figure-ground organization where attention does not operate on the output of the figure-ground process, but it is directly involved in the decision which surface should be the figure. This is an extension of the biased competition account of the attentional influence on the neural activity to the perceptual organization (Desimone & Duncan, 1995). However, his assertions were not corroborated by specific neural implementation. We claim that our model implements the idea of object-based biased competition in the surface network. The surface network combines the bottom-up boundary information and the top-down signal from the parietal cortex in order to arrive at a correct figural assignment.

Anatomical and physiological considerations

We presented a new architecture for visual processing based on neurophysiological findings about the interaction between the ventral and the dorsal processing stream. Interactions between streams occur at feedforward or horizontal connections as they are evident in more complex types of receptive fields with sensitivities to combination of attributes such as color-orientation cells or color-disparity cells (Sincich & Horton, 2005; Ts’o & Roe, 1995). Another possibility is that streams communicate via feedback connections from higher visual areas to V1 and V2. Hupé et al. (1998, 2001) discovered that MT area may directly influence figure-ground organization as measured by the difference in the activity of V1 and V2 neurons. They inactivated MT while the monkey was looking at the moving bar on a static background. In this case, saliency is almost completely determined by the motion signals. Many cells in V1 and V2 showed up to 40% reduction in the firing rate following MT inactivation. Based on these findings, Bullier (2001) developed an integrated model of visual processing where ventral stream and feedback connections from dorsal stream jointly contribute to the figure-ground assignment. First, the fast feedforward connections enable the computation of saliency in the parietal cortex. The activity from the parietal cortex is retro-injected into V1 and V2 which serve as active blackboards. In his model of visual search, Hamker (2003, 2004) proposed a similar idea of reentry of signals from the dorsal stream into the ventral stream. He suggested that activity in the oculomotor areas, which are responsible for planning eye movements, sensitizes cells in the ventral stream prior to eye movement is performed.

Recent models of figure-ground organization are focused on border ownership assignment (Craft et al., 2007; Jehee et al., 2007; Sakai & Nishimura, 2006; Zhaoping, 2005). They are based on recent findings that cells in monkey V2 show different responses depending on the relative position of the figure with respect to the figure-ground border (Qiu & von der Heydt, 2005; Zhou et al., 2000). However, physiological studies did not test many important stimulus configurations which are known to affect figure-ground organization. For instance, it is not clear whether the border ownership cells are sensitive to lower region, top–bottom polarity or whether they will exhibit preference for horizontal–vertical axis. Furthermore, by focusing on the borders alone, important information available on the surfaces is ignored as demonstrated by the effect of spatial frequency content on figure-ground organization (Klymenko & Weisstein, 1986). Recently, Palmer and Ghose (2008) showed that luminance gradients arising from surface curvature are powerful cues for figural assignment. It is not clear how the models of border ownership would handle stimuli like those shown in Figures 3, 6, and 7.

Several models of border ownership rely on integration of local boundary signals using large-scale detectors (Craft et al., 2007; Jehee et al., 2007; Sakai & Nishimura, 2006). A similar approach was used in the present model with cells in the parietal cortex. However, in the presented model, inputs to the parietal cortex are fuzzy boundary signals from the dorsal stream and not sharp boundary signals from the ventral stream. In this way, present model achieves better sensitivity to the weak boundary signals.
arising from the interior of the surfaces which helps to explain spatial frequency effect and attentional cuing effect on figural assignment.

Filling-in

Our model is based on previous work of Grossberg and colleagues (Grossberg, 1994, 1997; Grossberg & Mingolla, 1985, 1987; Grossberg & Todorović, 1988) who provide a detailed computational analysis of how boundary information and surface information complement each other. Several psychophysical investigations confirm the existence of separate perceptual representations for boundaries and surfaces (Elder & Zucker, 1998; Gilchrist, Humphreys, Riddoch, & Neumann, 1997; Rogers-Ramachandran & Ramachandran, 1998). In the model, boundary computation creates an outline of the input image enhanced with the representation of illusory contours and global grouping of boundary segments. We used a simplified version of boundary computation without sensitivity to illusory contours and global boundary configurations. In the original model, the final computational step, which provides surface representation, is called filling-in. It allows spreading of brightness and color signals in the homogenous areas within boundary outlines. The model’s filling-in stage is considered to be a neural implementation of the perceptual filling-in which occurs at blind spots, stabilized images, and neon color spreading (Pessoa, Thompson, & Noé, 1998). We suggest that, besides brightness and color filling-in, there should be a more abstract form of filing-in which allows figural label spreading along the whole surface. Such a coding scheme assumes that the visual cortex must contain a special group of cells sensitive to homogenous surfaces. Physiological assumptions of the model were critically examined in the next section.

Limitations of the model

Although the proposed model addressed many psycho-physical aspects of figure-ground assignment, the model of the surface network makes several assumptions not supported by neurophysiological investigations. First, there is no direct evidence that surfaces are encoded in the visual cortex with different firing rates. Several studies found cells in V1 which are responsive to surfaces with uniform luminance (Kinoshita & Komatsu, 2001; MacEvoy et al., 1998; Rossi et al., 1996). Sensitivity to luminance alternation extends 10 to 20 degrees of visual angle beyond the classical receptive field. In a similar vein, it was shown that the textured surface increases the firing rate of neurons coding the interior of the texture (Lamme, 1995; Lee, 2003; Zipser et al., 1996). However, activity modulation in the figural region is modest and it does reach maximal firing rates as postulated by the model. Furthermore, Rossi et al. (2001) suggest that interior enhancement might be an artifact of small stimulus sizes used in the experiments. They employed larger stimuli and they did not find evidence for activity enhancement corresponding to the interior of the figure.

Second limitation for the surface network is that its operation is based on the process of neural filling-in or spreading activation. Perceptual filling-in could be conceptualized as a manifestation of neural activity spreading in early topographically organized visual areas or as a symbolic interpretational process operating on higher visual areas (Pessoa et al., 1998). A recent review suggests that evidence for neural filling-in is ambiguous and that there might be different neural mechanisms related to different types of filling-in percepts (Komatsu, 2006). Single-unit recordings in monkeys showed increased neural activity in V1 region corresponding to the blind spot when the stimulus partially covered the blind spot (Fiorani, Rosa, Gattas, & Rocha-Miranda, 1992; Komatsu, Kinoshita, & Murakami, 2000; Matsumoto & Komatsu, 2005). Craik-O’Brien-Cornsweet stimulus induces increased neural activity in V2 (but not in V1) in homogenous regions corresponding to the filling-in of color information (Roe, Lu, & Hung, 2005). On the other hand, von der Heydt, Friedman, and Zhou (2003) did not find any evidence for a change in neural activity in V1 and V2 related to filling-in percepts. Brain imagining studies using fMRI on humans are also inconclusive. Several studies revealed changes in neural activity in early visual areas V1 and V2 consistent with filling-in percepts (Boyaci, Fang, Murray, & Kersten, 2007; Meng, Remus, & Tong, 2005; Sasaki & Watanabe, 2004), while others report no such activity change (Cornelissen, Wade, Vladusich, Dougherty, & Wandell, 2006; Perna, Tosetti, Montanaro, & Morrone, 2005). We suggest that some of the discrepancies between the studies might arise from the fact that they are looking only for brightness or color filling-in and they do not consider the possibility of figural filling-in. In these experiments, usually, only brightness appearance of the surface is changed, but the figural status of the surface remains constant. Therefore, absence of the change in neural activity related to the change in the surface appearance should not be taken as evidence that filling-in does not occur. Unchanged neural activity may simply signal that the figure-ground relationship remained the same. Further studies are needed to disentangle brightness or color filling-in and the figural filling-in.

Third limitation for the surface network is neural activation spreading speed. Physiological studies reveal very short latencies (about 30 ms from the stimulus onset) of the border ownership responses in V2 (Zhou et al., 2000). Moreover, response latencies are invariant to changes in the size of the stimulus. Based on the finding that propagation speed in intra-cortical axons in monkey V1 is about 0.3 m/s (Girard, Hupé, & Bullier, 2001), Craft et al. (2007) calculated that the model based on horizontal connections would need more then 70–90 ms to propagate
figural signals along the borders of the stimulus whose size is 8° of visual angle. This estimation is based on the fact that maximal length of the horizontal connections is about 3–4 mm (Angelucci et al., 2002). On the other hand, in the surface network cells are allowed to make contact only with their nearest neighbors. Otherwise, crosstalk might occur between representations of different surfaces. Therefore, in the surface network, neural activity will need to propagate along the longer cascade of cells compared to the estimation of Craft et al. (2007). Furthermore, shorter axons might exhibit slower signal propagation (Li, 2005). All these facts suggest that the total time needed for the surface network to reach the stable state would be well above 100 ms. Moreover, larger surfaces would require longer time to complete activity spreading.

Although the behavior of the surface network contradicts physiological findings about border ownership responses in V2, it is interesting to note that the model is consistent with psychophysical investigations which show that surface representation construction actually requires time (Paradiso & Hahn, 1996; Paradiso & Nakayama, 1991). Recently, Hung, Ramsden, and Roe (2007) attempted to measure directly whether neural activity in V1 really spreads from borders to the interior of the surface representation. They used stimuli of various sizes and found that enhanced activity in the representation of the interior of the surface appears later in time for larger surfaces exactly as it would be predicted from the surface network. However, in their experiment, activity spreading was not related to the brightness perception as there was no evidence that the neural activity changes in accordance with perceptual experience. It should be mentioned that Hung et al. (2007) did not consider the possibility that the filling-in process could be related to figure-ground organization because the figural status of the stimuli was not changed during the experimental trials.

Finally, the surface network heavily relies on inhibitory mechanisms. It contains two distinct classes of inhibitory cells for every excitatory cell which is contradicted by the anatomical finding that about 80% cells in the visual cortex are pyramidal cells and only 20% are inhibitory interneurons (Gabbott & Somogyi, 1986). Moreover, excitatory cells make most of the synaptic contacts with other excitatory cells and not with inhibitory interneurons (Braitenberg & Schuz, 1991). Furthermore, the model assumes that inhibition is distributed with equal strength to all network locations without distance fall-off. It is an open research question whether it is possible to construct model with similar behavior but more consistent with physiological constraints described above.

**Depth**

The presented model does not have the ability to represent depth because it is single-scale architecture. Although it was not explicitly treated as a principle of figure-ground organization, it is clear that the figure is perceived as closer to the observer when compared to the background. Grossberg (1994, 1997) provided a detailed account of depth and surface perception where the figure is segregated from the background as a part of the process of surface construction in different depth planes. We suggest that the single-scale surface network presented here could serve as an input to the multi-scale network for depth stratification. Information about depth might be recovered from the activity amplitudes in the surface network where stronger activity encodes closer surface (figure-near, ground-far). In this case, the process of figure-ground organization is completed before the depth assignment, which is in contrast with previous proposal (Grossberg, 1994, 1997). Therefore, our model makes a testable prediction concerning the relation between the depth perception and figure-organization. Vecera et al. (2002) in Experiment 7 provided evidence that the depth perception could be the cause of the principle of the lower region. However, we showed how this finding can be accommodated within our single-scale model. We used the gradient of synaptic weights as a bias toward the lower region. Therefore, we disagree with their assertion that the results in Experiment 7 implied precedence of depth perception relative to the figural assignment. A stronger test should be performed in order to disentangle these two contrasting hypotheses.

**Conclusion**

In conclusion, we presented a model of figure-ground assignment which is able to account for how classical factors such as size, contrast, surroundedness and convexity and recently discovered principles of the lower region and the top–bottom polarity influence the perception of the figure. The model is based on the interaction between two complementary processing streams. The ventral stream computes sharp object boundaries which are used to construct surface representation. On the other hand, the dorsal stream computes saliency based on the blurred version of the boundary signals. The dorsal stream provides feedback to the process of surface construction. The surface network enables all neurons that code the same surface to attain the same activity level due to the local excitatory connections. Also, the surface network enables different surfaces to be represented with different activity levels by global long-range inhibition which implements an object-based biased competition. Our model shows how the combination of feedforward, feedback and horizontal (or recurrent) interactions enable the visual system to decide which surface should be a figure based on the combination of bottom-up and top-down cues.
Appendix A

We used two types of model neurons: additive and multiplicative. Model neurons in LGN and the surface network are described with shunting (multiplicative) non-linear differential equations which approximate membrane potential of real neurons. In this model, neurons have upper and lower activity bounds. On the other hand, neurons in the primary visual cortex and the parietal cortex are modeled as feedforward summation units whose output is proportional to the result of convolution of their spatial masks with input. For simplicity, we omitted lateral (or recurrent) connections in V1 and the parietal cortex.

LGN

Inputs, $I_{ij}$, may assume values between 0 (black surface) and 2 (white surface). An exception to this is a cue used in modeling exogenous attention (Figure 9). The cue was set to $I_{ij} = 4$ which reflects the fact that the sudden appearance of a new object in the visual scene attracts attention and it gets priority over other objects in the visual scene. The input is first processed by on-center off-surround and off-center on-surround cells. These cells detect local luminance discontinuities and suppress the uniform luminance areas. Also, they compute local luminance ratios of neighboring surfaces by using the shunting mechanism. Following Grossberg and McLoughlin (1997) activity of on-center, $x_{ij}^+$, and off-center, $x_{ij}^-$, cells are described with

$$\frac{dx_{ij}^+}{dt} = -A_1 x_{ij}^+ + (B_1 - x_{ij}^+) C - (x_{ij}^+ + D_1) S$$  \hspace{1cm} (A1)$$

$$\frac{dx_{ij}^-}{dt} = -A_1 x_{ij}^- + (B_1 - x_{ij}^-) C - (x_{ij}^- + D_1) S$$  \hspace{1cm} (A2)$$

where $C(S)$ denotes the central (surround) part of the on-center cell’s receptive field. Off-center cells are modeled as a spatial inverse of on-center cells so their center is equivalent to the surround of on-center cells and vice versa. Parameter $A$ describes passive decay which drives activity toward zero if there is no input; $B(D)$ defines the excitatory (inhibitory) saturation point, that is, an upper (lower) bound for an activity level that can be obtained. The center and the surround part of receptive fields are obtained by convolution of spatial filters with input

$$C = \sum_{(p,q)} C_{pq} I_{i+p,j+q}$$ \hspace{1cm} and \hspace{1cm} $$S = \sum_{(p,q)} S_{pq} I_{i+p,j+q}, \hspace{1cm} (A3)$$

where $C_{pq}$ and $S_{pq}$ are Gaussian spatial filters given by

$$C_{pq} = \frac{M_1}{2\pi\sigma_1^2} \exp\left\{-\frac{1}{2\sigma_1^2} (p^2 + q^2)\right\}, \hspace{1cm} (A4)$$

and

$$S_{pq} = \frac{M_2}{2\pi\sigma_2^2} \exp\left\{-\frac{1}{2\sigma_2^2} (p^2 + q^2)\right\}. \hspace{1cm} (A5)$$

$-P_1 < p, q < P_1$ defines the size of the receptive field. $M_1$ ($M_2$) is a constant describing height, and $\sigma_1$ ($\sigma_2$) describes width of the Gaussian filter for the central (surround) part of the receptive field. Due to the fact that LGN is modeled as a feedforward network, equations for the activity of on-center and off-center cells are solved at equilibrium, $dx_{ij}/dt = 0$. In simulations, they are represented in algebraic form

$$x_{ij}^+ = \frac{\sum_{(p,q)} BC_{pq} I_{i+p,j+q} - DS_{pq} I_{i+p,j+q}}{A + \sum_{(p,q)} (C_{pq} + S_{pq}) I_{i+p,j+q}}, \hspace{1cm} (A6)$$

and

$$x_{ij}^- = \frac{\sum_{(p,q)} BS_{pq} I_{i+p,j+q} - DC_{pq} I_{i+p,j+q}}{A + \sum_{(p,q)} (C_{pq} + S_{pq}) I_{i+p,j+q}}. \hspace{1cm} (A7)$$

The non-linear shunting mechanism of LGN cells allows ratio computation which is necessary in order to achieve invariant response in conditions of varying illumination. The difference between the activity of on-center and off-center cells is used as an output from the LGN to simple cortical cells,

$$X_{ij}^+ = \left[x_{ij}^+ - x_{ij}^-\right]^+, \hspace{1cm} (A8)$$

and

$$X_{ij}^- = \left[x_{ij}^- - x_{ij}^+\right]^+. \hspace{1cm} (A9)$$

$[x]^+ = \max(x, 0)$ describes the half-wave rectification. The parameters for on-center and off-center cells were set to: $A_1 = 10$, $B_1 = D_1 = 1$, $P_1 = 4$, $\sigma_1 = 0.5$, $\sigma_2 = 1.5$, $M_1 = 1$, $M_2 = 1.03361$. At the level of LGN we did not make a distinction between the magnocellular and the parvocellular pathways for computational convenience. Branching between the processing streams starts at the level of simple cortical cells.
Simple and complex cells

Simple cortical cells exhibit orientation selectivity which is simulated by convolution of the oriented spatial filter with the input from LGN. The difference-of-offset-Gaussian is used as a spatial filter to detect oriented boundaries. Simple cells have two poles, the right and the left, making simple cells sensitive to the direction of contrast,

\[ R_{rjk} = \sum_{(p,q)} \left[ \text{DoG}_{pq}^k \right]^+ \left( x_{i+p,j+q}^+ \right) \] (A10)

\[ L_{rjk} = \sum_{(p,q)} \left[ -\text{DoG}_{pq}^k \right]^+ \left( x_{i+p,j+q}^- \right). \] (A11)

\(-P_2 < p, q < P_2\) defines the size of the receptive field, and \(\text{DoG}_{pq}^k\) is an oriented spatial filter given by

\[ \text{DoG}_{pq}^k = \exp \left[ -\frac{1}{2\pi\sigma_3^2} \left( (p + \delta \cos\theta)^2 + (q + \delta \sin\theta)^2 \right) \right] - \exp \left[ -\frac{1}{2\pi\sigma_3^2} \left( (p - \delta \cos\theta)^2 + (q - \delta \sin\theta)^2 \right) \right]. \] (A12)

The angle of rotation of the receptive field is described with \(\theta = \pi(k - 1)/K\), where \(k\) ranges from 1 to 2\(K\), and \(K\) is the total number of orientations. Following Lanyon and Denham (2004) we used two spatial resolutions, \(r\), which corresponds with two different settings of \(\delta\), and \(\sigma_3\), which controls the sharpness of the receptive field. For \(r = 1\), we set \(\delta = \sigma_3/2\) and \(\sigma_3 = 0.5\) in order to simulate a low resolution processing in the dorsal stream. For \(r = 2\), we set \(\delta = \sigma_3/4\) and \(\sigma_3 = 2\) in order to simulate a high resolution processing in the ventral stream. Other parameters were equal for both processing streams and were set to \(P_2 = 5\) and \(K = 12\). The response of the simple cell is described as

\[ \text{SIM}_{rjk} = \alpha \left[ R_{rjk} + L_{rjk} - |R_{rjk} - L_{rjk}| \right]^+ \] (A13)

The scaling parameter is set to, \(\alpha = 12\). The activity of simple cells with opposite contrast polarity at the same spatial position are summed into a complex cell, whose response is insensitive to the contrast polarity,

\[ \text{COM}_{rjk} = \text{SIM}_{rjk} + \text{SIM}_{rjk}(k + K). \] (A14)

Outputs from the complex cells which are sensitive to all orientations at the same spatial position \(i, j\) are combined

\[ T_{r} = \sum_{k} \text{COM}_{rjk} \] (A15)

in order to provide the total input into the parietal cortex, \(T_{1ij}\), and into the surface network, \(T_{2ij}\).

The parietal cortex

The parietal cortex computes the saliency based on the boundary signals from V1. The activity in the parietal cortex is given by

\[ P_{ij} = \text{Blur}_{ij} + \text{Cor}_{ij} + J_{ij} + R_{ij}. \] (A16)

Terms, \(\text{Blur}\) and \(\text{Cor}\) denote two spatial maps whose goal is to provide information about density of boundary signals and presence of corners, respectively. The \(\text{Blur}\) is described with

\[ \text{Blur}_{ij} = \sum_{(p,q)} G_{pq} T_{1,i+p,j+q} \] (A17)

where \(-P_3 < p, q < P_3\). \(G_{pq}\) is a Gaussian kernel defined by

\[ G_{pq} = \frac{1}{2\pi\sigma_4^2} \exp \left\{ -\frac{1}{2\sigma_4^2} \left( p^2 + q^2 \right) \right\}, \] (A18)

with \(\sigma_4\) denoting the width of the parietal cell’s receptive field. Sensitivity to corners is described as

\[ \text{Cor}_{ij} = \sum_{(p,q)} H_{pq} g \left( T_{1,i+p,j+q} - \omega_1 \right) \times \sum_{(p,q)} V_{pq} g \left( T_{1,i+p,j+q} - \omega_1 \right), \] (A19)

where \(-P_4 < p, q < P_4\). Horizontal kernel, \(H_{pq}\), is given by \(H = 1\) if \(p < 2\) and \(H = 0\) if \(p \geq 2\) for all \(q\). Vertical kernel, \(V_{pq}\), is defined analogously by \(V = 1\) if \(q < 2\) and \(V = 0\) if \(q \geq 2\) for all \(p\). Before the convolution with the horizontal and the vertical kernels takes place, the dorsal boundary signals, \(T_{1ij}\), are pre-processed by non-linear function, \(g()\), defined as \(g(r) = 1\) if \(r > 0\) and \(g(r) = 0\) if \(r \leq 0\). Term, \(\omega_1\), describes a threshold for decision whether a boundary signal is present or it is not present at particular location. Strong non-linearity is used here in order to make corner cells insensitive to contrast variations. Therefore, they are responsive only to geometrical properties of the boundary signals (presence of corners or T-junctions). The gradient of activity, \(J_{ij}\), is given by

\[ J_{ij} = \beta * j \text{ for } i = 1, \ldots, M, \] (A20)
and $R_{ij}$ represents random value drawn from a uniform distribution in the range $[0,1]$. The gradient, $J_{ij}$, simply increases the activity of the cells progressively from the top to the bottom locations allowing surfaces in the lower part of the visual field to become surfaces. Parameters for the parietal network were set to the following values: $P_3 = 12$, $P_4 = 8$, $\sigma_4 = 4$, $\omega_1 = 5$, $\beta = 0.1$. The width of the receptive field for Gaussian kernel is set to a relatively small value just to illustrate parametric robustness of the model. Any higher value will also yield the same results. Only if the width is set to a value $<1$, blur cells would not be able to bring competitive advantage to the correct surface. On the other hand, parameter choice for corner cells is more restrictive in order to obtain desired results.

The surface network

The surface network encodes figure and ground surfaces by different activity amplitudes or firing rates. The figure is always a surface with the maximal amplitude. The surface representation is constructed by combining sharp boundary signals from the ventral stream and saliency signals from the parietal cortex (or the dorsal stream). The activity in the surface network is given by

$$\frac{dy_{ij}}{dt} = -A_2y_{ij} + (B_2 - y_{ij}) \left( P_{ij}(t) + w_{ij}[y_{ij}]^+ \right) - (D_2 + y_{ij}) \sum_{pq} g\left( [y_{pq}]^+ - [y_{ij}]^+ - E_{ij} \right).$$

(A21)

Parameters $A_2$, $B_2$, and $D_2$ have the same physiological interpretation as in Equations A1 and A2. The transient input from the parietal cortex, $P_{ij}(t)$, initializes network activity, which is allowed to spread within boundaries set up by the signals from the ventral stream. It is defined as $P_{ij}(t) = P_{ij}$ if $t < t_m$ and $P_{ij}(t) = 0$ for $t \geq t_m$ where $t_m > 0$. Self-excitation, $y_{ij}$, enables the cell to remain active after the input from the parietal cortex ceases. If there is no inhibition, the cell will be driven to the saturation level, $B_2$. The strength of self-excitation is controlled by synaptic weights, $w_{ij}$. In the model, there are two types of inhibition: lateral inhibition and dendritic inhibition. Lateral inhibition between cells is given by $y_{pq}$. We assumed a network with full connectivity and the sum is taken over the whole network, $p = 1, \ldots, M$; $q = 1, \ldots, N$ where $M$ and $N$ are network dimensions. Function $g()$ is same as defined previously in the description of the parietal cortex. Dendritic inhibition, $-y_{ij}$, arising from the target cell inhibits lateral inhibitory influences from the other cells in the network. In this way, dendritic inhibition disables inhibition from the cells which have the same or lower activity level relative to the target cell. The interaction between lateral and dendritic inhibition allows cells coding different surfaces to attain different activity levels. Therefore, the firing rate is used in the model as a representational format for objects or surfaces. A detailed description of the surface network and its operation is given in Domijan (2004). Term, $E_{ij}$, describes interaction between boundary signals and local excitation among neighboring cells responsible for activity spreading,

$$E_{ij} = \sum_{mn} V_{mnij}2\left([y_{mn}]^+ - [y_{ij}]^+ - \omega_2\right)^+,$$

(A22)

where indexes $m$ and $n$ are four nearest neighbor locations defined as $m = \{(i+1, j), (i-1, j)\}$, $n = \{(i, j+1), (i, j-1)\}$. Excitation, $y_{mn}$, is allowed to influence the target cell only if $V_{mnij} = 1$, which is true in the case when there is no boundary signal. Excitatory signals from neighboring cells are also subject to dendritic inhibition from the target cell, $-y_{ij}$, which prevents excessive excitation. Term, $\omega_2$, is a threshold which allows activity spreading when the activity difference between $y_{mn}$ and $y_{ij}$ is small. Non-negative gating by boundary signals is described with

$$V_{mnij} = 1 - g\left( \sum_p \nu_{m+p,a,i+p,j} - \omega_3 \right)$$

(A23)

for horizontally displaced neighboring cells and by

$$V_{mnij} = 1 - g\left( \sum_q \nu_{m,n+q,i,j+q} - \omega_3 \right)$$

(A24)

for vertically displaced neighboring cells. The sum is taken over the range $P_5 < p, q < P_5$. Therefore, excitation is prevented from spreading even if the boundary signal is not present at the location $(mnij)$ but it is present in its vicinity. In this way, we circumvent the problem of a weak boundary response at the surface corners which may induce leakage of excitatory activity between surfaces (Grossberg & Todorovic, 1988). Threshold, $\omega_3$, controls how much evidence must accumulate in order to signal the presence of the boundary. Multiplicative interaction between outputs of the neighboring boundary cells from the ventral stream is used,

$$v_{mnij} = g(T_{2mn} - \omega_4) \times g(T_{2ij} - \omega_4),$$

(A25)

in order to provide sharp distinction between different surfaces in the surface network. A similar mechanism for achieving sharpness of boundary signals has been proposed by Neumann, Pessoa, and Hansen (1999). Multiplication is assumed to occur at dendritic branches (Häußser & Mel, 2003; London & Häußser, 2005).
Differential equations are numerically solved using the Euler method with step size = 0.0005. The parameters were set to: $A_2 = 1; B_2 = 300; D_2 = 0; w_{ij} = 1$ for all $i$ and $j; P_5 = 2; \omega_2 = 0.1; \omega_3 = 2; \omega_4 = 0.05$. In Figure 5B, one model neuron, coding the background surface, was set to $y_{ij} = 300$ simulating the endogenous cue. In Figure 5C, self-excitation was set to $w_{ij} = 1$ for neurons coding the figure simulating the habituation of the neurotransmitter release. After the switch between the surfaces, self-excitation was restored to the initial value $w_{ij} = 1$. In all simulations, the network dimensions were set to: $M = N = 50$ in all network layers. In order to avoid spurious edge effects at the network ends, we extended the network dimensions for the size of the respective receptive field in all directions.

### Acknowledgments

We would like to thank two anonymous reviewers for their helpful comments which significantly improved the manuscript and Sanja Pehnec for help during the preparation of the manuscript. This work was supported by the Bial Foundation grant 80/06 and by the Croatian Ministry of Science, Education and Sport grant 009-0362214-0818.

Commercial relationships: none.

Corresponding author: Dražen Domijan.

Email: ddomijan@ffri.hr.

Address: Department of Psychology, Faculty of Arts and Sciences, University of Rijeka, Ivana Klobučarića 1, 51000 Rijeka, Croatia.

### References


Qiu, F. T., & von der Heydt, R. (2005). Figure and ground in the visual cortex: V2 combines stereoscopic cues with gestalt rules. Neuron, 47, 155–166. [PubMed] [Article]


Roelfsema, P. R., Lamme, V. A., Spekreijse, H., & Bosch, H. (2002). Figure-ground segregation in a recurrent network architecture. Journal of Cognitive Neuroscience, 14, 525–537. [PubMed]


