

Boundary assignment in a recurrent network architecture

Janneke F.M. Jehee^{a,b,*}, Victor A.F. Lamme^{a,c}, Pieter R. Roelfsema^{c,d}

^a Department of Psychology, University of Amsterdam, Amsterdam, The Netherlands

^b Center for Visual Science, Department of Computer Science, University of Rochester, 242 Meliora Hall, Rochester, NY 14627-0270, USA

^c Netherlands Institute for Neuroscience, Amsterdam, The Netherlands

^d Department of Experimental Neurophysiology, Center for Neurogenomics and Cognitive Research, Free University, Amsterdam, The Netherlands

Received 4 April 2006; received in revised form 8 December 2006

Abstract

We describe a model and simulations of boundary assignment by cortical neurons, a process that assigns edges to figural image regions, as opposed to the background regions on the other side of the edge. The model is composed of several areas, resembling the hierarchical feedforward–feedback organization of areas in the visual cortex. In each successive area along the hierarchy, the visual image is represented at a coarser resolution. Model neurons tend to assign edges to convex image regions. Because of high spatial resolution, information about convexity is not immediately available to all neurons in lower-level areas. In higher-level areas, however, spatial resolution is low, and convexity is coded more reliably. Feedback connections propagate this information to the high-resolution neurons of lower-level visual areas, making it available at all network levels and at all spatial resolutions. The proposed connection scheme assigns edges faster and more reliable to objects than one with only horizontal connections. The model accounts for both psychophysical and neurophysiological data on figural assignment.

© 2007 Elsevier Ltd. All rights reserved.

Keywords: Computational; Figure–ground; Feedback; Feedforward; Edge assignment

1. Introduction

The brain segregates visual scenes into objects and background, as a first processing step towards visual perception. Essential in this process is the detection of boundaries and their grouping with the adjoining region. The Gestalt psychologists were among the first to study these processes. They pointed out that visual perception tends to assign dividing edges to objects, not to the background (Koffka, 1935; see Baylis & Driver, 1995a, 1995b; Driver & Baylis, 1995; for later studies of one-sided edge assignment). Fig. 1a demonstrates this principle, where the edge is perceived either as the contour of a bottle or as the contour of one of the two glasses.

Which of the two regions along the edge is perceived as figure (and hence ‘owns’ the dividing edge) depends on many factors. It has been suggested that convexity along the edge could be used, assigning edges to the more convex parts (Hoffman & Richards, 1984). This proposal is closely related to the older Gestalt notion of closure, in which the grouping of regions occurs in a way that favors the more closed shapes. Other factors include contrast, because regions that contrast strongly with the general illumination are biased to become figure, and size, because small shapes tend to become figure (Koffka, 1935).

A satisfactory computational theory of how the visual cortex assigns edges to objects is still lacking, but recent studies have uncovered neurophysiological correlates of edge assignment in the visual cortex. The visual cortex is hierarchically organized (Felleman & van Essen, 1991). Neurons in early visual areas have small receptive fields and are tuned to relatively simple features. They represent the various edges and surface elements in the image with

* Corresponding author. Address: Center for Visual Science, Department of Computer Science, University of Rochester, 242 Meliora Hall, Rochester, NY 14627-0270, USA. Fax: +1 585 271 3043.

E-mail address: jjeh@cvsc.rochester.edu (J.F.M. Jehee).

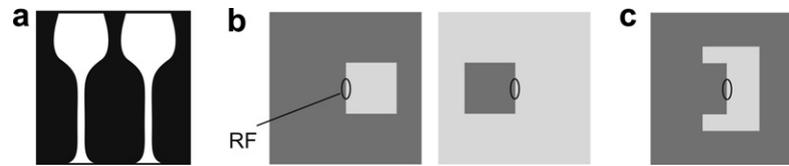


Fig. 1. Examples of one-sided edge assignment. (a) Bottle and glasses. This ambiguous figure demonstrates the tendency of the visual system to assign a dividing edge to one of its sides only. Perception switches back and forth, and the boundaries are assigned to either the bottles, or the glasses. (b) Schematics of the stimuli used in the Zhou et al., 2000 study. The circle denotes the classical receptive field of the recorded neuron. If the neuron is tuned to the left side of a figure, its response will be stronger to the elements that form the left side of the light gray square than to identical elements that form the right side of the dark gray square. (c) Concave corners provide local evidence for the incorrect assignment of boundaries.

high spatial detail. Neurons at higher hierarchical levels have larger receptive fields and are tuned to more complex features. Neuronal correlates of edge assignment have been observed in lower-level as well as in higher-level areas of the visual cortex. Zhou, Friedman, and von der Heydt (2000), for example, studied neurons in areas V1, V2 and V4, which are at a relatively low hierarchical level of the visual cortex, while a contrast-defined edge was presented in the neurons' receptive field. The edge belonged to one or the other surface (Fig. 1b). Many neurons encoded which side was figure, even when the information that determined the figural side was located far from their receptive field. This type of behavior was found in all three lower-level cortical areas. Thus, these neurons coded how boundaries are assigned to figural regions. Neuronal correlates of edge assignment have also been observed in higher visual areas. Baylis and Driver (2001) investigated the neuronal responses in the inferotemporal cortex of monkeys, while the animals saw a contour that either belonged to one or the other bounding surface. Also in this study, the neuronal responses were strongly modified by changes in edge-assignment, much more than, for example, changes in the contrast of the two adjoining surfaces.

The cortical mechanisms underlying this selectivity to edge assignment are currently not well-understood. The present study explores a computational model of the interactions between neurons in lower and higher areas of the visual cortex to shed light on edge assignment. Cortico-cortical connections run from lower visual areas to higher ones (feedforward connections) and back (feedback connections) as well as within areas (horizontal connections), and each of these connections, or a combination, could mediate neuronal tuning to edge assignment. A first possibility is that the tuning to the side of the figure depends on horizontal connections between neurons in the same visual area. However, the aforementioned study of Zhou et al. (2000) provided some evidence against the hypothesis that horizontal connections are the only connections involved in edge assignment. Modulation of the neuronal responses by the edge assignment process occurred very early, within 10–25 ms after the onset of the visual response. The image manipulations used in that study that determined which side was perceived as figure were relatively far from the receptive field (Fig. 1b). Horizontal connections would therefore have to mediate interactions between neurons

separated by relatively large cortical distances. This seems inconsistent with the early occurrence of the response modulation because horizontal connections have a low conduction velocity, i.e. a median of 0.1–0.33 m/s (Bringuier, Chavane, Glaeser, & Fregnac, 1999; Grinvald, Lieke, Frostig, & Hildesheim, 1994). As a result, using lower-level horizontal connections only, it would take tens of milliseconds to reach neurons located only 1° away in the center of the visual field (Bullier, 2001). In contrast, cortico-cortical feedback connections have a median conduction velocity of 2 m/s (Girard, Hupé, & Bullier, 2001) and can in that respect better account for these early response modulations. A further issue that arises in models that only use horizontal connections is that local cues can be inconsistent with the global layout of the figure components (local concavities, see Fig. 1c). The integration of many local and partly inconsistent cues into a consistent global interpretation using only horizontal connections is hard. Indeed, many computational neural network models of figure-ground organization that only extract local cues experience difficulties in the assignment of larger borders and rely on simulated annealing or other iterative approaches to resolve local conflicts in the assignment of boundaries at different locations in the image (e.g., Vecera & O'Reilly, 1998). Such approaches typically take hundreds of processing cycles before completion, which makes them implausible as models for the much faster visual brain.

Hierarchical networks that are composed of several areas (like the visual cortex) have a number of advantages for edge assignment. First, lower visual areas have high magnification factors (Dow, Snyder, Vautin, & Bauer, 1981; Gattas, Gross, & Sandell, 1981; Rosa, 1997). Thus, lower-level neurons represent the image with great spatial detail, but cells with receptive fields on different parts of an elongated edge can be separated by many synapses. Higher areas have lower magnification factors (Rosa, 1997), and neurons can be directly connected even if they represent image locations that are farther apart. Edge assignment can therefore proceed at a faster pace in these areas. Feedback connections then propagate edge assignment signals to lower-level visual areas, where such feedback effects can occur at a relatively early moment in time (Angelucci & Bullier, 2003; Girard et al., 2001; Hupé, James, Girard, Payne, & Bullier, 2001). Second, local concavities pose a problem in early visual areas by producing

signals supporting the erroneous assignment of a boundary. Neurons in higher areas with their larger receptive fields can interpret these local concavities correctly. They can then feed this information back to lower areas to resolve the conflicts. The interactions between lower and higher levels can eventually result in an integrated representation of different levels of spatial detail, thereby enabling correct figural assignments at all network levels and at all levels of spatial detail.

It is our aim to propose a hierarchical neural network model that uses different spatial scales and feedforward and feedback connections for their integration. Such interactions between lower and higher visual areas are in accordance with anatomy. Moreover, feedforward–feedback interactions have been observed in neurophysiology (Hupé et al., 1998; Lamme, 1995; Roelfsema, Lamme, & Spekreijse, 1998) and have been associated with the integration and processing of different levels of spatial detail (Hochstein & Ahissar, 2002; Lee, Mumford, Romero, & Lamme, 1998; Roelfsema, Lamme, & Spekreijse, 2000). The model is composed of several areas, resembling the hierarchical feedforward–feedback organization of areas in the visual cortex. In each successive area along the hierarchy, the visual image is represented at a coarser resolution. Each area is subdivided into contour extraction layers and boundary assignment layers. The boundary assignment layers tend to assign boundaries to convex image regions. Because of high spatial resolution, information about convexity is not immediately available to all boundary assignment neurons in lower areas. In higher areas, however, spatial resolution is low, and convexity is coded more reliably. Feedback connections propagate this information to the high resolution neurons of lower visual areas. The model reproduces neurophysiological results obtained in monkey V1, V2 and V4. Moreover, the model parallels findings in psychophysics by assigning edges to objects, not to the background (Baylis & Driver, 1995a, 1995b; Driver & Baylis, 1995).

2. Model

2.1. Architecture

The model (Fig. 2) is composed of five areas, corresponding to areas V1 and V2 and the ventral stream areas V4, TEO and TE. Each area is subdivided into contour extraction layers and boundary assignment layers. Units in the contour extraction layers are selective to contours but the neuronal mechanisms leading to orientation selectivity are not modeled explicitly (see Olshausen & Field, 1996; Rao & Ballard, 1999; Somers, Nelson, & Sur, 1995; for models of orientation selectivity). Also in the Zhou et al. study (2000), a fraction of the neurons was selective to oriented contours without signaling border ownership. The boundary assignment layers use four selectivities, referred to as ‘left boundary’, ‘right boundary’, ‘top boundary’, and ‘bottom boundary.’ Contour extraction units and

boundary assignment units are implemented in separate layers here but are in reality presumably intermingled. Contour extraction layers propagate activity to higher contour extraction layers via feedforward connections. Boundary assignment layers propagate activity in the reverse direction. In each area, units in the boundary assignment layers receive input from the contour extraction layers via interlaminar connections. We emphasize, however, that feedforward connections between boundary assignment layers, feedback connections between contour extraction layers, and/ or horizontal connections within the areas can be added without changing the main results; we merely choose to implement the model using a minimum of connections here (thereby emphasizing the connections that matter in this task). Area V1 contains 64×64 units for each receptive field selectivity (thus, there are a total of $4 \times 64 \times 64$ boundary assignment units and an additional 64×64 contour extraction units in area V1). At each higher-level in the model, the number of units decreases (by a factor four, Burt & Adelson, 1983) and the size of the receptive fields increases. The image is thus represented at a coarser resolution in each successive area of the hierarchy. The receptive fields of neighboring contour extraction neurons overlap by 30% in all areas higher than V1. Receptive fields of neighboring boundary assignment units overlap by 60% in all areas of the model.

2.2. Connections and weights

Each unit in layer L of the contour extraction pathway receives input from nine contour extraction units in the $L - 1$ layer (connection type 1 in Fig. 2 with weight w_1). The strength of these connections depends on the distance (measured in units) between the receptive field centers according to a Gaussian distribution. In model area V1, contour extraction units are stimulated by a single pixel in the input.

Boundary assignment units in each layer assign boundaries according to the assumption that figures tend to be convex. Thus, each boundary assignment unit in layer L is excited by contour extraction units in layer L that respond to the ‘inner area’ of local L-junctions and inhibited by contour extraction units in layer L that respond to the ‘outer area’ of local L-junctions. For example, a left boundary unit receives excitatory input from contour units on its right (6 units of each contour selectivity; connection type 2 with weight w_2) and inhibitory input from contour units on its left (3 units of each contour selectivity; connection type 3 with weight w_3) (Fig. 2c). The strength of these connections depends on the distance between the receptive field centers according to a Gaussian distribution (asymmetrically truncated to six or three connections). Each boundary unit furthermore receives excitatory feedback from units with similar boundary selectivity in the next higher area (connection type 4 with weight w_4) and inhibitory feedback from units with the opposite boundary selectivity in the next higher area (‘left’ from ‘right’, ‘top’ from

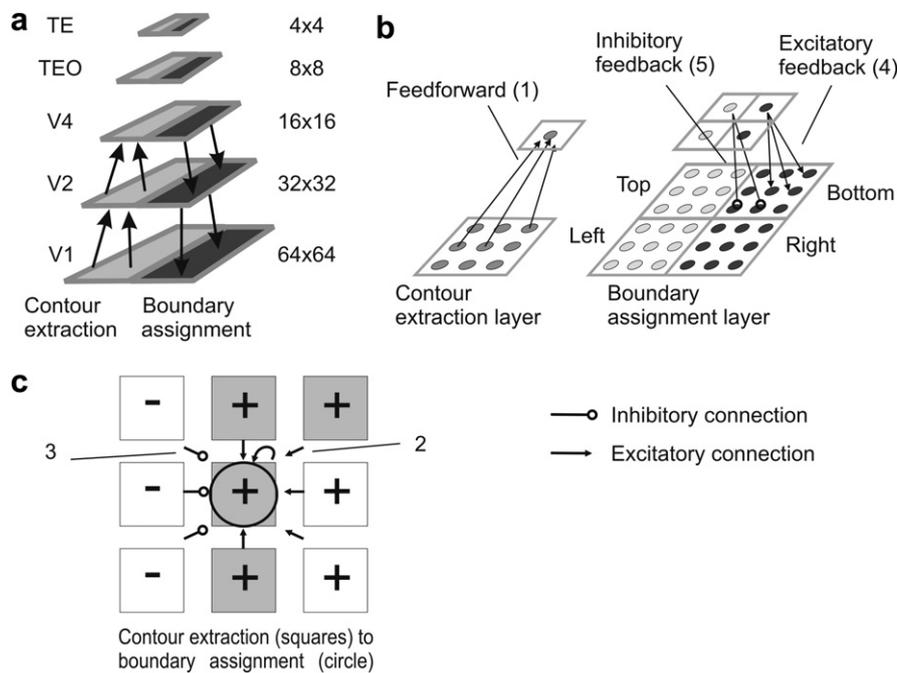


Fig. 2. Architecture of the connections of the model. (a) The model is composed of five areas. At each higher level in the model, the number of units decreases (by a factor four, Burt and Adelson, 1983), and the size of the receptive fields increases, thus the image is represented at coarser resolution. Each area is subdivided into contour extraction layers (light gray) and boundary assignment layers (dark gray). These are shown as separate layers here but would, in reality, be intermingled. (b) Connections between areas. Units in the contour extraction layers are selective to contours. The boundary assignment layers use four selectivities, referred to as 'left boundary', 'right boundary', 'top boundary', and 'bottom boundary.' Units of the contour extraction layers excite units at the next higher level (connection type 1). Units of the boundary assignment layers excite units with similar boundary selectivity in the next lower area (connection type 4) and inhibit units with the opposite boundary selectivity in the next lower area (connection type 5). To improve readability of the figure, only several connections are shown. Receptive fields of neighboring contour extraction neurons overlap by 30% in all model areas higher than area V1. (c) Connections from contour extraction layers to boundary assignment layers. The circle in the middle denotes a neuron that responds to the left boundaries of a figure, superimposed are contour extraction neurons (squares) from which the boundary assignment neuron receives input. Gray denotes the contour extraction units that are excited by (part of) a contour. Boundary assignment units in each layer assign boundaries according to the assumption that figures tend to be convex. Thus, each boundary assignment unit is excited by contour extraction units that respond to the 'inner area' of local L-junctions (connection type 2) and inhibited by contour extraction units in layer L that respond to the 'outer area' of local L-junctions (connection type 3). For example, this left boundary unit receives excitatory input from contour units on its right and inhibitory input from contour units on its left. Intra-areal connections in the other three boundary assignment layers are rotation symmetrical. Arrow, excitatory connection; circle, inhibitory connection.

'bottom', and vice versa; connection type 5 with weight w_5) (see Chey, Grossberg, & Mingolla, 1997; Finkel & Edelman, 1989; Hahnloser, Douglas, Mahowald, & Hepp, 1999; Roelfsema, Lamme, Spekreijse, & Bosch, 2002; for similar connection schemes). These connections are motivated by experimental studies showing that modulatory effects resulting from feedback interactions can be enhancing as well as suppressive (Bair, Cavanaugh, & Movshon, 2003; Hupé et al., 1998, 2001; Lamme, Supér, & Spekreijse, 1998; Murray, Kersten, Olshausen, Schrater, & Woods, 2002).

2.3. Activity and updating

The activity of the network units is described by continuous variables, representing the mean activity of a group or pool of functionally similar neurons in physiology. The appendix summarizes the equations. The model was updated synchronously. Each time step of the model was set to 1 millisecond. Because the model did not incorporate

the retina and the LGN, we added 40 milliseconds to the data points in the figures to account for the delays before area V1 (Nowak, Munk, Girard, & Bullier, 1995).

3. Results

3.1. The feedforward pathway assigns corners to convex image regions

To illustrate the implications of the local connection scheme between contour extraction- and boundary-assignment neurons, we will first discuss the model's response when there are only feedforward and horizontal connections. As soon as the first activity evoked by an image reaches model V1, contour extraction units in area V1 will start to fire (Fig. 3a). The boundary assignment neurons will receive this activity through the interlaminar connections. The image will activate boundary assignment neurons tuned to left, right, top, and bottom boundaries. Contour extraction and boundary assignment

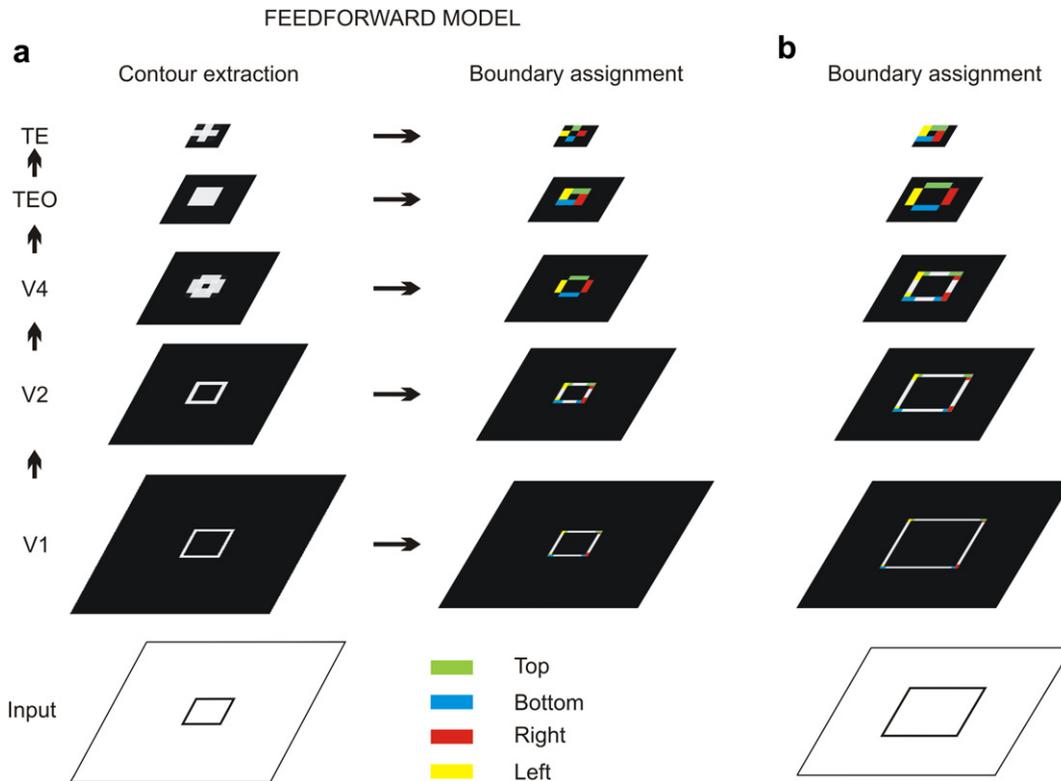


Fig. 3. Activity in the feedforward model, 90 ms after presenting the stimulus. (a, left) Activity in the contour extraction layers. In this and subsequent figures, light shades correspond to regions of high activity, dark regions correspond to regions of low activity. Higher areas represent the image at a coarser resolution. To improve visibility of the figure, areas are scaled differently (i.e. higher areas are depicted larger than they actually are when compared to lower areas). (a, right) Activity in the boundary layers. In this and subsequent figures, green, blue, red, and yellow denote the image regions where the model assigns boundaries unambiguously (i.e. only one boundary unit type is activated) to, respectively, the top, bottom, right and left side of the figure. For other (non-colored) activated neurons, the assignment of boundaries is ambiguous (i.e., left and right unit, or top and bottom unit, are both activated). Higher-level areas instantaneously represent the correct assignment of the boundaries. The network level at which correct edge assignment occurs first depends on the size of the square (compare a and b).

units are shown as separate maps here but are in reality presumably intermingled in a single retinotopic map. Note that in lower model areas, boundaries can only be unambiguously assigned around the corners of the figure. This is further illustrated in the upper panel of Fig. 4, which shows the temporal response profile of V1 boundary assignment units with a receptive field on the middle of a straight contour. The upper and lower horizontal contour of a square both evoke strong responses in a unit that is tuned to boundaries that belong to the lower adjoining region.

Fig. 3 shows that units in higher areas represent the same image at a coarser resolution. Therefore, if the hierarchy is ascended, the contours will ‘blend’ into blobs of elevated activity, until an area is reached where the image elicits activity in a single or a few units. Here, all units are close to a corner, and information about convexity is available to all activated boundary assignment neurons, enabling them to correctly assign all edges to the figure. Note that the network level where correct edge assignment occurs first depends on the size of the square (compare Figs. 3a and b). For larger figures, correct edge assignment occurs at a higher hierarchical level.

3.2. The feedback pathway assigns boundaries to figural regions

Let us now consider what happens when feedback connections between higher-level and lower-level boundary assignment neurons are added to the model. The resolution in higher visual areas is coarser than in lower visual areas, implying that the feedback signal in any area is of lower spatial resolution than the feedforward signal. To make sure that feedback does not influence model neurons that were not excited by the contour extraction signal, excitatory feedback is therefore gated by this signal. More precisely, the model’s feedback connections are multiplicative (see also Fukushima, 1988; Grossberg, 1999; Roelfsema et al., 2002). This ‘gating’ of feedback signals by feedforward activity is supported by physiological data showing that most neurons can only be influenced by contextual stimuli outside their receptive field if the cells are stimulated by another stimulus inside the receptive field (Hupé et al., 2001; Kapadia, Ito, Gilbert, & Westheimer, 1995; Knierim & van Essen, 1992; Sillito, Grieve, Jones, Cudeiro, & Davis, 1995; Zipser, Lamme, & Schiller, 1996). Similar gating of feedback signals has been shown for the attentional modula-

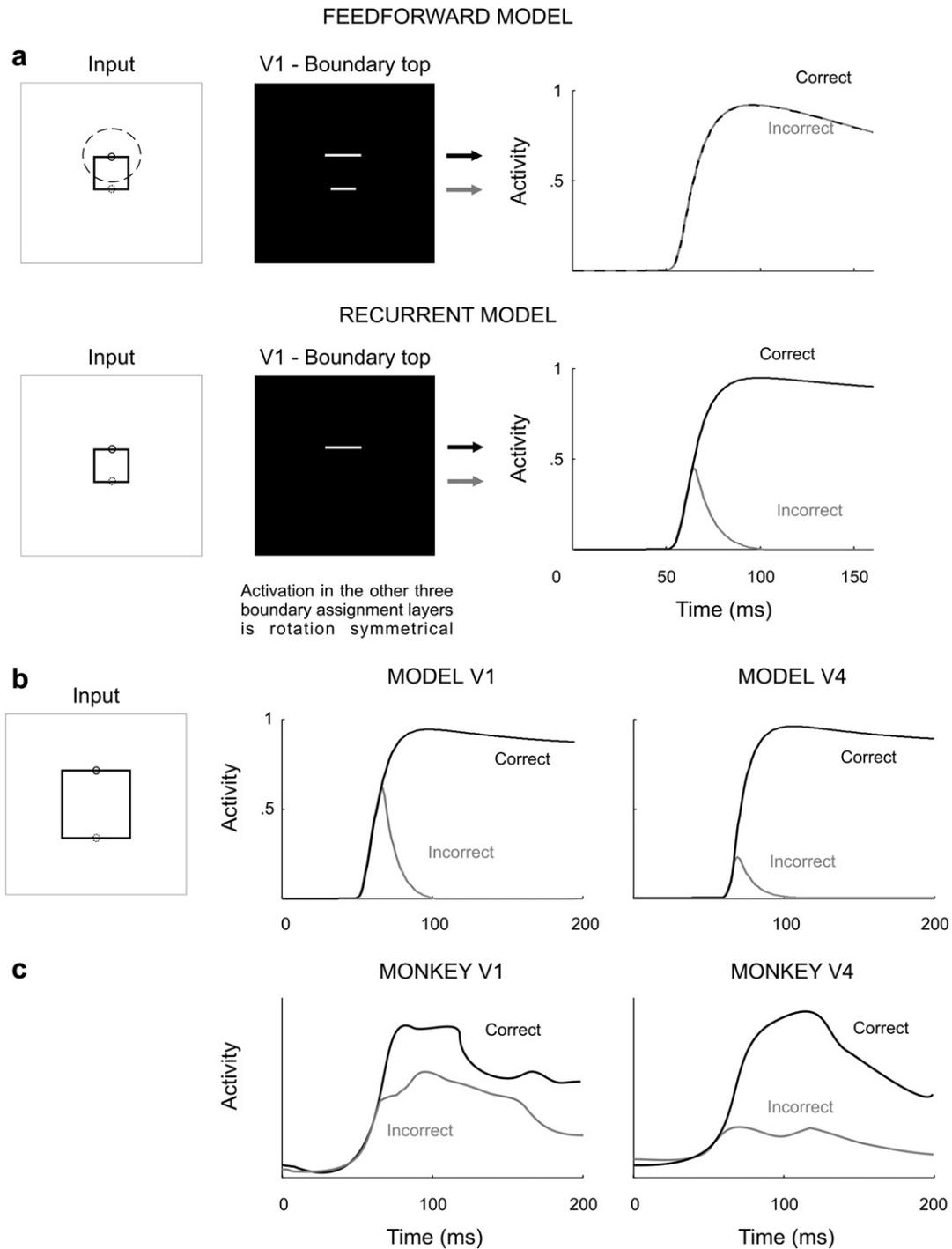


Fig. 4. Activity of model units. (a, top) Purely feedforward model. (Middle) Activity profile of the V1 boundary assignment layers 130 ms after presenting the contour (left). Boundaries are unambiguously assigned to the figure around the corners of the image only. Units activated by the straight boundaries do not distinguish between figure or background. (Right) Temporal activity profile of model V1 boundary assignment neurons whose receptive field cover (part of) the straight boundary. Black, activity profile of a top boundary assignment unit whose receptive field covers the top of the contour (continuous circle in left panel); model TE receptive field size (dashed circle in left panel) is shown for comparison. Gray, activity profile of a top boundary assignment unit whose receptive field covers the bottom of the contour (dotted circle in left panel). The unit's response does not distinguish between the two stimuli. (a, bottom) Fully recurrent model. (Middle) With feedback connections, all of the boundaries are assigned to the figure. (Right) Time-course of the response of a V1 unit tuned to boundaries that belong to the top of a figure evoked by a contour that is at the top (black curve) or bottom (gray curve) of a figure. Note that correct assignment, which depends on feedback from the higher areas, occurs at a delay. (b) Correct assignments are not limited to small figures. (c) Responses of neurons in monkey V1 and V4 after presenting the squares of Fig. 1b. Modified with permission from Zhou et al. (2000).

tion of neuronal responses (McAdams & Maunsell, 1999; Motter, 1994a, 1994b; Treue & Martínez Trujillo, 1999; reviewed in Roelfsema, 2006).

The spatially blurred, but correctly assigned, blob of activity in the highest area feeds back to model neurons in the next lower area that also respond to the contour.

Neurons with a similar boundary preference are excited by this feedback, and neurons with the opposite boundary preference are inhibited (see Section 2). In other words, figural assignments at the highest level enhance figural assignments in the next lower level, while the assignment of contours to the background is prevented. This activity subsequently feeds back to the next lower level and in the end also reaches model area V1. The lower panel of Fig. 4a shows the responses of model V1 neurons to the same square in the fully recurrent network. With feedback connections, all boundaries are correctly assigned to the figure. This is further illustrated by the temporal response profile of the model boundary assignment neurons, which differentiates between background and figural boundaries. Note that correct assignments are not limited to small figures (Fig. 4b).

For comparison, Fig. 4c reproduces data obtained by Zhou et al. (2000) in monkey V1 and V4 after presenting the squares of Fig. 1b. As in the model, responses to figural and background boundaries diverge almost from the beginning in both areas, and their relative difference is larger in V4 than in V1. We also determined the latency of model responses to both figural and background boundaries and that of their difference. The latency of response onset is 53 ms in model V1 and 61 ms in model V4. The latency of the difference is 69 ms in model V1 and 66 milliseconds in model V4. These numbers are comparable to those reported in the Zhou et al. (2000) study, namely 57 ms in monkey V1 and 63 ms in V4 for the response, and 69 ms in V1 and 73 ms in V4 for the difference. We remark, however, that many parameters that influence these latencies are represented by a single time constant in our model (τ_1 in Eqs. (1) and (6) in the appendix). Thus, while the qualitative correspondence between model and neurophysiology is reassuring, it should be realized that model latencies can be easily shifted by changing the time constant or by including more detailed mechanisms into the model.

3.3. Boundary assignment to complex shapes

We next investigated the recurrent model's response to more complex shapes. The activity of the model boundary assignment neurons at early network levels depends on the local shape of an image: convexities support the assignment of boundaries to the figure, while the assignment of straight borders to either figure or background receives equal support. When a convex shape such as a square is presented to the model, figural assignments are thus at least as strongly activated as background assignments, for each location in the visual field and for all hierarchical levels (Fig. 3). The situation is different for shapes that contain local concavities, such as the two shapes in Fig. 5. At local concavities, units that assign boundaries to the background receive the strongest support from their vicinity and are activated more strongly than units that assign boundaries to the figure. This is best seen in the activity profile of

model V1 boundary assignment neurons 60 ms after stimulus presentation, when higher-level areas in the model do not yet provide feedback: the model V1 boundary assignment neurons indeed assign the figures' concave corners to the background (Fig. 5a—boundary assignment after 60 ms: bottom boundary units are activated where top boundary units should have been activated, and so forth).

However, this wrongful assignment occurs in lower visual areas only. In the model's higher visual areas, the shapes are represented at reduced spatial resolution, and local concavities are not represented. The coarser resolution enables the model's higher visual areas to assign all edges correctly to the figure. This correctly assigned edge activity in higher areas reaches the high spatial resolution assignment neurons of the lowest area via the model's feedback connections. This leads the small receptive fields of model area V1 to overcome their inappropriate assignment of borders to the background, allowing the network to represent the contours at maximal spatial resolution (Fig. 5a—boundary assignment after 130 ms).

The resolution of a conflict between higher-level and lower-level model areas is demonstrated in Fig. 5b, which shows the temporal response profile of model V1 neurons after presenting the U-shape of Fig. 5a (the continuous circle denotes the classical receptive field of the recorded model neurons). At first, model neuronal responses that assign the boundary to the background are enhanced, while responses of model neurons that assign the boundary to the figure are inhibited. This wrongful decrease in responses to the figure and wrongful increase in responses to the background is caused by feedback from model areas V2 and V4. In these model areas, neurons that assign edges to the background are activated strongly because their receptive field sizes match the size of the opening of the U-shape and extract it as a candidate object. Feedback from higher areas, however, eventually overrules these wrongful assignments in model V2 and V4. Neuronal activity that assigns the boundary to the figure then becomes dominant in model V2 and V4 and propagates further down to the lowest levels of the model, making correct assignments available at all network levels and at all spatial resolutions.

The first three plots of Fig. 5c show responses of neurons in model areas V1, V2 and V4 after presentation of the U-shape. For comparison, the two leftmost plots of Fig. 5d show responses of neurons in monkey V2 and V4 to a U-shape (Fig. 1c). As in monkey V2 and V4, responses of model neurons that correctly assign edges to the figure are enhanced relative to responses of model neurons that assign edges to the background. The rightmost plot of Fig. 5c contrasts responses of model V2 boundary assignment neurons when presented with a square with responses of the same model neurons when presented with a U-shape. Assignment of edges to a square causes no conflicts between lower-level and higher-level model neurons, and responses are therefore higher than those to the conflict-causing U-shape. This parallels results obtained in monkey V2 (Fig. 5d, rightmost plot), where the neuronal response

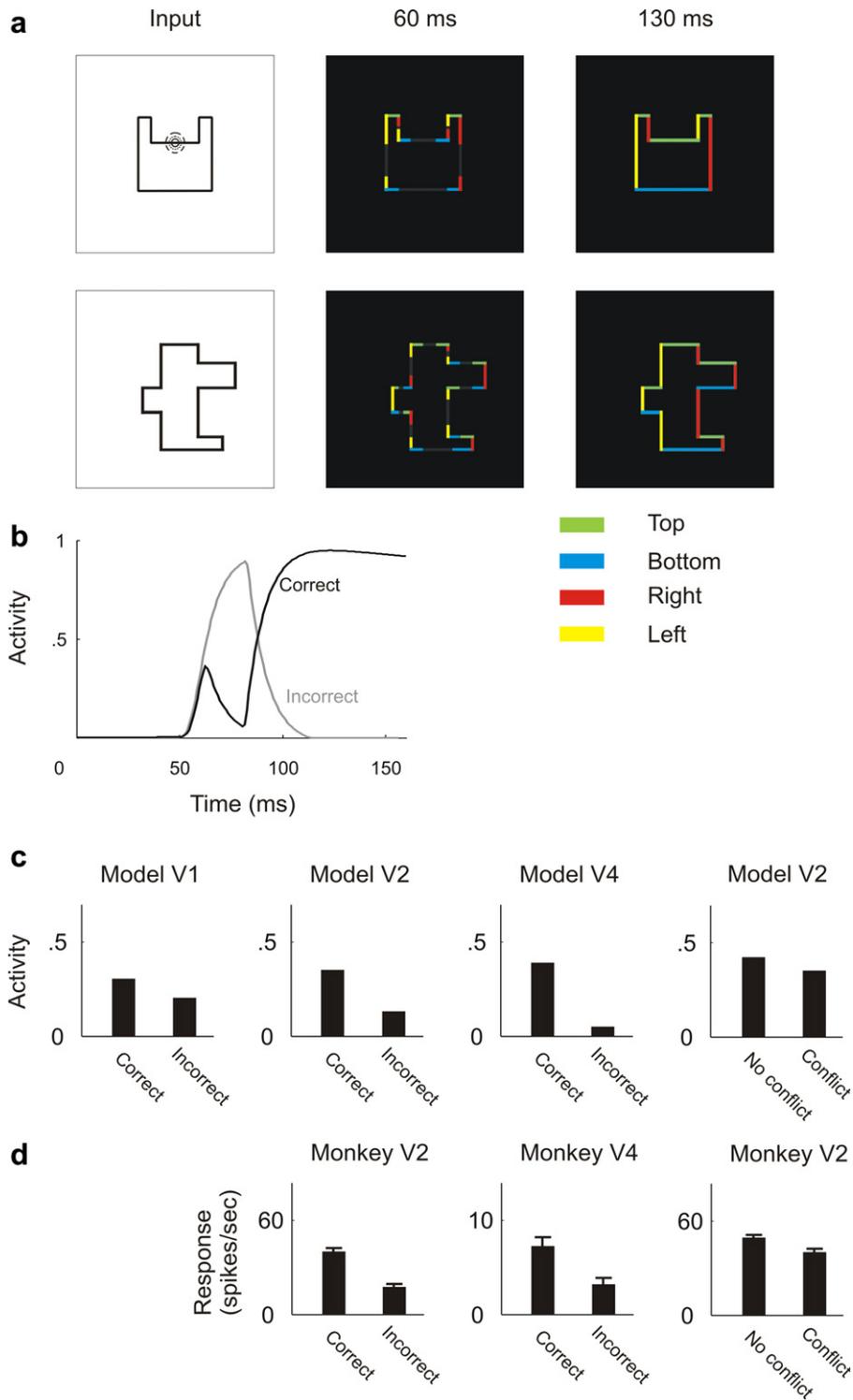


Fig. 5. (a) Activity in the boundary assignment layers of area V1 evoked by more complex shapes (left). Circles denote classical receptive fields. (Middle) Fully recurrent model after 60 ms. Boundaries are correctly assigned to the figure around convex corners but wrongfully to the background at concavities. (Right) Fully recurrent model after 130 ms. Feedback from higher areas eventually overcomes the wrongful assignment of borders to the background by neurons in model area V1 with small receptive fields. (b) Temporal response profile of model V1 neurons whose receptive fields are denoted by the continuous circle in (a). (c) The first three plots show responses of cells in model areas V1, V2 and V4 whose receptive fields are denoted by, respectively, the continuous, dotted and dashed circle in (a). The rightmost plot contrasts responses of model V2 boundary assignment neurons to a square ('no conflict') with responses of the same neurons to the U-shape ('conflict'). Responses were averaged over 125 ms. (d) The two leftmost plots depict responses of monkey V2 and V4 neurons to a U-shape (Fig. 1c). The rightmost plot contrasts responses of monkey V2 neurons to a square (Fig. 1b) with those to a U-shape. Modified with permission from Zhou et al. (2000; the rightmost plot is a combination of Fig. 23 1A and Fig. 23 3B, p. 6607).

to a square (Fig. 1b) is higher than that of the same neuron to a U-shape (Fig. 1c) (Zhou et al., 2000, compare Fig. 23 1A with 23 3B).

4. Discussion

The current model illustrates a general architecture for assignment of boundaries to a figural region. Boundary assignment was computed using local information in areas of different spatial resolution and communicated by feed-forward and feedback connections. This architecture promotes the assignment of boundaries to the figure at all hierarchical levels. Neuronal activity in the model mimicked neurophysiological results obtained in monkey V1, V2 and V4 (Zhou et al., 2000). Moreover, the model paralleled findings in psychophysics by assigning edges to objects, not to the background (Baylis & Driver, 1995a, 1995b; Driver & Baylis, 1995).

The present results can thereby account for the Gestalt rule of convexity (Koffka, 1935), which states that convex image regions tend to be perceived as figures, whereas concave regions tend to be perceived as background (see also Kanizsa & Gerbino, 1976). Selectivity for convexity emerges in the model as the result of a local connectivity scheme (Fig. 2c). This is in accordance with neurophysiological results, as neurons in area V2 and V4 of the monkey have been shown to be sensitive to the convexity of contours (Hedg e & van Essen, 2001; Pasupathy & Connor, 1999).

Our model uses only convexity to determine the figural regions. We suggest, however, that other cues that can bias image regions to be perceived as figure or background, like disparity and luminance, could be incorporated in our model without fundamentally changing the nature of the algorithm. Qui and von der Heydt (2005) indeed demonstrated that the responses of many V2 boundary assignment neurons are relatively independent of the cue that differentiates between figure and background.

Previous modeling accounts of the boundary assignment process (Heitger, von der Heydt, Peterhans, Rosenthaler, & K ubler, 1998; Peterhans & Heitger, 2001; Vecera & O'Reilly, 1998; Zhaoping, 2005) used a single hierarchical level for the computation of figural borders. In these models, image regions are grouped via horizontal connections, resulting in longer processing times (e.g., Vecera & O'Reilly, 1998) or relatively complex connectivity schemes (e.g., Zhaoping, 2005). Although we consider a detailed comparison between Zhaoping's model (2005) and ours as a topic of future research, we predict that our model gives comparable results when the size of the visual object is varied. The model of Zhaoping (2005) implements only a single spatial scale and is therefore sensitive to changes in the size of the objects that are presented. Additionally, Zhaoping's model (2005) uses many different connections to achieve boundary assignment; for instance, many different kinds of inhibitory connections between neurons signaling boundaries that do not belong to the same figure and many different

kinds of excitatory connections between neurons that signal boundaries of the same figure. In contrast, the model presented here uses a very simple mechanism to compute border ownership (Fig. 2c), repeated across all hierarchical levels. Grossberg (1994, 1997) and Kelly and Grossberg (2000) proposed a model of boundary detection and subsequent 'filling in' (detection) of the figural region by surface detectors, segregating boundaries and figural regions from the background. Their model, however, does not include explicit border ownership neurons. As such, model neurons encode whether the boundary belongs to the figure, but not whether it forms, for instance, the left side of the figure. In contrast, the present approach uses neurons that are selective for a specific side of a figure and, hence, assign boundaries exclusively to one figural side, in accordance with neurophysiology (Qui & von der Heydt, 2005; Zhou et al., 2000).

We used only relatively simple stimuli in our simulations, as opposed to, for example, stimuli with T-junctions and X-junctions, and images with partial occlusions. To robustly handle such more complex stimuli our model would have to be adjusted by incorporating mechanisms for the detection of colinearity (to group well aligned contours that are occluded by another surface) and by including units tuned to various corner configurations, to depth, and depth ordering, as well as many new types of connections. We deliberately did not include these details that would make our model much more complex and that would distract from the essence of the idea proposed here: border ownership can be computed easily in a hierarchy of areas with different spatial resolutions. The model presented here therefore uses only a small set of parameters, i.e. only one set of synaptic weights, time constants, etc., replicated across all hierarchical levels. Despite its minimal complexity, the model captures neurophysiological results obtained in monkey V1, V2 and V4 (Zhou et al., 2000).

In the model, the detection of convexity occurs at multiple network levels and therefore at multiple spatial scales. In the model, correct assignment of boundaries occurs first at a level where figural size best matches the size of the receptive fields. Correct assignment of boundaries might fail if the figure is larger than the receptive fields at the highest level. This problem can be circumvented by including areas in which receptive fields are still larger, but we note that the unambiguous assignment of figure and ground may become impossible if figure size exceeds the entire field of view. Whenever the figure size matches the receptive field size at some hierarchical level, feedback connections propagate the figural assignments to lower areas with finer resolutions. This connection scheme is much faster than one in which only horizontal connections are used (e.g., Vecera & O'Reilly, 1998). In schemes with only horizontal connections, conflicts that arise from local concavities of a figure can only be resolved by slow relaxation schemes, whereas the feedback signal immediately overrules the incorrect boundary assignments. Moreover, pyra-

midal cortico-cortical feedback connections have a median conduction velocity of around 2 m/s and are much faster than horizontal connections that have a median conduction velocity of about 0.1–0.33 m/s (Bringuier et al., 1999; Girard et al., 2001; Grinvald et al., 1994). The proposed framework allows for very rapid processing at all network levels: higher-level areas directly assign edges during the first feedforward sweep of activity, and feedback mechanisms rapidly communicate these to lower-level areas. Indeed, neurophysiological studies have revealed very short latencies of boundary assignment effects in both higher level (Baylis & Driver, 2001) and lower-level areas (Zhou et al., 2000).

Peterson, Harvey, and Weidenbacher (1991) demonstrated that familiar regions are more likely to be perceived as figure than less familiar regions (Peterson et al., 1991), suggesting that cortical boundary assignment processes are affected by familiarity. Our model can account for these effects if shape selective neurons are added to the highest levels of the model. Such neurons have been found in inferotemporal cortex (Tanaka, 1995). Moreover, previous studies have shown that a neural network model can account for familiarity effects if shape selective units excite lower-level cells coding for features consistent with that shape (Vecera & O'Reilly, 1998; Fukushima, 1988). Such a connectivity scheme is compatible with the one proposed here.

The present work illustrates the usefulness of multi-scale analyses in cortical processing (see also Koenderink, 1984). Receptive fields in lower visual areas have the best spatial resolution but are too small to directly assign an edge with certainty. Higher visual areas present the world at a coarser resolution, reducing the chance of false assignments. The model proposed here therefore assigns edges at coarser resolution first and uses these higher-level computations for its edge assignments at finer scales. It suggests that it can be easier to recognize a figure when some of its details are lost. This has also been found for classification, where visual features of coarser resolution do better than detailed ones (Ullman, Vidal-Naquet, & Sali, 2002). Koenderink (1984) showed how it is possible to bring image features that are represented at a low spatial resolution into correspondence with the same features at a higher resolution. Interactions between units that represent the same image at various resolutions have also been exploited in a model for texture-segregation (Roelfsema et al., 2002), and similar ideas have been applied in computer vision (Ballard, 1981). These models used simple rules for their feedforward, lateral and feedback connections, suggesting that these techniques could indeed be implemented by cortical neurons.

Consistent with recent theories, the model shows how feedback connections can provide input to lower visual areas in situations where it is necessary to process spatially detailed information (Hochstein & Ahissar, 2002; Lee et al., 1998; Roelfsema et al., 2000). In these theories, feedforward processing involves rapid and automatic processes that provide basic object categorizations, but give little spa-

tial detail. A detailed and complete representation could be found in lower visual areas using feedback mechanisms. Lower-level receptive fields are smaller than higher-level receptive fields and are in that respect more suitable for the coding of spatial detail. Our model is in accordance with these theories, as local boundary assignments become available only after subsequent feedforward–feedback passes between different levels of spatial resolution. In the model, using only coarse, higher-level computations would increase the probability that regions with different boundary ownership values are merged in the same regions. The small receptive fields of area V1 enable the network to represent the image at maximal spatial resolution, without smoothing over boundary ownership values.

In conclusion, the present model illustrates a general scheme by which cortical boundary assignment could be implemented. Following the visual brain's organization, the model uses a hierarchy of different spatial resolutions and recurrent interactions for their integration. The model accounts for both psychophysical and neurophysiological data on figural assignment and suggests a computational role for feedforward–feedback interactions in processing of spatially detailed information.

Acknowledgments

Part of this work was supported by a HFSP Young Investigators grant awarded to PRR, a grant of NWO-Exact, and an EU FP-6 STREP Grant (No. 027198).

Appendix A

Activity in the contour extraction pathway is updated according to the following equations:

$$\tau_1 \frac{d}{dt} C_{L,i} = -C_{L,i} + f_{15}^{0.15}(\text{Inp}C_{L,i}) - 0.25A_{L,i}^C \quad (1)$$

$$\tau_2 \frac{d}{dt} A_{L,i}^C = -A_{L,i}^C + C_{L,i} \quad (2)$$

where $C_{L,i}$ stands for the activity of unit i in the contour extraction layer of area L , $\text{Inp}C_{L,i}$ reflects input into unit i , and f is a squashing function:

$$f(x)_s^{\theta_C} = 0.5(1 + \tanh(s(x - \theta_C))) \quad (3)$$

where $\theta_C = 0.15$ is the threshold in the contour detection layers and $s = 15$ is the slope of the squashing function. The model is not very sensitive to the value of s nor to small deviations from θ_C ; substantially higher values for θ_C prevent higher-level contour extraction layers in the model from being activated. The slope of the initial response transient is determined by τ_1 , a time constant that was set to 10 ms. After this transient response, the activity of the neurons is reduced by a local inhibitory process, $A_{L,i}^C$, with time constant $\tau_2 = 100$ ms. This process was included to model the transient responses of visual cortical neurons (e.g., Maunsell & Gibson, 1992). Input $\text{Inp}C_{L,i}$ comes from

the next lower layer (for all contour extraction layers except V1):

$$\text{Inp}C_{L,i} = \sum_{m_1 \in V_1} H(i, m_1, V_1, w_1, \sigma_{V_1}) C_{L-1, m_1} \quad (4)$$

where V_1 is a neighborhood of 9 units in the $L - 1$ layer, and H is a Gaussian distribution $g(d(i, m_j), \sigma)$, in which $d(i, m_j)$ is the distance between unit i and neighboring unit m_j , normalized such that the sum of its values equals w_i :

$$H(i, m_j, V_k, w_i, \sigma_{V_k}) = w_i g(d(i, m_j), \sigma_{V_k}) / \sum_{m_j \in V_k} g(d(i, m_j), \sigma_{V_k}) \quad (5)$$

in which $\sigma_{V_1} = 0.85$ interunit distance (chosen such that the total output weight is equal for each unit in layer $L - 1$). Contour extraction units in model area V1 are stimulated by a single input unit with activity a_i , $\text{Inp}C_{V1,i} = a_i$. The activity of the input units was obtained by setting all input units overlaying the contour to 1 and all other input units to 0.

Activity in the boundary assignment layers is updated according to the following equations:

$$\tau_1 \frac{d}{dt} B_{L,i}^B = -B_{L,i}^B + f_{15}^{0.85}(\text{Inp}B_{L,i}^{\text{tot}}) - 0.25A_{L,i}^B \quad (6)$$

$$\tau_2 \frac{d}{dt} A_{L,i}^B = -A_{L,i}^B + B_{L,i}^B \quad (7)$$

$$\text{Inp}B_{L,i}^{\text{tot}} = \text{Inp}B_{L,i}^P (1 + \text{Inp}B_{L,i}^Q) - \text{Inp}B_{L,i}^R \quad (8)$$

where $B_{L,i}^B$ stands for the activity of unit i with boundary selectivity B in the boundary assignment layer of area L . A local inhibitory process $A_{L,i}^B$ was included to model the transient responses of visual cortical neurons. The threshold $\theta_B = 35$ in the squashing function ensures that only boundary units with the same indices as stimulated contour extraction units are activated (vs. also some boundary units around these indices that are also connected to the activated contour assignment units), but different values give similar results. $\text{Inp}B_{L,i}^{\text{tot}}$ in Eq. (6) stands for the total input received by the boundary assignment unit. It can be subdivided into three parts (superscripts P , Q , R ; Eq. (8)), of which the first part ($\text{Inp}B_{L,i}^P$) reflects excitatory input from contour extraction units that respond to the ‘inner area’ of local L-junctions (see Fig. 2c and text for further details):

$$\text{Inp}B_{L,i}^P = \sum_{m_2 \in V_2} H(i, m_2, V_2, w_2, \sigma_{V_2}) C_{L, m_2} \quad (9)$$

where V_2 in $\text{Inp}B_{L,i}^P$ is a neighborhood in the contour extraction layer of area L that contains 6 units.

A further source of input into the boundary assignment unit ($\text{Inp}B_{L,i}^Q$ in Eq. (8)) comes from boundary assignment units in area $L+1$ ($B_{L+1,i}^B$):

$$\text{Inp}B_{L,i}^Q = \sum_{m_4 \in V_4} H(i, m_4, V_4, w_4, \sigma_{V_4}) B_{L+1, m_4}^B \quad (10)$$

where V_4 is a neighborhood of units in the $L + 1$ layer and contains units with similar boundary selectivity. To make sure that input $\text{Inp}B_{L,i}^Q$ does not influence neurons that were not excited by the contour extraction signal, it was multiplied (‘gated’) by signal $\text{Inp}B_{L,i}^P$ (Eq. (8)).

A third source of input into the boundary assignment unit ($\text{Inp}B_{L,i}^R$ in Eq. (8)) is inhibitory and comes from contour extraction units (see Fig. 2c and text for further details) and from boundary assignment units in area $L + 1$ ($B_{L+1,i}^B$):

$$\text{Inp}B_{L,i}^R = \sum_{m_3 \in V_3} H(i, m_3, V_3, w_3, \sigma_{V_3}) C_{L, m_3} + \sum_{m_5 \in V_5} H(i, m_5, V_5, w_5, \sigma_{V_5}) B_{L+1, m_5}^{B'} \quad (11)$$

where V_3 is a neighborhood in the contour extraction layer of area L that contains 3 units that respond to the ‘outer area’ of local L-junctions and V_5 is a neighborhood of units in the $L + 1$ layer that contains units with mirrored boundary selectivity B' (‘left’ receives connections from ‘right’, ‘top’ from ‘bottom’, and vice versa).

Connection weights w_i for connection types i were as follows: $w_1 = 1$, $w_2 = 1.5$, $w_3 = 1.5$, $w_4 = 1$, $w_5 = 1$. $\sigma_{V_1} = \sigma_{V_4} = 0.85$, $\sigma_{V_2} = \sigma_{V_3} = 0.8$, and $\sigma_{V_5} = 2.5$ interunit distance. All parameter values were kept constant throughout all simulations. To avoid ‘boundary effects’ at the bounds of the layers, all values outside the bounds of the layers are computed by assuming that the layer is periodic.

References

- Angelucci, A., & Bullier, J. (2003). Reaching beyond the classical receptive field of V1 neurons: horizontal of feedback axons? *Journal of Physiology (Paris)*, 97, 141–154.
- Bair, W., Cavanaugh, J. R., & Movshon, J. A. (2003). Time-course and time-distance relationships for surround suppression in macaque V1 neurons. *Journal of Neuroscience*, 23, 7690–7701.
- Ballard, D. H. (1981). Strip trees: a hierarchical representation for curves. *Communications of the ACM*, 24, 310–321.
- Baylis, G. C., & Driver, J. (1995a). Obligatory edge assignment in vision: the role of figure and part segmentation in symmetry detection. *Journal of Experimental Psychology: Human Perception and Performance*, 21, 1323–1342.
- Baylis, G. C., & Driver, J. (1995b). One-sided edge-assignment in vision.1. Figure-ground segmentation and attention to objects. *Current Directions in Psychological Science*, 4, 140–146.
- Baylis, G. C., & Driver, J. (2001). Shape-coding in IT cells generalizes over contrast and mirror reversal, but not figure-ground reversal. *Nature Neuroscience*, 4, 937–942.
- Binguier, V., Chavane, F., Glaeser, L., & Fregnac, Y. (1999). Horizontal propagation of visual activity in the synaptic integration field of area 17 neurons. *Science*, 283, 695–699.
- Bullier, J. (2001). Integrated model of visual processing. *Brain Research Reviews*, 36, 96–107.
- Burt, P. J., & Adelson, E. H. (1983). The Laplacian pyramid as a compact image code. *IEEE Transaction on Communications*, 31, 532–540.
- Chey, J., Grossberg, S., & Mingolla, E. (1997). Neural dynamics of motion grouping: from aperture ambiguity to object speed and direction. *Journal of the Optical Society of America A: Optics, Image Science, and Vision*, 14, 2570–2594.

- Dow, B. M., Snyder, A. Z., Vautin, R. G., & Bauer, R. (1981). Magnification factor and receptive field size in foveal striate cortex of the monkey. *Experimental Brain Research*, *44*, 213–228.
- Driver, J., & Baylis, G. C. (1995). One-sided edge-assignment in vision. 2. Part decomposition, shape description, and attention to objects. *Current Directions in Psychological Science*, *4*, 201–206.
- Felleman, D. J., & van Essen, D. C. (1991). Distributed hierarchical processing in the primate cerebral cortex. *Cerebral Cortex*, *1*, 1–47.
- Finkel, L. H., & Edelman, G. M. (1989). Integration of distributed cortical systems by reentry: a computer simulation of interactive functionally segregated visual areas. *Journal of Neuroscience*, *9*, 3188–3208.
- Fukushima, K. (1988). A neural network for visual pattern recognition. *IEEE Computer*, *21*, 65–75.
- Gattas, R., Gross, C. G., & Sandell, J. H. (1981). Visual topography of V2 in the macaque. *Journal of Comparative Neurology*, *2001*, 519–539.
- Girard, P., Hupé, J. M., & Bullier, J. (2001). Feedforward and feedback connections between areas V1 and V2 of the monkey have similar rapid conduction velocities. *Journal of Neurophysiology*, *85*, 1328–1331.
- Grinvald, A., Lieke, E. E., Frostig, R. D., & Hildesheim, R. (1994). Cortical point-spread function and long-range lateral interactions revealed by real-time optical imaging of macaque monkey primary visual cortex. *Journal of Neuroscience*, *14*, 2545–2568.
- Grossberg, S. (1994). Theory and evaluative reviews: 3-D vision and figure-ground separation by visual cortex. *Perception & Psychophysics*, *55*, 48–120.
- Grossberg, S. (1997). Cortical dynamics of three-dimensional figure-ground perception of two-dimensional pictures. *Psychological Review*, *104*, 618–658.
- Grossberg, S. (1999). The link between brain learning, attention, and consciousness. *Consciousness and Cognition*, *8*, 1–44.
- Hahnloser, R., Douglas, R., Mahowald, M., & Hepp, K. (1999). Feedback interactions between neuronal pointers and maps for attentional processing. *Nature Neuroscience*, *2*, 746–752.
- Hedg e, J., & van Essen, D. C. (2001). Selectivity for complex shapes in primate visual area V2. *The Journal of Neuroscience*, *20*, RC61.
- Heitger, F., von der Heydt, R., Peterhans, E., Rosenthaler, L., & K ubler, O. (1998). Simulation of neural contour mechanisms: representing anomalous contours. *Image and Vision Computing*, *16*, 407–421.
- Hochstein, S., & Ahissar, M. (2002). View from the top: hierarchies and reverse hierarchies in the visual system. *Neuron*, *36*, 791–804.
- Hoffman, D. D., & Richards, W. A. (1984). Parts of recognition. *Cognition*, *18*, 65–96.
- Hup e, J. M., James, A. C., Girard, P., Payne, B. R., & Bullier, J. (2001). Feedback connections act on the early part of the responses in monkey visual cortex. *Journal of Neurophysiology*, *85*, 134–145.
- Hup e, J. M., James, A. C., Payne, B. R., Lomber, S. G., Girard, P., & Bullier, J. (1998). Cortical feedback improves discrimination between figure and background by V1, V2 and V3 neurons. *Nature*, *394*, 784–787.
- Kanizsa, G., & Gerbino, W. (1976). Convexity and symmetry in figure ground organization. In M. Henle (Ed.), *Art and artifacts*. New York: Springer.
- Kapadia, M. K., Ito, M., Gilbert, C. D., & Westheimer, G. (1995). Improvement in visual sensitivity by changes in local context: Parallel studies in human observers and in V1 or alert monkeys. *Neuron*, *15*, 843–856.
- Kelly, F., & Grossberg, S. (2000). Neural dynamics of 3-D surface perception: figure-ground separation and lightness perception. *Perception & Psychophysics*, *62*, 1596–1618.
- Knierim, J. J., & van Essen, D. C. (1992). Neuronal responses to static texture patterns in area V1 of the alert macaque monkey. *Journal of Neurophysiology*, *67*, 961–980.
- Koenderink, J. J. (1984). The structure of images. *Biological Cybernetics*, *50*, 363–370.
- Koffka, K. (1935). *Principles of Gestalt psychology*. London: Routledge and Kegan Paul.
- Lamme, V. A. F. (1995). The neurophysiology of figure-ground segregation in primary visual cortex. *Journal of Neuroscience*, *15*, 1605–1615.
- Lamme, V. A. F., Sup er, H., & Spekreijse, H. (1998). Feedforward, horizontal, and feedback processing in the visual cortex. *Current Opinion in Neurobiology*, *8*, 529–535.
- Lee, T. S., Mumford, D., Romero, R., & Lamme, V. A. F. (1998). The role of the primary visual cortex in higher level vision. *Vision Research*, *38*, 2429–2545.
- Maunsell, J. H., & Gibson, J. R. (1992). Visual response latencies in striate cortex of the macaque monkey. *Journal of Neurophysiology*, *68*, 1332–1344.
- McAdams, C. J., & Maunsell, J. H. R. (1999). Effects of attention on orientation-tuning functions of single neurons in macaque area V4. *Journal of Neuroscience*, *19*, 431–441.
- Motter, B. C. (1994a). Neural correlates of attentive selection for color or luminance in extrastriate area V4. *Journal of Neuroscience*, *14*, 2178–2189.
- Motter, B. C. (1994b). Neural correlates of feature selective memory and pop-out in extrastriate area V4. *Journal of Neuroscience*, *14*, 2190–2199.
- Murray, S. O., Kersten, D., Olshausen, B. A., Schrater, P., & Woods, D. L. (2002). Shape perception reduces activity in human primary visual cortex. *Proceedings of the National Academy of Sciences of the United States of America*, *99*, 15164–15169.
- Nowak, L. G., Munk, M. H. J., Girard, P., & Bullier, J. (1995). Visual latencies in areas V1 and V2 of the macaque monkey. *Visual Neuroscience*, *12*, 271–384.
- Olshausen, B., & Field, D. J. (1996). Emergence of simple-cell receptive field properties by learning a sparse code for natural images. *Nature*, *381*, 607–609.
- Pasupathy, A., & Connor, C. E. (1999). Responses to contour features in macaque area V4. *Journal of Neurophysiology*, *82*, 2490–2502.
- Peterhans, E., & Heitger, F. (2001). Simulation of neuronal responses defining depth order and contrast polarity at illusory contours in monkey area V2. *Journal of Computational Neuroscience*, *10*, 195–211.
- Peterson, M. A., Harvey, E. M., & Weidenbacher, H. J. (1991). Shape recognition contributions to figure-ground reversal: which route counts? *Journal of Experimental Psychology, Human Perception and Performance*, *17*, 1075–1089.
- Qui, 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.
- Rao, R. P. N., & Ballard, D. H. (1999). Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. *Nature Neuroscience*, *2*, 79–87.
- Roelfsema, P. R. (2006). Cortical algorithms for perceptual grouping. *Annual Reviews in Neuroscience*, *29*, 203–227.
- Roelfsema, P. R., Lamme, V. A. F., & Spekreijse, H. (1998). Object-based attention in the primary visual cortex of the macaque monkey. *Nature*, *395*, 376–381.
- Roelfsema, P. R., Lamme, V. A. F., & Spekreijse, H. (2000). The implementation of visual routines. *Vision Research*, *40*, 1385–1411.
- Roelfsema, P. R., Lamme, V. A. F., Spekreijse, H., & Bosch, H. (2002). Figure-ground segregation in a recurrent network architecture. *Journal of Cognitive Neuroscience*, *14*, 525–537.
- Rosa, M. (1997). Visuotopic organization of primate extrastriate cortex. In K. S. Rockland, J. H. Kaas, & A. Peters (Eds.), *Extrastriate cortex in primates* (vol. 12, pp. 127–203). New York: Plenum Press.
- Sillito, A. M., Grieve, K. L., Jones, H. E., Cudeiro, J., & Davis, J. (1995). Visual cortical mechanisms detecting focal discontinuities. *Nature*, *378*, 492–496.
- Somers, D. C., Nelson, S. B., & Sur, M. (1995). An emergent model of orientation selectivity in cat visual cortical simple cells. *Journal of Neuroscience*, *15*, 5448–5465.

- Tanaka, K. (1995). Neuronal mechanisms of object recognition. *Science*, 262, 685–688.
- Treue, S., & Martínez Trujillo, J. C. (1999). Feature-based attention influences motion processing gain in macaque visual cortex. *Nature*, 399, 575–579.
- Ullman, S., Vidal-Naquet, M., & Sali, E. (2002). Visual features of intermediate complexity and their use in classification. *Nature Neuroscience*, 5, 682–687.
- Vecera, S. P., & O'Reilly, R. C. (1998). Figure-ground organization and object recognition processes: an interactive account. *Journal of Experimental Psychology: Human Perception and Performance*, 24, 441–462.
- Zhaoping, L. (2005). Border ownership from intracortical interactions in visual area V2. *Neuron*, 47, 143–153.
- Zhou, H., Friedman, H. S., & von der Heydt, R. (2000). Coding of border ownership in monkey visual cortex. *The Journal of Neuroscience*, 20, 6594–6611.
- Zipser, K., Lamme, V. A. F., & Schiller, P. H. (1996). Contextual modulation in primary visual cortex. *Journal of Neuroscience*, 16, 7376–7389.