CHAPTER SIX

NEUROCOMPUTATIONAL MODELS OF FIGURE-GROUND ORGANIZATION

DRAŽEN DOMIJAN[1]

1. Introduction

Any complex system can be analysed at many different levels and visual perception is not an exception. Traditional psychological explorations of human experience when we are exposed to a particular pattern of sensory stimulation are now being augmented with neuroscientific and computational investigations. Neuroscience offers an opportunity to identify brain structures and mechanisms responsible for visual processing. Computational work attempts to cross the boundaries between neural and psychological levels of analysis and to provide a unified explanation for why we experience the visual world as we do. In order to gain detailed insight into the nature of perceptual processes, all three levels should be investigated in collaboration (Palmer, 1999).

The aim of the present review is to describe recent progress in our understanding of neurocomputational mechanisms responsible for figure-ground organization. In recent years, several models have been developed with different assumptions about the processes underlying figural assignment. In order to better elucidate the strengths and weaknesses of a particular approach, it is important to review the relevant empirical findings that can constrain their psychological or neurophysiological plausibility. The present review is divided into three parts reflecting three different levels of analysis. First, a review of relevant psychological phenomena is provided, starting with the seminal contribution by Gestalt psychologists who identified a set of principles that govern figural assignment. Also, new developments are sketched, highlighting the discoveries of novel variables

[1] Associate Professor, Department of Psychology, Faculty of Humanities and Social Sciences, University of Rijeka, Rijeka, Croatia (ddomijan@ffri.hr).

affecting figural status of the surface. Second, neurophysiological evidence is provided that sheds new light on how real cortical networks respond to figure and background.

Finally, neurocomputational models are described. The models are divided into two groups based on their adherence to a biological or psychological level of explanation. Biological models propose that the border ownership response in V2 is a major neural event responsible for figural assignment. These models are further subdivided into two groups based on the assumed source of border ownership signals. One group suggests that border ownership is the consequence of recurrent interaction within V2. Other groups propose that border ownership arises from the interaction between V2 and higher visual areas. Psychological models rely on the process of filling-in that enables construction of surface representation. In these models, interaction between boundary and surface information is considered essential for proper figure-ground segregation. In this part of the chapter, I will analyze the ability of proposed computational models to account for complex patterns of empirical findings. In doing so, I will try to identify important unresolved issues that require further investigation.

2. Psychological (behavioural) perspective

Figure-ground separation is part of perceptual organization which enables parsing a visual image into two distinct components:

1) Figure – a region of interest to which our attention should be directed
2) Background – remaining space that is largely ignored.

Gestalt psychologists were the first to realise the importance of figure-ground separation for visual perception. Rubin's famous face-vase stimulus illustrates basic points that need to be addressed (Rubin, 1915/1958). The figure is segregated from the background by the edge (or border) that defines the shapes that we recognise as either two faces or one vase. The central question is how we assign figural status to a surface from one side of the border and not to the other. At any one point in time, we see either a vase or faces as a figure but never both. Perceptual organization switches easily from one interpretation to another, but the faces and the vase are never perceived simultaneously. This is an example of a bi-stable stimulus. More importantly, we perceive the boundary that divides surfaces as belonging always to the figure and not to the ground.

When the perceptual interpretation alters, the border ownership of the boundary also alters. Therefore, we always perceive a figure as a surface with a definite shape while the background is shapeless or unstructured. Furthermore, a figure appears to be located in front of the background (Rubin, 1915/1958). Gestalt psychologists discovered several principles (figural cues) that govern the process of assigning figural status to surfaces in visual fields (Bahnsen, 1928; Harrower, 1936; Kanizsa & Gerbino, 1976; Koffka, 1935; Metzger, 1953; Rubin, 1915/1958; for a review, see Pomerantz & Kubovy, 1986; and Palmer, 1999). The surface most likely to be perceived as a figure is the one that is:

1) surrounded by other surfaces
2) smaller in size
3) higher in contrast
4) oriented along a horizontal and a vertical axis
5) convex
6) symmetrical
7) composed of parallel boundaries

It should be noted that these principles are descriptive in nature. They do not offer quantitative relations between relevant variables. For instance, it is not immediately clear which principle will dominate when they are directly put against each other. An exception was the work of Kanizsa and Gerbino (1976) that showed that convexity is a powerful cue that overrides symmetry when the two cues are in conflict. Reviewed principles do not exhaust all possible image cues registered by the visual system that can potentially influence figural assignment. As pointed out by Vecera *et al.* (2002), a coffee cup is perceived as a figure on a background of a cluttered desk, although it may be neither symmetric nor convex.

Further work led to discoveries of new principles that govern figure-ground assignment. Klymenko and Weistein (1986) discovered that surfaces with higher spatial frequency content had a tendency to be perceived as a figure. Peterson (1994; Peterson & Gibson, 1994a; 1994b) discovered the role of top-down processes in figure-ground perception. She showed that a familiar surface has a greater chance to be perceived as a figure compared to a less familiar surface. She argued that object recognition contribution to the figure-ground assignment derives from the extraction of boundary signals which directly access visual memory. She showed that object recognition does not always dominate other figural cues such as symmetry. Rather, it is just one of the potential cues that the visual system uses to disambiguate surface relations. Object recognition

completely determines figural assignment only when there are no other competing cues present in the image. However, when other cues are present, object recognition might be overridden.

Spatial attention may alter figural assignment depending on which surface is cued (Baylis & Driver, 1996). Spatial attention can be directed by either endogenous or exogenous cues. An endogenous (or central) cue is a symbolic cue (i.e., an arrow pointing to the location of the upcoming stimulus) that is interpreted by the observer and that allows him/her to voluntarily orient attention to the relevant location in visual space. An exogenous (or peripheral) cue is a brief flash of light that automatically draws attention to its position. When endogenous cuing was employed on a display of ambiguous figures, attention biased figural assignment in the direction of the cue. Baylis and Driver (1996) found that exogenous cueing is not effective in biasing figural assignment.

In the new millennium, there is a resurgence of interest in figure-ground organization. New figural cues have been discovered, including lower region, top-bottom polarity, and extremal edges (Figure 6-1). Lower region refers to the tendency to assign figure to a surface in the lower part of the visual field (Vecera *et al.*, 2002). Such a tendency might arise from the fact that we gain more information from the lower part of the visual field where more objects are present (e.g., compare sky vs. ground) or from asymmetry in the visual capabilities for upper and lower visual fields (Previc, 1990). Vecera *et al.* (2002) argued that there is a close relationship between pictorial depth perception and figure-ground assignment. Figures are typically perceived as being closer to the observer as compared to the background. Thus, figural status depends on depth assignment. In particular, the principle of lower region is influenced by relative position. This is a pictorial depth cue where a surface positioned lower in a visual space appears to be closer to the observer as compared to the horizon line. Such a position makes the surface a likely candidate for figural status.

Top-bottom polarity describes the tendency to perceive a surface with a wide base and a narrow top as a figure. Examples of objects with a wide base and a narrow top are pyramids, cones, roofs, and hills. Of course, pyramids and cones can be rotated in space but their canonical position is with the wide part lying on the ground and the narrow part pointing up. Hulleman and Humphreys (2004) argued that top-bottom polarity constitutes an independent source of influence on figural assignment that cannot not be reduced on previously described principles.

Extremal edges are defined as the set of points whose sight lines (i.e., lines extending from the observer's eye to the boundaries of a visible surface) are tangent to the smoothly curved convex surface. A Surface with extremal edges tends to be perceived as being closer to the observer as compared to a surface without them. Extremal edges offer a strong cue for figural assignment that overrides some of the classical figural cues such as size or surroundness (Palmer & Ghose, 2008).

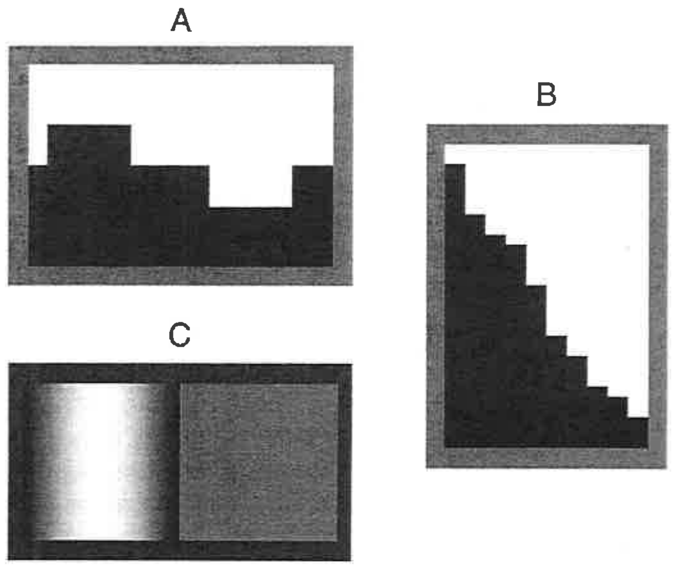

Figure 6-1. Recently discovered figural cues. A) Lower region induces the tendency to perceive lower surface as a figure (in this case, the black surface); B) Top-bottom polarity induces the tendency to perceive a surface with a wide base and a narrow top as a figure (in this case, the black surface); C) Extremal edges force us to assign figural status to the left surface.

With respect to attention, Vecera *et al.* (2004) showed that exogenous attention may also alter figural assignment. They noted that previous work by Baylis and Driver (1995) failed to find evidence for exogenous cues because cuing was performed inadequately, without respect to the surface borders. Baylis and Driver put the cue near the target surface and not on the target surface. When the cue was put on the target surface, it was perceived as a figure. It is interesting to note that Gestalt principles of figure-ground organization possess ecological validity, as demonstrated by Fowlkes *et al.* (2007) who showed that people segment real digital images and assign figural status to objects that are small, convex and positioned in the lower part of an image. As can be seen from the preceding review, psychophysical investigations offer a wealth of data that should be accounted for by a computational model of visual perception.

3. Neurophysiological perspective

The seminal work of Hubel and Wiesel (1968, 1977) on neuronal selectivity in the primary visual cortex laid the foundations for our understanding of the neural basis of shape perception. They discovered neurons with orientation selectivity, that is, neurons responding to the particular orientation of elongated bars or lines. Neurons showed maximal activity when the bar was oriented in one way and reduced activation when oriented in another. The reduction in activity was proportional to the degree of deviation from the optimal stimulus. For instance, neurons that showed a preference for a vertical bar would also show elevated activity for a bar oriented 10° from the vertical axis, higher than the activity for a horizontal bar. Based on these response properties, Hubel and Wiesel classified neurons with orientational selectivity into three groups: a) simple cells; b) complex cells; and c) hyper-complex or end-stopped cells. Simple cells are characterized by distinct excitatory and inhibitory regions within their receptive field and which extend along the axis of preferred orientation. Optimal stimulus for a simple cell is a luminance step (or edge) with appropriate orientation. Simple cells exhibit sensitivity to the polarity of edges, namely, they preferentially respond to white-black edges and not to black-white edges. Simple cells are usually grouped in pairs so that for the same orientation preference there is a simple cell with excitation on the left of the edge and another cell selective for opposite polarity (excitation on the right of the edge). Complex cells show orientation tuning but without distinct excitatory and inhibitory regions within their receptive fields. In other words, a luminance edge could be placed anywhere within the receptive field in order to invoke cell response. Hyper-complex or end-stopped cells also show orientation tuning, but with the additional property of selectivity for the size of the edge. End-stopped cells are particularly responsive to line ends and corners. According to Hubel and Wiesel, the three classes of orientation-selective cells are arranged in a hierarchy of visual processing. Input from the lateral geniculate nucleus (LGN) is first processed by the simple cells. They pool unoriented responses from the LGN[2] along the preferred axis of orientation, thereby producing a characteristic elongated receptive field with orientation selectivity. Simple cells with the same orientation

[2] LGN cells have concentric receptive fields with centre-surround antagonism.

preference and opposite contrast polarity feed their output to the complex cell. End-stopped cells further process output from the group of complex cells.

An important problem for the early work on neuronal selectivity in the visual cortex is that it did not relate clearly to visual perception. Responses of simple, complex or end-stopped cells just signal the presence of an edge at a particular location in the visual field but this does not allow us to discriminate figure from ground. In other words, the neurons do not take into account the context in which a particular edge is embedded (Albright & Stoner, 2002). Although the shape representation formed by neurons with orientation selectivity is an important first step in visual processing, it could not possibly be the only one. Many psychophysical investigations suggest the existence of an intermediate step involved in surface construction before object recognition takes place (Nakayama, He, & Shimojo, 1995).

In an attempt to examine in more detail the neural basis of figure-ground segregation, Lamme (1995) used a texture segregation task to monitor neural responses in V1. The display consisted of a figure region defined by the different orientation of line segments as compared to the background region. He found neurons that exhibited an enhanced firing rate to the interior of the figure region as compared to the control situation where no figure was presented. The stimulus was positioned in such a way that the receptive field of these cells was far from the boundary of the figure. Therefore, they could not pick up an orientation difference at the border between figure and ground. An enhanced firing rate to the interior of the figure was observed late in the neural responses (150 ms after stimulus onset), suggesting that feedback from higher visual areas is responsible for this effect (Lamme & Roelfsema, 2000). A later study showed that interior enhancement occurs also for a figure defined by a difference in luminance or depth (Zipser *et al.*, 1996). However, these findings were criticised by Rossi *et al.* (2001) who showed that interior enhancement is an artefact of the small surfaces used in the prior studies and that there is no such response when large figures are used as stimuli.

Zhou *et al.* (2000) found neurons in the V2 cortex selective for an oriented edge but whose response is modulated by the position of the figure relative to the edge. This property is termed border ownership selectivity. For instance, a neuron with a vertically oriented receptive field could show a strong response if the figure is on the left side but a weak response if the figure is on the right side of the border. Such a neuron is usually accompanied by another neuron that shows the same orientational selectivity at the same spatial location but with an opposite pattern of response to the location of a figure relative to the border (i.e., weak response if the figure is on the left side of the border and strong response if the figure is located on the right side of the border). Some of these neurons showed sensitivity to the polarity of the edge (different response to the black-white or white-black edge) but most responded independently of the polarity of the edge, indicating that their response truly reflects figural assignment. The timing of the border ownership response is fast, occurring in the range of 30-60 ms from stimulus onset. In a subsequent study, Qui *et al.* (2005) showed that border ownership selectivity is combined with the depth ordering of surfaces, as indicated by stereoscopic cues.

4. Computational perspective

Computational (i.e., neural network) models are designed to bridge the gap between the behavioural and the neural levels of explanation. The basic idea is to show how computational principles derived from neurophysiology can give rise to such behavioural phenomena as visual perception. The models vary greatly in scope, underlying assumptions, and biological plausibility. I will tentatively distinguish two approaches in modelling visual perception: 1) the bottom-up approach or biological approach which starts with known physiological facts and tries to work out how the identified neuron types and neural mechanisms produce relevant behaviour; and 2) the top-down or psychological approach, which starts from behavioural phenomena and tries to work out how detailed patterns of behavioural data could be fitted using known or hypothetical neural mechanisms.

4.1 Biological (border ownership) models

Biologically inspired models applied to figure-ground segregation are concerned with the explanation for how border ownership arises from the interaction of network elements[3]. Within this approach, I will distinguish two classes of models, depending on the type of the cortical interaction they propose underlies figure-ground perception. In the models by Zhaoping (2005) and Sakai and Nishimura (2006), border ownership is a consequence of intra-cortical or recurrent connections within a single visual area. On the other hand, Jehee *et al.* (2007) and Craft *et al.* (2007)

---

[3] Here, I will focus only on the models of figural assignment. For a review of models dealing with other apsects of boundary computation such as the formation of illusory contours or amodal completion, see Lesher (1995) and Grossberg (1997).

proposed that between-cortical or feedback connections from higher visual areas to lower visual areas are responsible for the emergence of border ownership responses in V2.

Zhaoping (2005) suggested that the border ownership response emerges from the interaction between excitatory and inhibitory nodes within recurrent neural networks. Mathematically speaking, recurrent networks are dynamical systems that describe the temporal evolution of neural activity. Dynamical systems are characterized by their initial state, energy landscape, and final state. The initial state of a recurrent network is determined by the input pattern. After initialization, network activity evolves until it converges to its final state, often-called stable state (attractor or fixed point). The energy landscape is an abstract description of the path that a dynamical system traverses from the initial state to the stable state (Figure 6-2). Each point in the landscape is associated with the amount of energy needed for the system to get to this point. Higher points in the energy landscape are more unstable (i.e., they require more energy) and the system attempts to settle to a less energy-demanding point. Stable states are local minima in the energy landscape that do not drive the network to change its activity any more. Stable states are often called attractors because they attract the network dynamics to their position in the state space. The stable state that the network will end up in depends on the initial activation of the model neurons and on the pattern of the synaptic weights that connect them. Connections in the network should be set up in such a way that the network favours the interpretation that is most consistent with the global configuration of the input.

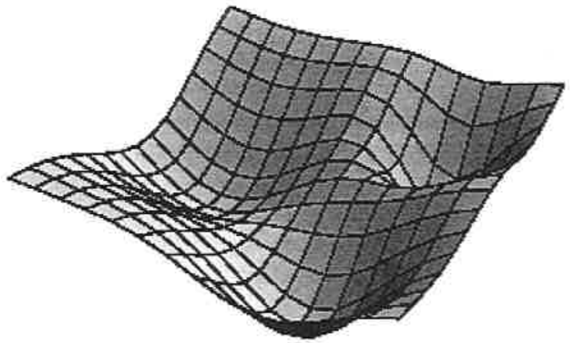

Figure 6-2. An example of the energy landscape of dynamical system. In this case, three stable points exist (i.e., three holes in the landscape). A dynamical system could be conceived as a ball that rolls on this landscape until it settles on one of the holes. The ball's starting position determines which hole it will be attracted to.

The model of a V2 cortex proposed by Zhaoping (2005) is a recurrent neural network composed of a 2-D map of nodes that are sensitive to different orientations and with the opposite pattern of border ownership response for each orientation at each spatial location on the map (see Figure 6-3). The nodes respond to the input image presented in Figure 6-3a. Each node has three types of excitatory connections and three types of inhibitory connections. Excitatory connections are projected (Figure 6-3b):

1) to collinear nodes, that is, nodes sensitive to the same orientation along the boundary also sensitive to the same side of the boundary;

2) to nodes located laterally, sensitive to perpendicular orientation but that point to the same side of the surface as it occurs at the corners and;

3) to nodes with a perpendicular orientation located anterior as it occurs on configurations of boundaries that are interpreted as an occlusion (e.g., on T-junctions).

Such boundary configurations usually occur when one object partially occludes the other (as in Figure 6-3a). In this case, the visual system favours the occluding object as a figure and not the occluded one. Inhibitory connections are projected in an opposite pattern from excitatory connections, that is, inhibition occurs among nodes which are selective for the opposite direction of border ownership along a straight border, at corners or at T-junctions (Figure 6-3c). When the input pattern is presented to such network, the initial activation of the border ownership nodes is ambiguous. However, the network dynamics will enhance the activation of the nodes that are consistent with one interpretation of the border ownership due to their mutual excitation and will suppress the activation of those nodes that signal an opposite interpretation of border ownership. Computer simulations showed that networks with the described pattern of excitatory and inhibitory connections do converge to the stable state that reflects the correct interpretation of the border ownership relationship in the input image.

Sakai and Nishimura (2006) attributed border ownership responses to the extra-classical component of the receptive fields. Neurophysiological research showed that, besides the stimulus within the classical receptive field, the activity of many neurons is modulated by the presence of stimuli outside the receptive field. Modulation could be facilitatory or suppressive depending on spatial relations and on the difference in orientation and contrast of the stimuli presented to the classical and the non-classical receptive fields. Modulation is effective only when there is simultaneous input within the classical receptive field. This means that a stimulus positioned outside the classical receptive field is unable to drive a neuron on its own but it can only modulate the neuron's response driven by the stimulus within the receptive field. The existence of such modulation is considered to be important for understanding the role of context in visual perception (Albright & Stoner, 2002). Sakai and Nishimura (2006) argued that the precise shape of the extra-classical receptive field is unimportant. It is sufficient that cells receive distinct subunits of excitatory and inhibitory modulatory influences in order to exhibit border ownership. They performed an extensive set of computer simulations to show that various configurations of extra-classical receptive fields are able to produce correct border ownership assignment.

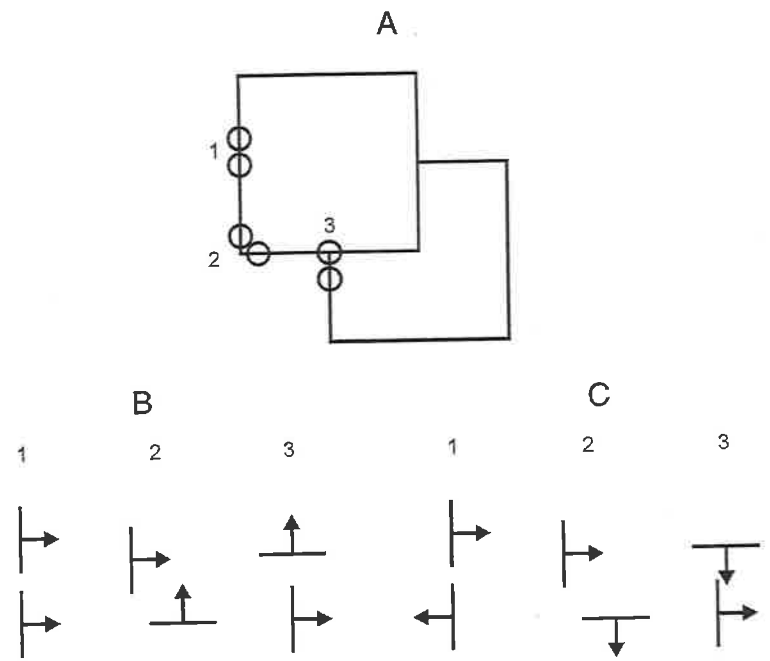

Figure 6-3. Recurrent excitatory and inhibitory interaction in the Zhaoping model (2005). A) Input image with two squares where one square partially overlaps the other. Circles denote the locations of the border ownership nodes and the numbers refer to the three boundary situations described in the text. 1 – straight line; 2 – corner; 3 – T-junction. B) Excitatory connections among nearby border ownership nodes. Lines indicate their preferred orientation and attached arrows indicate their figural preference (e.g., arrow pointing to the left means that the figure is on the left side); C) Inhibitory connections among same border ownership nodes.

Craft *et al.* (2007) and Jehee *et al.* (2007) claimed that border ownership might arise from the interaction between V2 and higher visual areas, such as V4 or IT. It is known that these cortical areas contain neurons with larger receptive fields and neurons that are selective for more complex shapes. Selectivity for complex shapes may contribute to the disambiguation of locally ambiguous boundary responses in V1 and V2. In the model by Craft *et al.* (2007), V2 projects its output to a layer of nodes selective to ring-shaped stimuli of various sizes. Activity from this layer is fed back to V2, where the competition among neurons tuned to the opposite side of border ownership is biased toward the solution that receives more support from the shape-selective nodes (Figure 6-4). Consider, for example, a stimulus with two squares positioned in such a way that one square partially overlaps the other (as in Figure 6-3a). Both squares activate a set of edge selective cells in V1. Output from V1 is projected to the border nodes (denoted in the model as B-nodes), which show orientational selectivity with the preference for one direction of border ownership. The B-nodes with the same orientational selectivity at the same spatial location but with the opposite direction of border ownership mutually inhibit each other. The B-nodes project their output to the grouping nodes (or G-nodes in the model). These operate as feature detectors. Their activation is proportional to the match of the shape of their receptive field with the shape of the input pattern (Figure 6-4a). In the case of two overlapping squares, stronger activity will be observed in the G-node positioned over the occluding square because this node receives more edge responses than the G-node positioned over the occluded square. At the location of an occluded object, part of the edge response is missing (due to the occlusion) so that the corresponding G-node will show less activation. The G-node sends excitation to the B-nodes that point to the direction of the border ownership that is consistent with the location of the G-node. Therefore, the B-nodes that point to the interior of the occluding surface will show greater activation (less feedback inhibition) than the B-node that points in the opposite direction. This indicates that the network assigns figural status to the occluding surface (Figure 6-4b). A similar idea has been put forward by Jehee *et al.* (2007), who noted that higher visual areas have lower spatial resolution and thus could detect more global aspects of the input pattern.

Another argument for the inter-cortical origin of border ownership is the speed of the signal flow within and between cortical areas. Physiological measurements on the conduction velocities of axons showed that neural activation spreads much faster between cortical areas than within. The recurrent connections within a cortical layer are relatively slow. Therefore, Craft *et al.* (2007) concluded that only the feedback models fit the temporal constraint imposed by the high speed of border ownership responses.

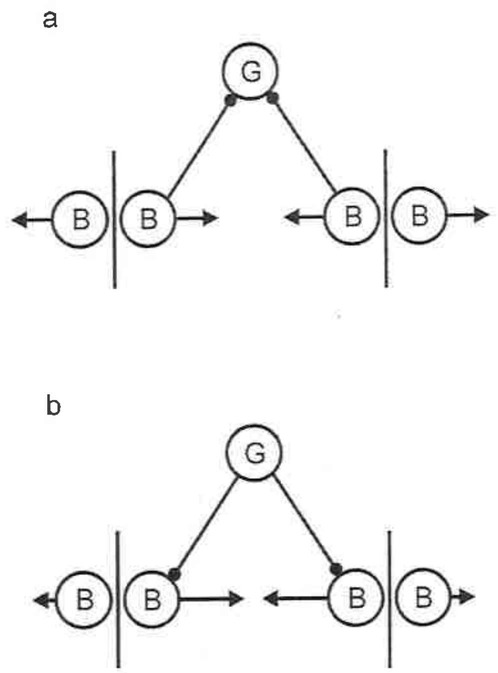

Figure 6-4. Interactions between the B-nodes and the G-nodes in the model of figural assignment proposed by Craft *et al.* (2007). Circles with letters denote the model's nodes. Arrows indicate the figural preference of the node and the length of the arrow indicates the amount of the node's activity. Straight vertical lines denote the locations of the edges. The lines between the B-nodes and the G-nodes are axons and the small black disks at the end of the lines denote the synaptic connections. Panel a) Feedforward excitatory connections from the B-nodes to the G-node enable the G-node to detect the presence of the global pattern (e.g., the square). Panel b) Feedback excitatory connections from the G-node to the B-nodes biases competition between nearby B-nodes with opposite border ownership preference. A consequence of the feedback bias is that the B-nodes pointing to the interior of the figure will win the competition, as indicated by the greater length of the arrows.

The reviewed models offer valuable insight into the process of figural assignment, but it is also necessary to examine their weak points. An important criticism of the border ownership models is that none of them

relates precisely to our subjective experience. Namely, we experience surfaces together with their colour, brightness or texture. We do not experience the world as a set of outlines or wireframes. Although shape information is important for object recognition, it is not the only information available. Another constraint for the border ownership models is that they have not yet demonstrated whether they are able to explain the role of the lower region, top-bottom polarity, extremal edges, spatial frequency, and horizontal-vertical orientation in figure-ground perception. The potential problem is that these models do not utilise information available on the interior of the surface. For instance, the spatial frequency content of the surface is difficult to read from the neural responses at the borders alone. Furthermore, border signal interactions might not be sufficient to disambiguate locally contradicting border ownership information in complex images involving several surfaces sharing a common border (Sakai & Nishimura, 2006). Also, border ownership models do not incorporate mechanisms of attentional enhancement, so it is not clear how endogenous and exogenous attention might bias figural assignments.

4.2 Psychological (filling-in) models

Psychological models of visual perception are concerned with the explanation of how the same computational mechanisms lead to a veridical perception in one circumstance and, at the same time, lead to such errors as visual illusions in other circumstances. During the 1980s and 1990s Grossberg (1994; 1997; Grossberg & Mingolla, 1985; Grosssberg & Todorović, 1988) developed a sophisticated model of visual perception called FACADE (Form And Colour And DEpth). He proposed two complementary processing streams, called boundary contour system (BCS) and feature contour system (FCS), which specialise in extracting different types of visual information. The BCS computes oriented edges and represents an outline of an image while it suppress featural information such as colour, brightness and texture. The FCS computes brightness and colour estimates that are independent of the variable illumination conditions. The FCS achieves colour and brightness constancy by computing luminance ratios at the surface borders and suppressing luminance gradients at the interior of the surfaces. BCS and FCS signals are combined at a final processing stage called filling-in. This stage is considered to be isomorphic with perception. It enables FCS signals from the borders to propagate along the whole surface. Signal propagation is stopped at the locations where BCS signals exist. Filling-in could be conceived as a neural implementation of the perceptual filling-in that occurs at the blind spot, in neon colour spreading, and during viewing of stabilised retinal images (Komatsu, 2006, Pessoa *et al.*, 1998). A great variety of perceptual phenomena such as geometrical and brightness illusions can be explained by the interaction between the BCS and the FCS (Grossberg & Mingolla, 1985; Grossberg & Todorović, 1988). Empirical predictions derived from the theory were corroborated in several studies that specifically looked for signatures of interaction between independent surface and boundary representation (Elder & Zucker, 1998; Gilchirst *et al.*, 1997; Rogers-Ramachandran & Ramachandran, 1998). However, despite the great success of the FACADE theory, the claim that there exists a neural process in the brain analogous to perceptual filling-in is still controversial (Komatsu, 2006). Some studies failed to find support for a spreading of neural activation during surface perception (von der Heydt *et al.*, 2003).

According to Grossberg (1994, 1997), figure-ground separation is solved as a part of the depth stratification process. Within FACADE, different depth planes are represented by different network layers. Each network layer is sensitive to a different spatial frequency content or spatial scale. A network sensitive to a small scale will pick up image details corresponding to surfaces close to the observer. The reason is that closer objects are seen more clearly than distant objects. Furthermore, a large-scale network will pick up only gross features of an image. Inhibitory interactions between network layers, where small-scale nodes inhibit corresponding locations in the large-scale network, produce a stratified representation. As a consequence of these network interactions, the surface represented in a network layer for a near-depth plane is perceived as a figure, while the surface on a layer for far depth plane is perceived as a ground. It should be noted that this description is an over-simplification. The actual model involves much more complex network mechanisms and interactions that provide an explanation of binocular depth perception, perception of transparency, and 3-D interpretation of 2-D images.

Another psychologically motivated approach to figure-ground perception was proposed by Domijan and Šetić (2008). They elaborated on the neural mechanisms proposed by Grossberg (1994, 1997) in order to account for classical and recently discovered principles of figure-ground organization. In their model, the ventral stream computes surface boundaries in a similar vein as the BCS. Output of the ventral stream is used in the dorsal stream to compute saliency of spatial locations. It is suggested that the parietal cortex, as a central area within the dorsal stream, uses neurons with large receptive fields in order to measure the density of boundary signals. A stronger response is observed at regions

with more boundary signals. Furthermore, the output from the dorsal stream is augmented with a spatial gradient that favours locations in the lower visual field. Such a gradient enables modelling of lower- region and top-bottom polarity. The final stage, where the output of ventral and dorsal stream output converges, is called figural filling-in. It is a neural network with lateral inhibition that assigns different activity levels to different surfaces. The system solves the feature-binding problem by using the firing rate or amplitude of neural activity as a code for surface representation (Domijan, 2004). Surface with the maximal activity is defined as a figure and all other surfaces represented by lower activity levels are background. The simulations showed that the proposed model was able to explain how figural assignment is influenced by the classical and the recently discovered cues and how attention alters the figural status of the surface (Domijan & Šetić, 2008). In evaluating the figural filling-in model from a neurophysiological perspective, it should be noted that filling-in is a slow process that requires time to complete. The time needed to complete a surface representation is proportional to the size of the surface. Therefore, models based on the filling-in process are unable to capture the fast dynamics of the border ownership neurons in V2.

A point of departure from the Grossberg's model is the claim that figure-ground segregation is solved independently of depth stratification within a single-scale neural network. Whether depth stratification precedes or is followed by figure-ground separation is still an open issue that deserves further empirical investigation. Although the concept of figural filling-in was successful in accounting for behavioural data, it still lacks biological plausibility because the evidence for the neural analogue of filling-in is lacking (von der Heydt *et al.*, 2003). One interesting possibility is that many previous studies were explicitly looking for evidence for brightness filling-in while, in fact, they were observing figural filling-in. For instance, some studies reported evidence for the spread of neural activity in V1, but this neural filling-in was not correlated with the brightness perception as expected (Hung *et al.*, 2007). However, since they did not manipulate the figural status of the surfaces, it is possible that they captured figural filling-in and not brightness filling-in. Further research is needed to examine these issues, along with the search for alternative computational mechanisms that can support filling-in and that are more consistent with neurophysiological findings (Komatsu, 2006).

5. Conclusions

Figure-ground organization is an important early step in visual processing. It separates the structured input to which processing efforts should be devoted (figure) from the less structured inputs in the background. Gestalt psychologists identified many variables that can influence figure-ground assignment including size, contrast, surroundness, convexity, symmetry, parallelism, and horizontal-vertical axes. Recently, several new factors have been identified, including lower region, top-bottom polarity, and extremal edges.

Several neurophysiological investigations (Lamme, 1995; Zipser *et al.*, 1996) of figure-ground assignment found enhanced firing rate responses to the figure in the primary visual cortex of monkeys. Enhanced activity was observed in the region corresponding to the perceived figure on a texture segregation task. Other researchers (Zhou *et al.*, 2000; Qui *et al.*, 2005) found a special group of neurons in V2 that detect border ownership. 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 with opposite figural selectivity will increase its firing rate.

Many computational models have been proposed to explain the psychological and neurobiological properties of figure-ground segregation. Several models focused on boundary assignment alone. Other models are concerned with explaining psychophysical findings (i.e., how figural cues give rise to figural assignment). Although the models share some common assumptions, there are also important differences. Border ownership models cannot explain the role of many of the identified figural cues. Moreover, psychophysical models make assumptions not supported by neurophysiological data. An important problem for future research is how to reconcile psychophysical and neurophysiological models in order to provide a unified account of figure-ground organization.

Acknowledgments

I would like to thank Danko Nikolić and the anonymous reviewer for their helpful comments and suggestions that improved the manuscript. This work was supported by grants received from the Croatian National Science Foundation (02.05/06), Croatian Ministry of Science, Education and Sport (009-0362214-0818) and the Bial Foundation (80/06).

References

Albright, T. D., & Stoner, G. R. (2002). Contextual influences on visual processing. *Annual Review of Neuroscience, 25*, 339–379.

Bahnsen, P. (1928). Eine Untersuchung über Symmetrie und Asymmetrie bei visuellen Wahrnehmungen. *Zeitschrift für Psychologie, 108,* 129-154.

Baylis, G.C., & Driver, J. (1995). One-sided edge assignment in vision: 1. Figure-ground segmentation and attention to objects. *Current Directions in Psychological Science, 4*, 140-146.

Craft, E., Schütze, H., Niebur, E., & von der Heydt, R. (2007). A neural model of figure-ground organization. *Journal of Neurophysiology, 97*, 4310-4326.

Domijan, D. (2004). Recurrent network with large representational capacity. *Neural Computation, 16*, 1917-1942.

Domijan, D., & Šetić, M. (2008). A feedback model of figure-ground assignment. *Journal of Vision, 8(7):10*, 1-27.

Elder, J., & Zucker, S. (1998). Evidence for boundary-specific grouping. *Vision Research, 38*, 143-152.

Fowlkes, C.C., Martin, D.R., & Malik, J. (2007). Local figure-ground cues are valid for natural images. *Journal of Vision, 7*, 1-9.

Gilchrist, I., Humphreys, G.W., Riddoch, M.J., & Neumann, H. (1997). Luminance and edge information in grouping: A study using visual search. *Journal of Experimental Psychology: Human Perception and Performance, 23*, 464-480.

Grossberg, S. (1994). 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., & Mingolla, E. (1985). Neural dynamics of form perception: Boundary completion, illusory figures, and neon color spreading. *Psychological Review, 92*, 173-211.

Grossberg, S., & Todorović, D. (1988). Neural dynamics of 1-D and 2-D brightness perception: A unified model of classical and recent phenomena. *Perception & Psychophysics, 43*, 241-277.

Harrower, M.R. (1936). Some factors determining figure-ground articulation. *British Journal of Psychology, 26*, 407-424.

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

Hubel, D.H., & Wiesel, T.N. (1977). Functional architecture of macaque monkey visual cortex. *Proceedings of the Royal Society of London (B), 198*, 1-59.

Hulleman, J., & Humpherys, G.W. (2004). A new cue to figure-ground coding: top-bottom polarity. *Vision Research, 44,* 2779-2791.

Hung, C.P., Ramsden, B.M., & Roe, A.W. (2007). A functional circuitry for edge-induced brightness perception. *Nature Neuroscience, 10,* 1185-1190.

Jehee, J.F.M., Lamme, V.A.F., & Roelfsema, P.R. (2007). Boundary assignment in recurrent network architecture. *Vision Research, 47,* 1153-1165.

Kanizsa, G., & Gerbino, W. (1976). Convexity and symmetry in figure–ground organization. In M. Henle (Ed.), *Vision and artifact* (pp. 25-32). New York: Springer.

Klymenko, V., & Weisstein, N. (1986). Spatial frequency difference can determine figure-ground organization. *Journal of Experimental Psychology: Human Perception and Performance, 12*, 324-330.

Koffka, K. (1935). *Principles of gestalt psychology*. New York: Harcourt Brace.

Komatsu, H. (2006). The neural mechanisms of perceptual filling-in. *Nature Reviews Neuroscience, 7*, 220-231.

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., & Roelfsema, P.R. (2000). The distinct modes of vision offered by feedforward and recurrent processing. *Trends in Neuroscience, 23*, 571-579.

Lesher, G. (1995). Illusory contours: toward a neurally based perceptual theory. *Psychonomic Bulletin and Review, 2*, 279–321.

Metzger, W. (1953). *Gesetze des Sehens*. Frankfurt, Germany: Waldemar Kramer.

Nakayama, K., He, Z.J., & Shimojo, S. (1995). Visual surface representation: A critical link between lower-level and higher-level vision. In S.M. Kosslyn & D.N. Osherson (Eds.), *Visual cognition: An invitation to cognitive science, Vol. 2* (pp. 1-70). Cambridge, MA: MIT Press.

Palmer, S.E. (1999). *Vision science*. Cambridge, MA: MIT Press.

Palmer, S.E., & Ghose, T. (2008). Extremal edges: A powerful cue to depth perception and figure-ground organization. *Psychological Science, 19*, 77-84.

Pessoa, L., Thompson, E., & Noë, A. (1998). Finding out about filling-in: a guide to perceptual completion for visual science and the philosophy of perception. *Behavioral and Brain Sciences, 21*, 723-802.

Peterson, M.A. (1994). Object recognition processes can and do operate before figure-ground organization. *Current Directions in Psychological Science, 3*, 105-111.

Peterson, M.A., & Gibson, B.S. (1994a). Must figure-ground organization precede object recognition? An assumption in peril. *Psychological Science, 5*, 253-259.

Peterson, M.A., & Gibson, B.S. (1994b). Object recognition contributions to figure-ground organization: Operations on outlines and subjective contours. *Perception & Psychophysics, 56*, 551-564.

Pomerantz, J.R., & Kubovy, M. (1986). Theoretical approaches to perceptual organization. In K.R. Boff, L. Kaufman, & J.P. Thomas (Eds.), *Handbook of perception and human performance: Vol. 2. Cognitive processes and performance* (pp. 1-46). New York: Wiley.

Previc, F.H. (1990). Functional specialization in the lower and upper visual fields in humans: Its ecological origins and neurophysiological implications. *Behavioral and Brain Sciences, 13,* 519–575.

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.

Roelfsema, P.R. (2006). Cortical algorithms for perceptual grouping. *Annual Review of Neuroscience, 29*, 203-227.

Rogers-Ramachandran, D.C., & Ramachandran, V.S. (1998). Psychophysical evidence for boundary and surface systems in human vision. *Vision Research, 38*, 71-77.

Rossi, A.F., Desimone, R., & Ungerleider, L.G. (2001). Contextual modulation in primary visual cortex of macaques. *Journal of Neuroscience, 21*, 1698-1709.

Rubin, E. (1958). Figure and ground. In D. C. Beardslee & M. Wertheimer (Eds.), *Readings in perception* (pp. 194-203). Princeton, NJ: Van Nostrand. (Original work published 1915)

Sakai, K., & Nishimura, H. (2006). Surrounding suppression and facilitation in the determination of border ownership. *Journal of Cognitive Neuroscience, 18*, 562-579.

Vecera, S.P., Flevaris, A.V., & Filapek, J.C. (2004). Exogenous spatial attention influences figure-ground assignment. *Psychological Science, 15*, 20-26.

Vecera, S.P., Vogel, E.K., & Woodman, G.F. (2002). Lower region: A new cue for figure-ground assignment. *Journal of Experimental Psychology: General, 131,* 194-205.

von der Heydt, R., Friedman, H., & Zhou, H. (2003). Searching for the neural mechanisms of color filling-in. In L. Pessoa, & P. De Weerd (Eds.), *Filling-in: From Perceptual Completion to Cortical Reorganization* (pp. 106-127), Oxford: Oxford University Press.

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. *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.

Summary

The phenomenon of figural-ground assignment is analysed from three complementary perspectives: psychological, neurophysiological, and computational. The classical and recent psychological findings that identified cues influencing figure-ground segregation are reviewed. Next, relevant neurophysiological data are provided that show how cortical structures encode the figure-ground relationship. Finally, recent neurocomputational models are described that attempt to explain previous psychological and neurophysiological findings. Based on these underlying assumptions, the models are grouped into biological (or border ownership) models and psychological (or filling-in) models. Biological models are focused on explaining how sensitivity for figural status of the boundary (border ownership) arises from the interaction of cortical areas. Psychological models attempt to explain the effectiveness of figural cues as a consequence of the spreading activation between an object's boundaries (filling-in). Analysis of the ability of the diverse models to handle a complicated set of empirical data suggests that no model to date provides a complete account of processes involved in figural assignment. An important problem for future research is how to reconcile psychophysical and biological models in order to provide a unified account of figure-ground organization.

Chapter Seven

Differentiation of Mental and Motor Components in a Psychomotor Task

Ilija Manenica,[1] Ana Brečić,[2A] and Ana Šuto[2B]

1. Introduction

All working activities include mental and motor components in different quantities, where some of the activities are predominantly mental and others predominantly motor (physical). Usually, psychomotor tasks do not require extreme mental or motor efforts, but they can often be rather demanding, nevertheless. According to Veltman and Gaillard (1996), some authors define mental load as the ratio between task requirements and channel capacity (O' Donnell, & Eggemeier, 1986; Kantovitz and Elvers, 1988), while others define it in terms of the information needed for efficient task performance (Manenica & Krošnjar, 1990).

Efficiency in psychomotor tasks is mainly limited by the mental load component, which is difficult to assess. There are two reasons for this. The first is the lack of an appropriate definition of mental load, so that it is difficult to quantify it at the input. The second difficulty in its assessment is the limited capacity of the human processing system, where overloading leads to a deterioration in performance. The mental processes such tasks entail are rather complex and they consist of at least three stages, i.e., information seeking, information processing, and decision-making. The

[1] Professor emeritus, University of Zadar, Zadar, Croatia (ilija.manenica@gmail.com).
[2] Department of Psychology, University of Zagreb, Zagreb, Croatia
[A] Ph.D. Student (brecic@gmail.com).
[B] Junior Assistant (anasuto@gmail.com).