BRILL

Art & Perception 5 (2017) 1–95

brill.com/artp

How Humans Consciously See Paintings and Paintings Illuminate How Humans See

Stephen Grossberg[1,*] and Lauren Zajac[2]

[1]Center for Adaptive Systems, Graduate Program in Cognitive and Neural Systems, Departments of Mathematics & Statistics, Psychological & Brain Sciences, and Biomedical Engineering, Boston University, Boston, MA 02215, USA

[2]Department of Anatomy & Neurobiology, Boston University School of Medicine, 72 East Concord Street (L 1004), Boston, MA 02118, USA

Received 14 September 2016; accepted 19 December 2016

Abstract

This article illustrates how the paintings of visual artists activate multiple brain processes that contribute to their conscious perception. Paintings of different artists may activate different combinations of brain processes to achieve their artist's aesthetic goals. Neural models of how advanced brains see have characterized various of these processes. These models are used to explain how paintings of Jo Baer, Banksy, Ross Bleckner, Gene Davis, Charles Hawthorne, Henry Hensche, Henri Matisse, Claude Monet, Jules Olitski, and Frank Stella may achieve their aesthetic effects. These ten painters were chosen to illustrate processes that range from discounting the illuminant and lightness anchoring, to boundary and texture grouping and classification, through filling-in of surface brightness and color, to spatial attention, conscious seeing, and eye movement control. The models hereby clarify how humans consciously see paintings, and paintings illuminate how humans see.

Keywords

Painting, neural networks, Baer, Banksy, Bleckner, Davis, Hawthorne, Hensche, Matisse, Monet, Olitski, Stella

> "…nobody wants to be told that somebody sees better than they do. But I believe it is the case. Some painters are experts…they have given their lives to it…Now that to me is the glory of art."
>
> (*Sergy Mann* et al., 2016, p. 267)

*To whom correspondence should be addressed. E-mail: steve@bu.edu

© S. Grossberg and L. Zajac, 2017
DOI: 10.1163/22134913-00002059
This is an open access article distributed under the terms of the Creative Commons Attribution-Noncommercial 4.0 Unported (CC-BY-NC 4.0) License.

Downloaded from Brill.com04/06/2021 01:46:22AM
via free access

1. Introduction

1.1. Paintings as Probes of Brain Processes

Whenever an artist manipulates a canvas, and experiences conscious percepts of an emerging painting, the artist is performing an experiment that probes different combinations of the brain processes whereby we see. Artists typically succeed in doing so without having explicit knowledge about the brain processes that mediate between painterly manipulation and percept. The particular interests and aesthetic sensibilities of different artists have led each of them to instinctively emphasize different combinations of these brain processes. These different combinations may be one hallmark of different artists' styles, and indeed of entire artistic movements, and they evolve with each artist's experiences of the world.

By working to achieve aesthetic and conceptual goals on two-dimensional surfaces, painters have hereby explored and exploited principles of the human visual system. As a result of this process, there is an incredible richness of issues that paintings elicit, both scientific and aesthetic. This article addresses various of these issues through a discussion of specific paintings by well-known artists that exploit different combinations of brain processes in order to achieve their aesthetic goals. Illustrative paintings or painterly theories by ten artists will be given a unified analysis in the light of neural design principles and mechanisms that have been articulated and computationally characterized by the most advanced neural models of how advanced brains consciously see. Where appropriate, the structure of the examined paintings will be tied to the artist's intentions, or to reviews of the artist's work written by art historians, curators, or critics.

There are many other sources available that link the science of vision to the appreciation of art. The informative and beautiful books by Margaret Livingstone (Livingstone, 2002) and Semir Zeki (Zeki, 1999) are particularly notable. However, these are contributions by leading experimental neuroscientists that describe fascinating facts about the brain and visual perception, but do not explain how brain mechanisms give rise to conscious visual percepts. There is also a vast field of art history and aesthetic theory, and many alternative neural models, that do not make the link between brain dynamics and conscious psychological experiences, and will thus not be further discussed.

Cavanagh (2005) has noted that "discrepancies between the real world and the world depicted by artists reveal as much about the brain within us as the artist reveals about the world around us" (p. 307). This article explains some of the ways in which visual percepts differ from the 'real world', notably how certain visual illusions have been exploited in paintings by artists. The main focus of the article is not, however, on discrepancies between the real world and the world depicted by artists, but is rather upon the brain organizational

Downloaded from Brill.com04/06/2021 01:46:22AM
via free access

principles and mechanisms that lead to conscious visual percepts, and how these principles and mechanisms have influenced the paintings of different artists.

Other writers have questioned the ability of science to clarify how humans appreciate the beauty of artworks. For example, Conway and Rehding (2013) write that "it is an open question whether an analysis of artworks, no matter how celebrated, will yield universal principles of beauty" and that "rational reductionist approaches to the neural basis for beauty… may well distill out the very thing one wants to understand… Its progress in uncovering a beauty instinct, if it exists, may be accelerated if the field were to abandon a pursuit of beauty per se and focus instead on uncovering the relevant mechanisms of decision making and reward and the basis for subjective preferences… This would mark a return to a pursuit of the mechanisms underlying sensory knowledge: the original conception of aesthetics."

The current article makes no attempt to describe "universal principles of beauty." Nor does it "focus on uncovering the relevant mechanisms of decision making and reward and the basic for subjective preferences." Such an analysis would require the study of how perceptual, cognitive, and *emotional* processes interact. Although there are some promising approaches to trying to understand aesthetic emotions by using mathematical models of the mind (e.g., Perlovsky, 2010), our goal is to first try to better understand the brain mechanisms of perception and cognition whereby humans see paintings, and whereby painters have achieved their aesthetic goals. There are also related modeling studies of how the perceptual and cognitive processes that are described herein interact with emotional processes to create conscious experiences of seeing, knowing, and *feeling* (e.g., Grossberg, 2013, 2017). We believe that these more comprehensive theoretical insights will more easily be applied to an appreciation of how visual art is experienced when it can build upon insights such as those described herein.

This article accordingly describes illustrative paintings or painterly theories of ten artists in light of neural models such as the Form-And-Color-And-DEpth (FACADE) model of 3D vision and figure–ground perception (e.g., Grossberg, 1994, 1997; Grossberg and McLoughlin, 1997; Kelly and Grossberg, 2000), and the 3D LAMINART model of how identified laminar circuits in visual cortex embody and extend FACADE design principles and mechanisms (e.g., Cao and Grossberg, 2005, 2012; Fang and Grossberg, 2009; Grossberg, 1999; Grossberg, Mingolla, and Ross, 1997; Grossberg and Yazdanbakhsh, 2005). These are currently the most advanced computational theories of how brains see. In particular, these models have provided unified explanations and predictions of much more psychological and neurobiological data about vision than other models. They are also the only available neural models that propose an explanation of what happens in a viewer's brain when

Downloaded from Brill.com04/06/2021 01:46:22AM
via free access

having a conscious visual experience, including the viewing of a painting. These models will therefore be used to link paintings, as visual inputs, to the brain mechanisms that create conscious visual percepts of the paintings. This discussion is organized in a way that may shed new light on artists' aesthetic struggles, and on how humans see the paintings that resulted from them.

The article thus begins with a review of key concepts and mechanisms from these models as a basis for discussing particular paintings by different artists. The article hereby continues in the tradition of the Grossberg (2008) article on *The Art of Seeing and Painting*, but with an expanded selection of artists to emphasize different combinations of brain processes. Using these examples, artists can, in the future, have at their disposal a more precise knowledge of how their painterly manipulations may give rise to the conscious percepts of a completed painting.

There is a vast experimental literature about vision that is relevant to the chosen paintings, but that cannot be comprehensively reviewed in this article for lack of space. Key data will nonetheless be described in order to make specific points. Interested readers can find many other data described and explained in the cited modeling articles, as well as various data that these models have successfully predicted. There are also alternative vision models that will not be discussed because they do not incorporate key concepts and mechanisms that will be needed below to carry out our analyses and do not provide a linking hypothesis to conscious seeing. Comparative analyses of various of these models can also be found in the cited modeling articles.

1.2. From Gist to Scene Understanding

When one first looks at any scene, whether it is in a painting or in the physical three-dimensional world, the information extracted from it most quickly is the *gist* (Friedman, 1979; Intraub, 1999; Oliva, 2005; Potter, 1976; Potter and Levy, 1969). Gist is capable of providing, within a single glance, sufficient information for recognizing what type of scene it is, whether of a city street, forest, mountain, coast, or countryside; cf., Oliva and Torralba (2001) and Grossberg and Huang (2009). Our discussions below will review how the gist of a scene can be computed from properties of the basic functional units of vision—boundaries and surfaces—as can a wide range of other scenic properties, including 3D shape, boundary groupings, texture, shading, color, brightness, and object identity. All of these properties can interact to influence how a 3D scene or a 2D painting can generate a context-sensitive 3D representation in the mind of a viewer.

Gist can be derived from a process that is primarily *bottom-up*; that is, one that can be derived from information coming from a scene to the retinas, which then send signals that propagate ever deeper into the brain. *Top-down* processes can also strongly influence the 3D representations that we see; that

Downloaded from Brill.com04/06/2021 01:46:22AM
via free access

is, processes that begin at higher levels of brain processing and send signals that propagate towards levels ever closer to our retinas. These top-down processes include volitionally-controlled spatial attention and scenic expectations that are based upon knowledge of objects and scenes that have been learned from previous experience. Our mechanistic account will invoke both bottom-up and top-down processes to derive its unified comparative analysis of different paintings.

The emphasis on top-down vs. bottom-up processes has a long history in visual perception, and has triggered an enduring controversy about how we see and recognize the world. Hermann von Helmholtz, one of the greatest scientists of the nineteenth century, advocated a top-down view when he proposed that we see using *unconscious inferences*, or learned expectations (Fig. 1), to see what we expect to see, based on past experiences. In contrast, Gaetano Kanizsa provided brilliant counterexamples to Helmholtz's hypothesis using images, such as the one in Fig. 2 (right panel), that violate expectations (Kanizsa 1955, 1974, 1979). Here, the percept of an emergent cross whose horizontal bars lie behind the emergent square, but whose vertical bars lie in front of it, should be impossible if our learned expectations about squares usually being flat determined the percept. Kanizsa emphasized the power of bottom-up visual processes, such as adaptive filtering and perceptual grouping, that act directly on visual scenes and images at early stages of brain processing.

Both Helmholtz and Kanizsa were partly correct. A more comprehensive understanding has arisen from neural models that explain how bottom-up and

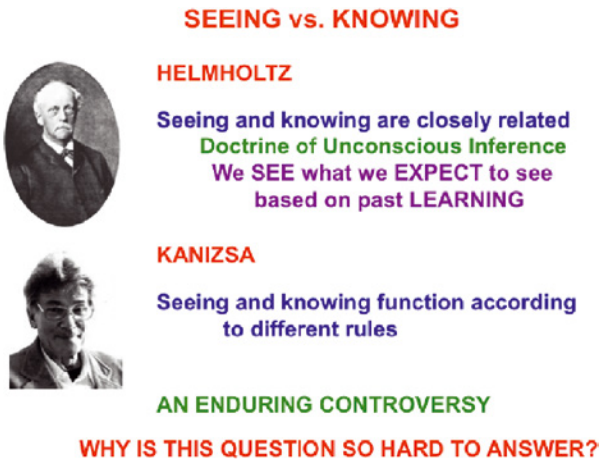


Figure 1. Helmholtz emphasized top-down effects on perception, which he called "unconscious inferences." This concept embodied the idea that we see what we expect to see based upon past experiences. Kanizsa, in contrast, emphasized that perceptual processes for seeing often do not lead to percepts that would be expected from learned familiarity.

Downloaded from Brill.com04/06/2021 01:46:22AM
via free access

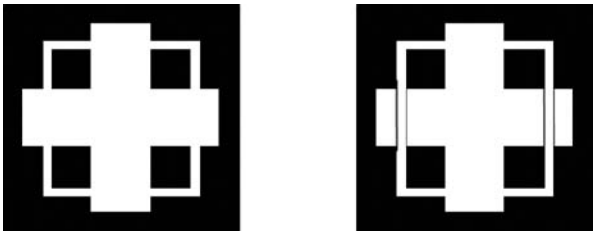


Figure 2. Kanizsa stratification images. (Left) The left image is most frequently seen as a cross in front of a partially occluded square. The percept is, however, bistable with the square occasionally seen in front of the cross. In each percept, the ambiguous white within the boundaries of the figure that is seen in front belongs to that figure, so that the other figure is completed amodally behind it. (Right) When vertical lines bound the vertical bars of the square, the corresponding bars are seen in front of the square. Despite this fact, the horizontal bars of the square are most frequently seen behind the vertical bars of the cross. This latter percept violates expectations that are based on past experiences with flat crosses and squares.

top-down processes work together to generate an attentive consensus, or *adaptive resonance*, between what is there in the world and what we expect to see based upon our past experiences. Such a consensus typically involves all the processes of adaptive filtering, perceptual grouping, and attentive matching of learned expectations with bottom-up data. This Adaptive Resonance Theory, or ART, also explains how we can rapidly learn about a changing world throughout life, and become conscious of events as we do so (Carpenter and Grossberg, 1987, 1991; Grossberg, 1976, 1980, 2013, 2017). Helmholtz and Kanizsa could not fully make these connections because they did not have the critical intuitive and mathematical concepts needed to express them clearly. ART concepts will be summarized in Sect. 2.3 and beyond as one of the tools in our analysis. To get started, key bottom-up processes first need to be reviewed.

2. Theoretical Introduction

This section reviews some of the basic neural principles and mechanisms that will be used to discuss paintings by several artists. Section 3 will build upon this introduction to provide additional information that will also be needed. Section 4 will apply this information to analyze paintings by Jo Baer, Banksy, Ross Bleckner, Gene Davis, Charles Hawthorne, Henry Hensche, Henri Matisse, Claude Monet, Jules Olitski, and Frank Stella, whose paintings benefit from different combinations of visual properties.

2.1. Boundary Completion and Surface Filling-In

The functional units of visual perception are 3D boundaries and surfaces, or more exactly, 3D representations of completed boundary groupings and of

Downloaded from Brill.com04/06/2021 01:46:22AM
via free access

Figure 3. Dalmatian in Snow. The initial percept of black splotches on white paper gradually becomes recognizable as a Dalmatian in snow as amodal boundaries get completed between the splotches of the Dalmatian's body.

filled-in surfaces (Grossberg, 1987a, b, 1994). The words 'completed' and 'filled-in' refer to processes that deal with the incomplete nature of boundary and surface information that is propagated bottom-up to the brain from each retina. The nature of the incompleteness may be due to the environment, or to the structure of the retina itself.

Figure 3 shows a famous example of a Dalmatian in Snow wherein the incompleteness of boundaries comes from the visual environment. When we first look at this picture, it may just look like an array of black splotches of different sizes, densities, and orientations across the picture. Gradually, however, we can recognize the Dalmatian in it when boundary groupings form in our brain between the black splotches. These emergent boundaries are visual illusions that are not in the image itself. They are created in the visual cortex. Despite being illusory, however, these boundary groupings are very useful in enabling us to recognize the dog. This is particularly remarkable because the emergent boundaries are perceptually invisible, or amodal; they are not lighter or darker than the white background, nor a different color, nor perceived at a different depth. This percept illustrates that we can *consciously recognize invisible boundaries*. Both invisible and visible boundary groupings help us to group image fragments into object boundaries that are sufficient for object recognition.

Downloaded from Brill.com04/06/2021 01:46:22AM
via free access

The Dalmatian in Snow example illustrates that *some* boundaries are invisible. Grossberg (1984, 1987a, b, 1994, 1997) predicted, moreover, that *"all boundaries are invisible"*, at least within the boundary formation stream of the visual cortex. Some other examples of this boundary property are provided in Fig. 4. Figure 4 (upper row) includes a Kanizsa square (left panel) and a reverse-contrast Kanizsa square (right panel).

The Kanizsa square is a visual illusion that is induced by boundary completion between pairs of collinear pacman edges. The square is visible because each of the four black pacmen induces bright contrasts within the square. These contrasts can then spread, or fill-in, surface brightness or color within the illusory square boundary until they hit either the real pacman boundaries or the illusory square boundaries. This flow of brightness and color behaves much like a fluid that diffuses away from a source. Boundaries act like a dam

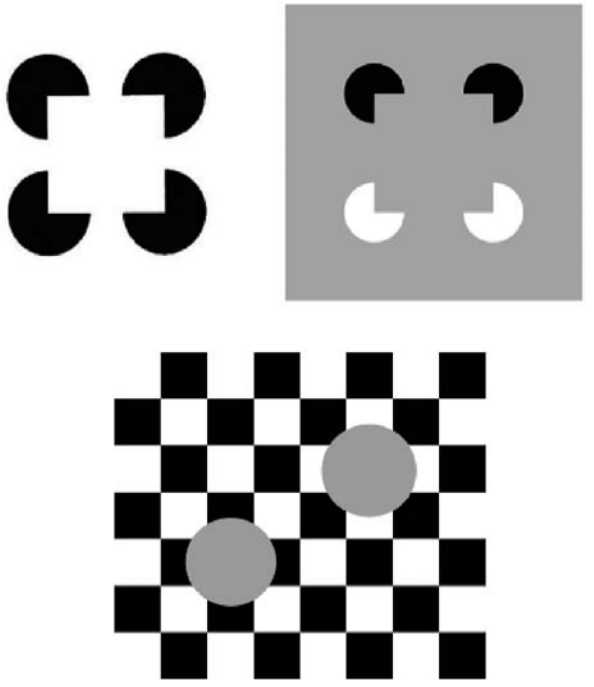


Figure 4. (Left, upper row) In the percept of a Kanizsa square, the illusory boundary square is rendered visible due to the enhanced brightness inside the square relative to the background. See the computer simulation in Fig. 6. (Right, upper row) The boundary of the reverse-contrast Kanizsa square can be recognized without being seen because the dark and light inducers due to the pacman figures generate a gray color during surface filling-in. See the computer simulation in Fig. 7. (Lower row) As the circumference of the gray disk is traversed, the polarity with respect to the background repeatedly reverses between dark–light and light–dark contrasts as black and white regions of the background are traversed. Complex cells in V1 can nonetheless respond at every position of the circumference because they pool opposite contrast polarities (i.e., are 'insensitive to contrast polarity') at every position.

Downloaded from Brill.com04/06/2021 01:46:22AM
via free access

that contains the flow of brightness and color, and keeps it from flowing outside the contours of the boundary. By acting like barriers, or obstructions, to the flow of brightness and color, boundaries can make themselves visible by causing a different brightness or color to occur on opposite sides of the boundary. Both types of boundaries, whether induced directly by pacman image contrasts or completed between pairs of pacmen, act as barriers to further spreading of brightness. After filling-in is complete, the interior of the square appears brighter than the background around it.

The reverse-contrast Kanizsa square is also a visual illusion that can be consciously recognized. As with the Kanizsa square, it is induced by boundary completion between pairs of collinear pacman edges. However, each edge in such a pair has an opposite contrast (dark/light or light/dark) relative to the gray background. Despite having opposite contrasts, these pacmen can still induce illusory contours between them. This can happen because, in response to like-oriented boundary signals at each position, the brain adds both dark/light and light/dark boundary inputs at a cell type that can respond to both polarities. This is accomplished in two processing stages. The first stage uses orientationally-tuned cells at each position whose receptive fields are sensitive to just one polarity, either dark/light or light/dark, but not both. These cells are called *simple cells* (Figure 5a). At each position, simple cells with similar

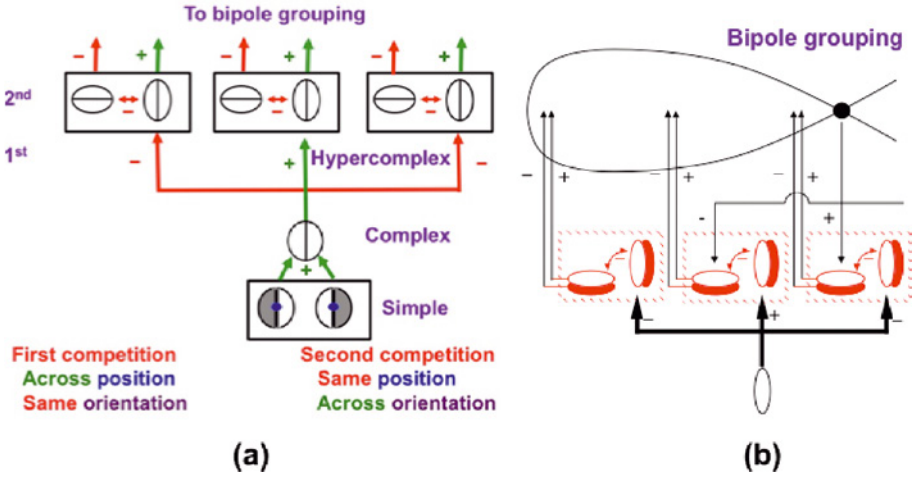


Figure 5. (Left) Oppositely polarized, but like oriented, simple cells at each position add their inputs at a complex cell. The complex cell, in turn, inputs to the first and second competitive stages. The first stage creates hypercomplex cells using an on-center off-surround network across position and within orientation. The second stage uses a push-pull tonically-active competition within position and across orientation. Together, the two competitive stages can create end cuts at line ends and other high curvature positions on contours. (Right) On-cells at the second competitive stage excite bipole cells with the same orientational preference. Off-cells (in solid red), that are transiently turned on when the on-cells shut off, inhibit the bipole cell.

Downloaded from Brill.com04/06/2021 01:46:22AM
via free access

orientational tuning, but opposite contrast polarity preference, add their inputs at the next processing stage. The cells that add these inputs are called *complex cells* (Figure 5a). Both simple and complex cells occur in the first region of visual cortical processing, which is called V1. Hubel and Wiesel (1968) received the Nobel prize for their discovery of simple and complex cells. Boundary completion occurs at a later processing stage than the complex cells. Because boundary completion cells receive their inputs after the stage of complex cells (Fig. 5a, b), boundary completion can occur between inputs with opposite contrast polarities, as in the reverse-contrast Kanizsa square.

Unlike the percept of the Kanizsa square, the interior brightness of the reverse-contrast Kanizsa square percept does not look significantly different from that outside the square. This is because the two white pacmen induce enhanced darkness within the square, whereas the two black pacmen induce enhanced brightness. When these opposing contrasts fill-in within the interior of the square, they tend to average across space and cancel out.

Figures 6 and 7 show computer simulations of how these percepts can be generated. These simulations were published in Gove *et al*. (1995). Figure 6b

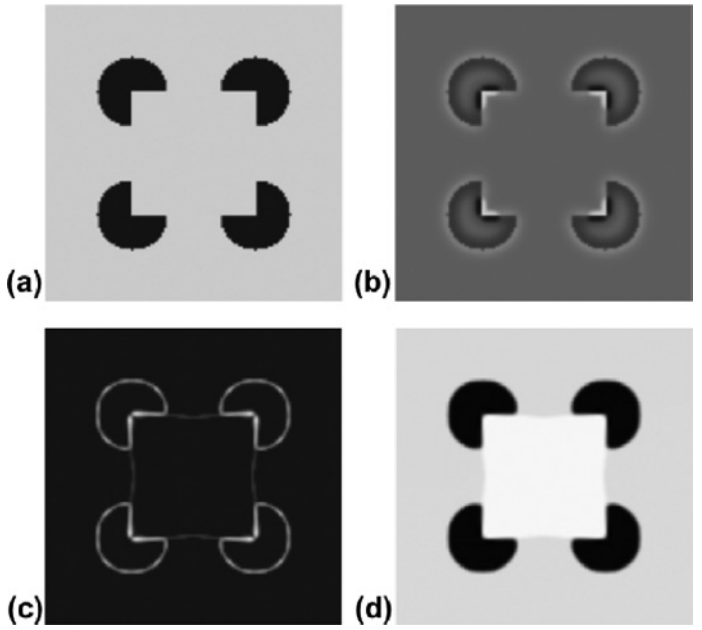


Figure 6. Simulation of a Kanizsa square percept. (Left, upper row) Kanizsa square input image. (Right, upper row) Feature contours computed via discounting of the illuminant. Note the 'brightness buttons' inside each of the four pacman figures. (Left, lower row) Completed boundaries. (Right, lower row) Filled-in surface brightnesses. Note that the interior of the square is brighter than its background due to the filling-in of the four brightness buttons within the square. (Reprinted with permission from Gove *et al*., 1995).

Downloaded from Brill.com04/06/2021 01:46:22AM
via free access

shows the *feature contours* that are induced just inside the pacman boundaries. These feature contours are the regions of enhanced brightness that are due to the black pacmen. The feature contours fill-in within the square boundary that is created by boundary completion (Fig. 6c) to create a percept of enhanced brightness throughout the square surface, as shown in Fig. 6d.

A simulation of the percept in response to the reverse-contrast Kanizsa square is shown in Fig. 7. Whereas bright feature contours are induced just inside the boundaries of the two black pacmen at the bottom of the figure, as in Fig. 6b, dark feature contours are induced just inside the boundaries of the two white pacmen at the top of the figure (Fig. 7b). Because these dark and bright feature contours are approximately balanced, the filled-in surface color inside the square is indistinguishable from the filled-in surface color outside of the square (Fig. 7d), so that the square boundary is recognized but not seen.

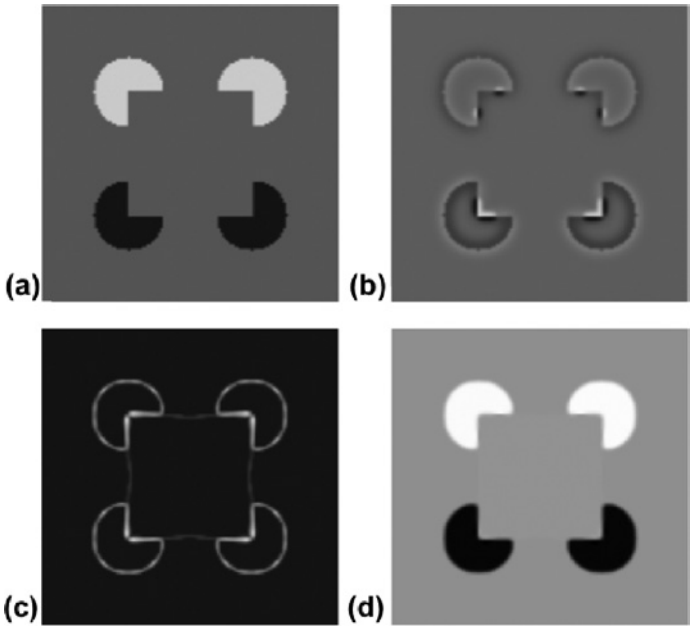


Figure 7. Simulation of a reverse-contrast Kanizsa square percept. (Left, upper row) Reverse-contrast Kanizsa square input image. (Right, upper row) Feature contours computed via discounting of the illuminant. Note the 'brightness buttons' inside each of the two black pacman figures at the bottom of the image, and the two 'darkness buttons' inside each of the two white pacman figures at the top of the image. (Left, lower row) Completed boundaries. (Right, lower row) Filled-in surface brightnesses. Note that the interior of the square has the same brightness as its background because the bright and dark buttons average out during filling-in to create an intermediate level of brightness. (Reprinted with permission from Gove *et al*., 1995).

Downloaded from Brill.com04/06/2021 01:46:22AM
via free access

The reverse-contrast Kanizsa square percept thus provides another example of invisible boundaries. These boundaries are invisible for two different kinds of reasons: To begin, because complex cells pool opposite contrast polarities at each position, they cannot distinguish between dark/light and light/dark contrasts; hence *"all boundaries are invisible"* within the boundary formation system. Indeed, complex cells are known to pool inputs from opposite polarities of both achromatic and chromatic inputs (e.g., Thorell *et al*., 1984). Secondly, in this particular example, the effects of the brightness inducers tend to cancel out after surface filling-in occurs, so the surface percept does not distinguish the brightness inside the square from that outside the square. As a result, in response to a properly balanced reverse-contrast Kanizsa square stimulus, surface brightness cannot demarcate the positions of the invisible boundaries using a visible filled-in brightness difference across the boundary. The salience, or strength, of the boundary signals can, however, do so, even if the boundaries cannot be consciously seen.

It should be noted that, although neurophysiologists like Thorell *et al.* (1984) reported experimentally how complex cells pool signals from different simple cell polarities, they believed that these data implied that complex cells "must surely be considered color cells in the broadest sense" (p. 768). In contrast, Grossberg (1984) predicted that "all boundaries are invisible" and that complex cells are amodal boundary detectors that pool together signals from multiple simple cell detectors in order to build the best possible boundary signals. This conclusion is derived from a theoretical analysis of how boundary completion and surface filling-in interact to generate consciously visible percepts.

The image in Fig. 4 (lower row) illustrates an important consequence of the fact that opposite polarity boundary signals are added by complex cells at each position. The gray disks in this image lie in front of a black and white checkerboard. As the circumference of each gray disk is traversed, the relative contrasts reverse, from black–white to white–black and back again. Because complex cells exist, a boundary can form around the entire circumference of the gray disk. If, instead, the brain computed separate black–white or white–black boundaries using only simple cells, these boundaries would have four big holes in them through which brightness could easily flow. Complex cells prevent this perceptual calamity from occurring. But they do so only at the cost that "all boundaries are invisible": Because they add signals from opposite contrast polarities at each position, they cannot signal whether the contrast goes from light-to-dark or dark-to-light.

The image in Fig. 4 (lower row) was yet another example that Kanizsa introduced to argue against Helmholtz's position that all seeing is based on a type of knowledge-based hypothesis testing, or unconscious inference. Kanizsa noted that our experiences with regular black-and-white checkerboards should

Downloaded from Brill.com04/06/2021 01:46:22AM
via free access

lead us to expect that a white square is occluded by the gray disk in the lower left of the image, and a black square is occluded by the gray disk in the upper right of the image. Instead, there is a strong percept of an amodally completed black cross behind the gray disk in the lower left, and of a white cross behind the gray disk in the upper right, again contradicting a purely top-down explanation of what observers see.

The fact that "all boundaries are invisible" has strongly influenced the painterly techniques and theories of many famous artists, as was illustrated in Grossberg (2008). More examples of its influence on the history of art are summarized herein.

Before moving on to such considerations, it is important to realize that the retina itself can create incomplete boundaries out of complete ones that are received from a visual scene. This is because each eye contains a blind spot and retinal veins (Fig. 8). The blind spot is the place on the retina where multiple pathways from photodetectors at other positions on the retina are gathered

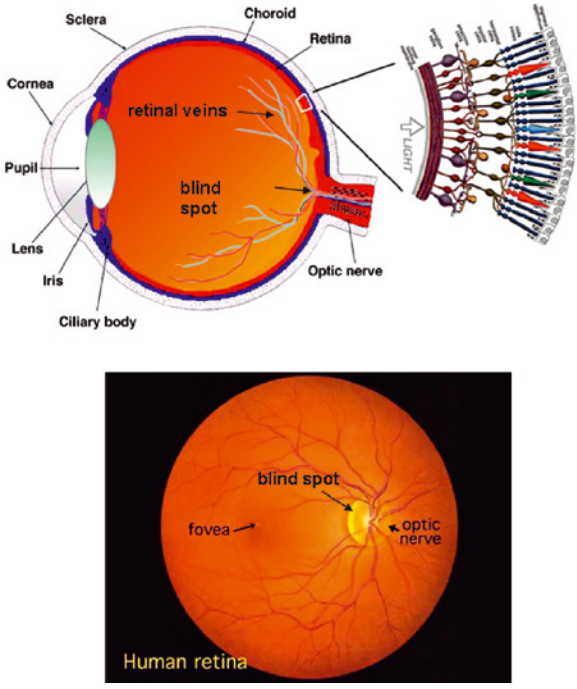


Figure 8. (Upper row) A cross section of the eye showing the blind spot and retinal veins, as well as the fact that light goes through all the retinal layers before hitting the photoreceptors. (Lower row) Top-down view of the retina, showing how big the blind spot and retinal veins are relative to the fovea, which is the high-resolution region of the retina. (Images adapted with permission from Webvision-University of Utah).

Downloaded from Brill.com04/06/2021 01:46:22AM
via free access

together to form the optic nerve. The optic nerve sends visual signals from the retina into the brain. The retinal veins cover multiple positions on the retina in order to nourish it.

The positions that are covered by the blind spot and the retinal veins cannot send visual signals to the brain from objects in the world. The brain needs to complete the boundaries in these occluded positions in the same way that it completes boundaries between the black splotches of the Dalmatian.

2.2. Boundaries and Surfaces Are Computationally Complementary

If "all boundaries are invisible", then how do we see anything at all, let alone a beautiful painting? Grossberg (1984, 1987a, b, 1994, 1997) predicted that *"all conscious percepts of visual qualia are surface percepts"*. To accomplish this, every visual scene is processed by two parallel processing streams in the visual cortex, as illustrated by the famous diagram in Fig. 9 due to DeYoe and Van Essen (1988). *Boundaries* are computed within the interblob cortical stream (shown in red), which passes from the retina through the lateral geniculate nucleus (LGN), onwards through the interblobs of cortical area V1 to the interstripes (or pale stripes) of cortical area V2, and up to cortical area V4. *Surfaces* are computed within the blob cortical stream (shown in green), which passes from the retina through the LGN, onwards through the blobs of

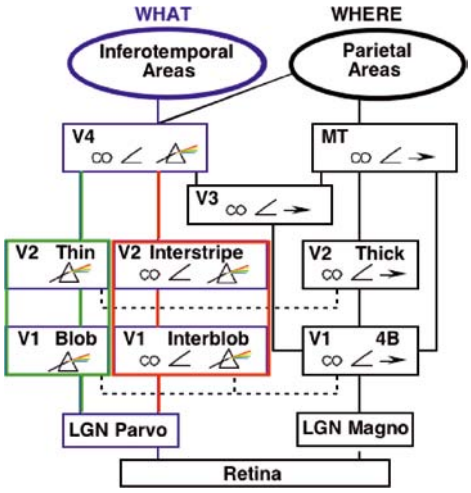
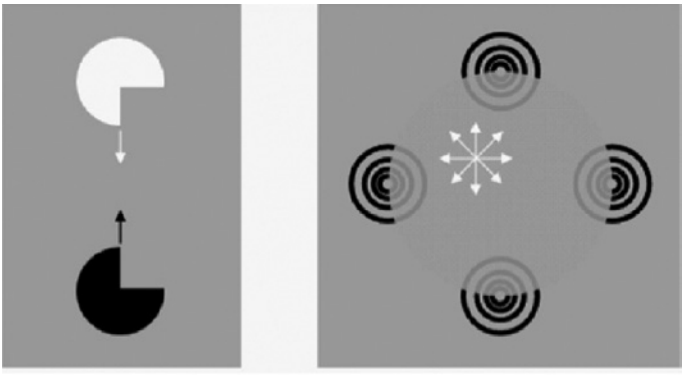


Figure 9. Diagram of anatomical connections and neuronal selectivities of early visual areas in the macaque monkey. LGN = Lateral Geniculate Nucleus; V1 = striate visual cortex; V2, V3, V4, MT = prestriate cortical areas. The surface stream goes through the blobs and thin stripes (in the green box) to cortical area V4 and inferotemporal areas. The boundary stream goes through interblobs and interstripes (in the red box) to V4. The motion stream goes through V1 and MT to the parietal areas. (Adapted with permission from De Yoe and Van Essen, 1988).

Downloaded from Brill.com04/06/2021 01:46:22AM
via free access

the cortical area V1 to the thin stripes of cortical area V2, and up to cortical area V4.

Why are two parallel cortical streams needed? This is true because boundary completion and surface filling-in have been demonstrated by neural models and supportive data to obey *computationally complementary* properties (Grossberg, 1984, 1994, 2000, in press). As summarized in Figure 10 and illustrated by the percepts generated by the images in Figs 2–4, boundaries form *inwardly* between pairs or greater numbers of inducers, are *oriented*, and are *insensitive* to direction-of-contrast. This last property is just another way of saying that the boundary system pools inputs at complex cells from opposite contrast polarity simple cells at each position (Fig. 5a). In contrast, surfaces fill-in *outwardly* from individual inducers in an *unoriented* way, and are *sensitive* to direction-of-contrast. This last property is just another way of saying that the surface stream can use contrast differences to generate conscious percepts of visual qualia. These three pairs of boundary and surface properties (outward vs. inward, unoriented vs. oriented, sensitive vs. insensitive to direction-of-contrast) are clearly complementary. In particular, the properties needed to complete a boundary cannot be used to fill-in a surface, and conversely. Interactions between both streams are needed to generate visual percepts.



Complementary Properties of Boundaries and Surfaces

Boundary Completion	**Surface Filling-in**
Inward	Outward
Oriented	Unoriented
Insensitive to direction-of-contrast	Sensitive to direction-of-contrast

Figure 10. Complementary properties of boundary completion and surface filling-in. See text for details.

Downloaded from Brill.com04/06/2021 01:46:22AM
via free access

2.3. Discounting the Illuminant

We have already seen that surface filling-in is needed to compensate both for retinal occlusions, like the blind spot and retinal veins, and for incomplete scenic information, like the Dalmatian in Snow. There is another reason why filling-in is needed, and that is because we see the world when it is illuminated by many different light sources that change their color and intensity both across a scene and throughout the day. Somehow the brain needs to compensate for this variability in lighting, or else it would confuse variable illumination information with unvarying properties of object shape and color. This compensatory process is called *discounting the illuminant*. Even Helmholtz was aware of the critical role of this process, but he did not understand how it worked. Later scientists like Edwin Land and his colleagues introduced models, such as the Retinex model, to better characterize its properties; e.g., Land (1964, 1977) and Land and McCann (1971). However, a computational neural model of how filling-in works in concert with boundary completion in response to both 2D pictures and 3D scenes was accomplished only later, by the FACADE and 3D LAMINART models (e.g., Bhatt *et al*., 2007; Cao and Grossberg, 2005; Fang and Grossberg, 2009; Grossberg, 1987a, b, 1994, 1997; Grossberg *et al.*, 2007; Grossberg and McLoughlin, 1997; Grossberg and Mingolla, 1985a; Grossberg and Swaminathan, 2004).

In particular, in order to eliminate (most of) the variability of illumination across a scene that could cause wildly unstable shape and color percepts, *feature contours* are computed at positions where luminance or color contrasts change rapidly enough across space. Such positions often occur along a surface's boundaries, also called *boundary contours* to distinguish them from the feature contours (Fig. 11, upper row). After discounting of the illuminant takes place, the surviving feature contours compute brightness and color signals that are relatively uncontaminated by varying illumination levels. They can do this because the contrast changes where they are computed are due primarily to changes in the material properties, called *reflectances*, of the underlying objects, whereas the illumination level changes little, if at all, across such a contrast change. The illuminant-discounted feature contour signals can then, at a subsequent processing stage, trigger filling-in of their brightnesses or colors across the surface until they hit the boundary contours that enclose the surface (Fig. 11, left, lower row).

Boundaries are computed at multiple depths, and trigger filling-in of colors and brightnesses only at those depths where the feature contours and boundary contours are parallel and adjacent to one another (Grossberg, 1994; Fig. 12). Filling-in occurs in networks within the surface stream that are called Filling-In-DOmains, or FIDOs. There are multiple FIDOs to enable filling-in of multiple opponent colors (red–green, blue–yellow) and achromatic

Downloaded from Brill.com04/06/2021 01:46:22AM
via free access

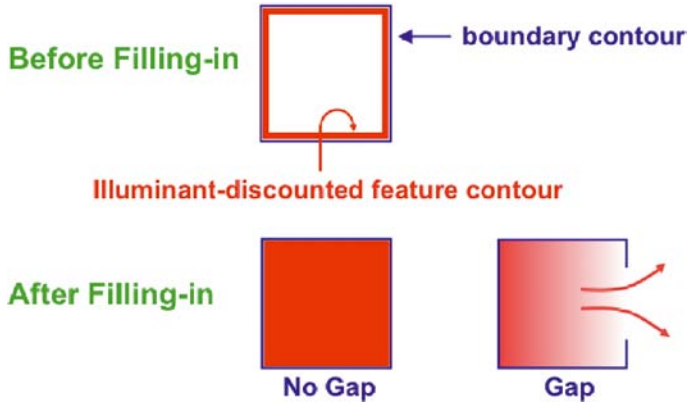


Figure 11. (Upper row) After discounting of the illuminant in response to a uniformly illuminated red square, a closed boundary (in blue) surrounds the square at the depth that it is perceived. Just within the boundary, feature contours (in red) code the surface color with the illuminant discounted. (Left, lower row) After surface filling-in occurs, the feature contours fill in the entire square boundary, and are prevented from spreading beyond it. (Right, lower row) If the boundary has a big enough break in it, then surface color can spread outside in during filling-in, thereby creating similar contrasts both inside and outside the boundary.

brightnesses (light–dark) at multiple depths. The three pairs of opponent FIDOs at every depth in Fig. 12 schematize this property.

Many painters have struggled to represent the difference between the real colors of objects and the illumination that is reflected from them and onto other objects. Their success in doing so is all the more remarkable considering that they see the world only after the illuminant has been discounted, which is also true of us as we look at their paintings.

2.4. Gist as a Coarse Boundary/Surface Texture Category

Given that boundaries and surfaces are the functional units of visual perception, how is gist perceived? This is an active area of research in both biological vision and computer vision. Grossberg and Huang (2009) proposed a simple biological answer to this question in their ARTSCENE model; in the same paper they also review other scene classification models and review comparative benchmark simulations of these models. The ARTSCENE model classifies the gist of a scene as a *texture category* that is derived from coarse boundary and surface representations that are derived from the scene. This texture representation was input to a network that could learn to categorize, or classify, textures into different scene types; e.g., coast, forest, mountain, countryside. The same kind of categorizing network can be used to classify both textures and objects. It will be reviewed in the next section.

Downloaded from Brill.com04/06/2021 01:46:22AM
via free access

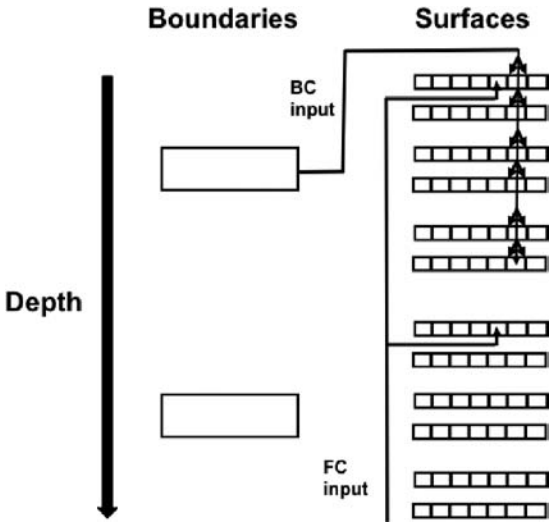


Figure 12. Multiple depth-selective boundary contours (BC) are sensitive to different ranges of receptive field sizes, or spatial scales. Each BC boundary can send topographic signals to the corresponding depth-selective surface Filling-In DOmain, or FIDO. These signals can trigger filling-in of feature contour (FC) signals that are positionally aligned with them within their BC. The BC signals reach all three double-opponent FIDOs (see three pairs of FIDOs) at their depth. The FC signals reach all depth-selective FIDOs, but are captured only at depths where there are BC signals that are positionally aligned with them. These BC signals act both as filling-in generators and filling-in barriers. BC signals can also influence FIDOs with nearby depth preferences, but less strongly, thereby enabling continuous perception of depth to occur in response to a finite number of BC and FC representations. These nearby BC-to-FC connections are not shown, for simplicity.

Grossberg and Huang (2009) further proposed how, after a glimpse of a scene enables its gist to be classified, the eyes may wander around the scene to accumulate additional information about it. ARTSCENE assumes that the eyes look at bigger regions of a scene with higher probability, other things being equal. For example, the largest region in a coast scene may be the ocean, so the eyes may look at the ocean with higher probability. ARTSCENE explains how spatial attention could select such a region from the rest of the scene and, after doing so, learn to classify that texture with its own recognition category.

Before going on, more should be said about what recognition categories do. A recognition category responds selectively to the distributed features of a scene or painting, and not to very different scenes or paintings. Typically, a single population of cells, or small number of such populations, defines such a category, much smaller than the number of cells in the scene or painting that they represent. In this sense, a category ‘compresses’ the scene or painting into a selective recognition event. It is a ‘symbol’ of the scene or painting. Whereas

Downloaded from Brill.com04/06/2021 01:46:22AM
via free access

individual pixels of a scene or painting have no meaning, the selective activation of a category by such a distributed input provides a context-sensitive internal index that a particular scene or painting is being viewed.

In the ARTSCENE model, after the gist of a scene is categorized, the model's attentional focus shifts two or three more times, thereby allowing spatial attention to focus upon a few of the biggest textures of the scene; e.g., ocean, sky, grass, trees, etc. These texture categories are also categorized, and all of their recognition decisions are allowed to 'vote' for the scenic category. This simple voting procedure generates a more accurate prediction than gist alone, indeed a prediction that achieves human levels of gist recognition, and that did better than other, more complicated, models for computing gist.

In summary, one of the first things that a viewer of a painting may notice is the gist of the painting. This may include information that it is a painting of a certain kind of scene, or that it is a painting by a particular artist. Section 4.13 notes that some artists, such as Claude Monet, created their paintings with this insight in mind.

2.5. Texture and Object Learning, Recognition, and Prediction by Adaptive Resonance

In order to recognize the contents in a painting, one needs to be able to recognize the textures and the objects of which it is composed, even if these objects are just simple lines and curves. Learning about any new texture or object requires solving the *stability–plasticity dilemma* (Grossberg, 1980). The stability–plasticity dilemma asks how the brain is able to learn quickly about new objects and events without just as quickly forgetting previously learned, but still useful, memories. Adaptive Resonance Theory, or ART, proposed how this problem can be solved (Carpenter and Grossberg, 1987, 1991; Grossberg, 1976, 1980, 2013). ART is currently the most advanced cognitive and neural theory in the specific sense that it has explained and predicted the most psychological and neurobiological data, and each of its foundational principles and mechanisms has been supported by such data, as reviewed in Grossberg (2013).

To solve the stability–plasticity dilemma, ART uses matching between bottom-up input patterns and learned top-down expectations at feature-selective networks of cells (Fig. 13). A top-down expectation is released by a currently active recognition category. Such expectations may be thought of as predictions that the brain releases in particular contexts. These predictions are matched against incoming data from the world.

In the absence of bottom-up inputs, top-down expectations usually only prime, sensitize, or modulate their targets cells with the object prototype that the top-down expectation has learned. The concept of priming can be understood by the example of asking a viewer of a painting to look for a particular

Downloaded from Brill.com04/06/2021 01:46:22AM
via free access

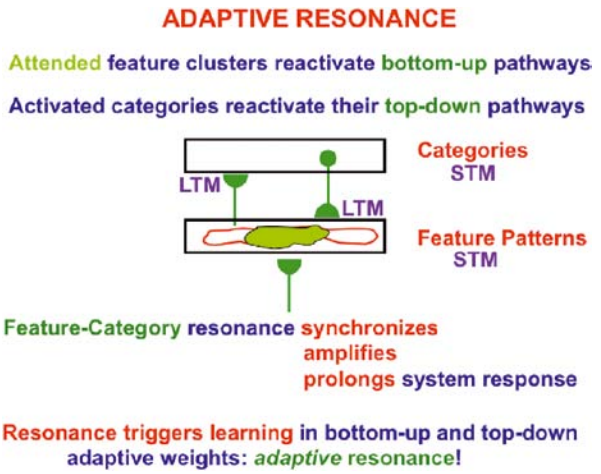


Figure 13. A feature–category resonance is triggered when bottom-up signals from a distributed feature pattern are sufficiently well matched with top-down signals from an activated category. This resonance synchronizes, amplifies, and prolongs the responses of the active cells—both the attended feature pattern and the active category—and can trigger learning in the adaptive weights, or long-term memory (LTM) traces, within the bottom-up adaptive filter pathways from feature pattern to category, and within the top-down expectation pathways from category to feature pattern. STM = Short-term memory.

object in it. If the viewer activates an expectation of seeing the object, the expectation, by itself, does not usually generate a conscious percept of the object. That would be akin to hallucinating the object. Rather, when the viewer's eyes happen to look at or near the object, the object will be more quickly recognized than if the expectation were not active. This happens because, when a top-down expectation is active, a bottom-up input pattern that matches the expectation at feature-selective cells can more rapidly and vigorously activate these cells to suprathreshold values than in the absence of the expectation.

How is such a top-down expectation learned? Before any top-down learning of the expectation has occurred, *all* top-down adaptive weights, or long-term memory (LTM) traces, within the pathways from the category level to the feature are large and uniformly distributed. These uniformly distributed large adaptive weights enable the top-down LTM traces of the category that is chosen on the first learning trial to match whatever feature pattern activated it, and thus to allow learning to begin. Prototype learning on subsequent learning trials *prunes* these weights to match the *critical feature pattern* (marked by the light green region at the feature level in Fig. 13) that is learned from the sequence of all the input patterns that can activate the category.

Learning occurs within an ART system when a good enough match occurs between a bottom-up input pattern and a top-down expectation on a given

Downloaded from Brill.com04/06/2021 01:46:22AM
via free access

learning trial. A good enough match triggers a context-sensitive *resonance* between the active cells in the feature and category levels. Such a resonance is supported by the mutual exchange of bottom-up excitatory signals within the adaptive pathways from active feature cells to the active category, and of top-down excitatory signals within the adaptive pathways from the active category cells to the active feature cells (Fig. 13). The resonance prolongs, synchronizes, and amplifies the activities of both the attended critical feature pattern and the active category; hence the name *feature–category resonance* for this dynamical state. A feature–category resonance triggers fast learning by the adaptive weights in both the bottom-up and top-down pathways that join these representations; hence the name *Adaptive* Resonance Theory. By also inhibiting irrelevant outlier feature cells that do not resonate during this match-based learning process, ART also solves the stability–plasticity dilemma.

2.6. ART Matching Rule: Attention, Biased Competition, and Predictive Coding

Attention within such an ART circuit obeys the ART Matching Rule. This Rule is embodied by top-down, modulatory on-center, off-surround circuits that read out the learned expectations to be matched against bottom-up input patterns (Fig. 14). This matching process has the following properties: When a bottom-up input pattern is received at a processing stage, it can activate its target cells, if nothing else is happening. When a top-down expectation pattern is received at this stage, it can provide excitatory modulatory, or priming, signals to cells in its on-center, and inhibitory signals to cells in its off-surround. The on-center is modulatory because the off-surround also inhibits the on-center cells (Fig. 14), and these excitatory and inhibitory inputs are approximately balanced ('one-against-one'). When a bottom-up input pattern and a top-down expectation are both active, cells that receive both bottom-up excitatory inputs and top-down excitatory priming signals can fire ('two-against-one') with amplified and synchronized activities, while other cells are inhibited (Fig. 14b). In this way, only cells can fire whose features are 'expected' by the top-down expectation, and an attentional focus starts to form at these cells.

Many psychological, anatomical, and neurophysiological experiments have provided support for the ART Matching Rule prediction of how attention works. These data support the predicted modulatory on-center, off-surround interactions; excitatory priming of features in the on-center; suppression of features in the off-surround; and amplification of matched data (e.g., Bullier *et al*., 1996; Caputo and Guerra, 1998; Downing, 1988; Hupé *et al*., 1997; Mounts, 2000; Reynolds *et al*., 1999; Sillito *et al*., 1994; Somers *et al*., 1999; Steinman *et al*., 1995; Vanduffell *et al*., 2000).

Downloaded from Brill.com04/06/2021 01:46:22AM
via free access

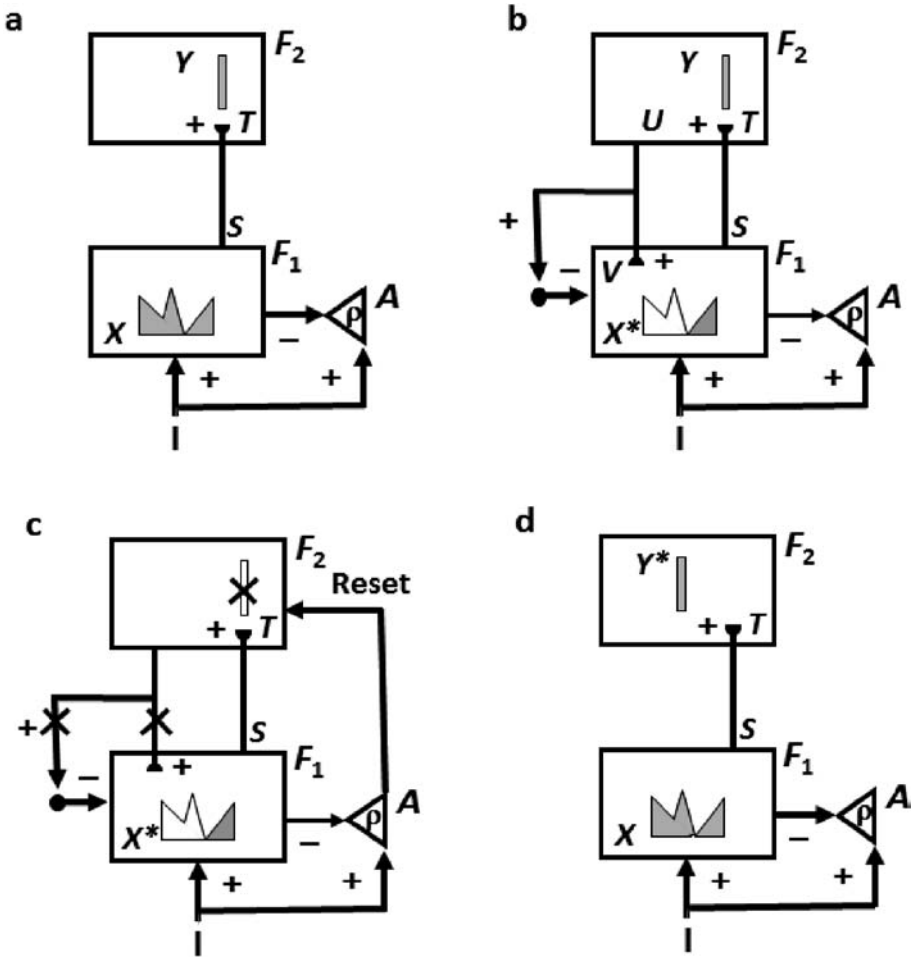


Figure 14. How an ART model searches for, learns, and recognizes a new recognition category using cycles of match-induced resonance and mismatch-induced reset. Active cells are shaded gray; inhibited cells are not shaded. (a) Input pattern I is instated across feature detectors at level F_1 as an activity pattern X, at the same time that it generates excitatory signals to the orienting system A with a gain ρ that is called the vigilance parameter. Activity pattern X generates inhibitory signals to the orienting system A as it generates a bottom-up input pattern S through the adaptive filter to the category level F_2. A dynamic balance within A between excitatory inputs from I and inhibitory inputs from S keeps A quiet. The bottom-up signals in S are multiplied by learned adaptive weights to form the input pattern T to F_2. The inputs T are contrast-enhanced and normalized within F_2 by recurrent lateral inhibitory signals that obey the membrane equations of neurophysiology, otherwise called shunting interactions. This competition leads to selection and activation of a small number of cells within F_2 that receive the largest inputs. In this figure, a winner-take-all category is chosen, represented by a single cell (or cell population). The chosen cells represent the category Y that codes for the feature pattern at F_1. (b) The category activity Y generates top-down signals U that are multiplied by adaptive weights to form a prototype, or critical feature pattern, V that encodes

Downloaded from Brill.com04/06/2021 01:46:22AM
via free access

The way in which the off-surround competition biases attention using the ART Matching Rule has led to the term 'biased competition' used by various experimental neurophysiologists (Desimone, 1998; Kastner and Ungerleider, 2001). The property of the ART Matching Rule that bottom-up signals may be enhanced when matched by top-down signals is supported by many neurophysiological experiments that have reported the facilitatory effect of attentional feedback (Luck *et al*., 1997; Roelfsema *et al*., 1998; Sillito *et al*., 1994).

Various other data have supported ART predictions about how processes of consciousness, learning, expectation, attention, resonance, and synchrony are related in the brain. For example, attention and learning are linked in a manner consistent with ART predictions about visual perceptual learning (e.g., Ahissar and Hochstein 1993, 1997; Ito *et al*., 1998; Lu and Dosher, 2004), auditory learning (e.g., Gao and Suga, 1998), and somatosensory learning (e.g., Krupa *et al*., 1999 and Parker and Dostrovsky, 1999). ART has also predicted links between attention and synchronous oscillations that were subsequently experimentally reported (e.g., Buschman and Miller, 2007; Engel *et al*., 2001; Gregoriou *et al*., 2009; Grossberg, 2009; Pollen, 1999), and between synchronous oscillations and consciousness (e.g., Lamme, 2006; Llinas *et al*. 1998; Singer, 1998).

The above citations are just a subset from a vast experimental literature. All the predictions of ART have analogous levels of experimental support. Illustrative experiments will be cited throughout the article, as appropriate.

Figure 14 (Continued).
the expectation that the active F_2 category has learned for what feature pattern to expect at F_1. This top-down expectation input V creates an excitatory modulatory signal at F_1 cells in its on-center, at the same time that it inhibits F_1 cells in its off-surround (ART Matching Rule). If V mismatches I at F_1, then a new STM activity pattern X^* (the gray pattern), is selected at cells where the patterns match well enough. In other words, X^* is active at I features that are confirmed by V. Mismatched features (white area) are inhibited. When X changes to X^*, total inhibition decreases from F_1 to A. (c) If inhibition decreases sufficiently, A releases a nonspecific arousal burst to F_2; that is, 'novel events are arousing'. Within the orienting system A, the vigilance parameter ρ determines how bad a match will be tolerated before a burst of nonspecific arousal is triggered. This arousal burst triggers a memory search for a better-matching category, as follows: Arousal resets F_2 by inhibiting Y. (d) After Y is inhibited, X is reinstated and Y stays inhibited as X activates a different category, that is represented by a different activity winner-take-all category Y^*, at F_2. Search continues until a better matching, or novel, category is selected. When search ends, an attentive resonance triggers learning of the attended data in adaptive weights within both the bottom-up and top-down pathways (see Fig. 13). As learning stabilizes, inputs I can activate their globally best-matching categories directly through the adaptive filter, without activating the orienting system. (Adapted with permission from Carpenter and Grossberg, 1987).

Downloaded from Brill.com04/06/2021 01:46:22AM
via free access

Before leaving the topic of how expectations embody predictions about what may or may not happen next in the world, it is important to realize that not all theories of predictive coding are supported by the experiments that confirm the ART Matching Rule. In particular, these data do not support Bayesian 'explaining away' models in which matches with top-down feedback causes only suppression, leaving only the unmatched data, or 'prediction error', for further processing (Mumford, 1992; Rao and Ballard, 1999). Such inhibitory matching does occur in the brain, but it typically controls spatial and motor representations (see Grossberg, 2013, for a review). There are fundamental reasons for this difference in matching schemes between perceptual–cognitive ART-like matching and spatial-motor inhibitory matching, also called VAM-like matching. One of the most important reasons is that only the ART Matching Rule solves the stability–plasticity dilemma. See Carpenter and Grossberg (1987) for a mathematical proof.

2.7. Symbol Grounding and Complementary Attentional and Orienting Systems

The resonant state combines computationally *complementary* properties of the network into a context-sensitive representation that can be the basis of an object's meaning. In particular, individual pixels or features of a painting or scene are meaningless. Pixels and features become meaningful only as part of the spatial context of the other pixels and features that surround them. Activating the category of a painting, or scene, provides context-selectivity because the category responds selectively to specific combinations of features which define that scene, but not to those that characterize different paintings. However, the category itself has no knowledge of what feature patterns cause it to fire, or indeed any information about the object or painting that activates it. The resonant state binds together the distributed features of the painting with its category. This *bound state* represents both the identity of the painting and the features that define it. In this way, a feature–category resonance provides a solution of the symbol grounding problem (Harnad, 1990).

Category learning in ART is controlled by interactions between an attentional system and an orienting system (Fig. 14). These two systems embody yet another type of computationally complementary laws (Grossberg, 1980, 2000). The attentional system (levels F_1 and F_2 in Fig. 14) carries out processes like attention, category learning, expectation, and resonance. The orienting system (level A in Fig. 14) enables the attentional system to learn about novel information using processes like reset, memory search, and hypothesis testing. The attentional system includes brain regions like temporal cortex and prefrontal cortex. The orienting system includes brain regions like the nonspecific thalamus and hippocampus.

Downloaded from Brill.com04/06/2021 01:46:22AM
via free access

If an input pattern causes a sufficiently bad mismatch to occur within the attentional system, it will activate the orienting system (Fig. 14c). The orienting system, in turn, resets the active category and initiates a search for a better-matching category, possibly an entirely new one (Fig. 14d).

Within the orienting system, a *vigilance parameter* ρ (see Fig. 14) determines the generality of the learned categories. If vigilance is high, then learning of a concrete or specific category occurs, such as recognition of a frontal view of a friend's face, or of the painting called *Impression, Sunrise* by Monet. If vigilance is low, then learning of an abstract or general category occurs, such as recognition that everyone has a face, or that one is viewing an Impressionist painting. In general, vigilance is chosen as low as possible to conserve memory resources, without causing a reduction in predictive success. Because the baseline level of vigilance is initially set at the lowest level that led to predictive success in the past, ART models try to learn the most general category that is consistent with the data. This tendency can lead to the type of overgeneralization that is seen in young children (Brooks *et al.,* 1999) until subsequent learning leads to category refinement (Tomasello and Herron, 1988).

When a given task requires a finer categorization, vigilance is raised. Indeed, vigilance can be automatically adjusted to learn either specific or general information in response to predictive failures, or disconfirmations, within each environment. Such a predictive failure could occur, for example, if a viewer of a painting believes that it is by Monet, whereas it is really by Sisley. Within ART, such a predictive disconfirmation causes a shift in attention to focus on a different combination of features that can successfully be used to recognize that this painting is by Sisley, and perhaps to recognize other paintings by him as well.

One type of automatic adjustment of vigilance in response to a predictive disconfirmation is the following. Suppose that on every learning trial, a predictive failure causes vigilance to increase by the smallest amount that can trigger a memory search for a new recognition category that can correct the error. As a result of such a memory search, a category will be learned that is just general enough to eliminate the error. When this happens, *match tracking* is said to occur, because vigilance tracks the degree of match between the input pattern and the expected prototype. Match tracking leads to *minimax learning*, or learning that can *minimize* predictive errors while it *maximizes* category generality. In other words, match tracking uses the minimum memory resources that are needed to correct the predictive error. Because vigilance can vary during match tracking in a manner that reflects current predictive success, recognition categories capable of encoding widely differing degrees of generalization or abstraction can be learned by a single ART system (Carpenter and Grossberg, 1987, 1991).

Downloaded from Brill.com04/06/2021 01:46:22AM
via free access

Since the concept of vigilance control was first published by Carpenter and Grossberg, quite a bit of new data has been published that supports ART predictions about how vigilance may be regulated in the brain, through brain regions like the nucleus basalis and neurotransmitters such as acetylcholine (Grossberg and Versace, 2008; Palma *et al.* 2012a, b). It has also been predicted that vigilance cannot dynamically adjust itself in some individuals, leading to problems of attention, learning, and recognition. Grossberg and Seidman (2006) have, for example, proposed that various individuals with autism have their vigilance stuck at an abnormally high value, leading to the learning of abnormally concrete and hyperspecific recognition categories, as well as to a correspondingly narrow focus of attention. It is known that there is abnormal cholinergic activity in the parietal and frontal cortices of autistic individuals that is correlated with abnormalities in the nucleus basalis (Perry *et al*., 2001), consistent with the predicted role of the nucleus basalis and ACh in regulating vigilance.

Much of the recognition learning that we do is *unsupervised.* In other words, it goes on without an explicit teacher telling us the answers. Unsupervised learning clusters different input patterns into the same category based upon the *similarity* of their feature patterns using the world itself as a teacher. However, there are many instances where *supervised* learning is needed. In other words, a teacher provides predictive answers. Here, the recognition category learns to activate the answer that the teacher provides. Supervised learning is needed to identify what type of scene one is viewing by answering with the name of the scene type; e.g., beach, forest, mountain, etc. Supervised learning is also needed to say what artist has created a particular painting, or what artistic movement it illustrates. When an ART system predicts that a painting by Sisley is really by Monet, and it is provided with the correct answer, the system can spontaneously drive a memory search leading to a new focus of attention that can better recognize the painting as a Sisley. Much experimental evidence supports the hypothesis that a similar process goes on in human observers, as reviewed in the above archival articles.

2.8. Invariant Category Learning, Recognition, and Search

The discussion in Sect. 2.4 mentioned the utility of scanning a scene or picture with eye movements in order to accumulate more information about it, such as when recognizing scene type using the ARTSCENE model. However, such scanning raises a fundamental problem. Much of our learning about objects takes place by classifying multiple views of a novel object into a view-invariant object category, without the benefit of a teacher, as our eyes freely scan a complicated scene, or picture. Why do not views from multiple objects get erroneously bound together to activate a single category? More generally, how does invariant category learning bind together different appearances of an object, not only as it is seen from the perspective of multiple

Downloaded from Brill.com04/06/2021 01:46:22AM
via free access

views, but also at multiple positions, and distances—and thus sizes—relative to our retinas? Moreover, having learned such invariant object categories, how do we use them to efficiently search a scene for a valued object, such as a much-loved picture in a museum? This latter problem is often called the Where's Waldo Problem. A solution to this suite of problems is needed every time we scan a painting from multiple perspectives.

The 3D ARTSCAN Search model (Figs 15 and 16) predicts how valued objects are learned, recognized, and searched as our eyes freely scan a 3D scene or 2D painting (Cao *et al.*, 2011; Chang *et al.*, 2014; Fazl *et al.*, 2009; Foley *et al.*, 2012; Grossberg, 2009; Grossberg *et al.*, 2011, 2014). Accomplishing this combination of basic competences requires interactions between multiple brain regions whose dynamics embody processes of spatial and object

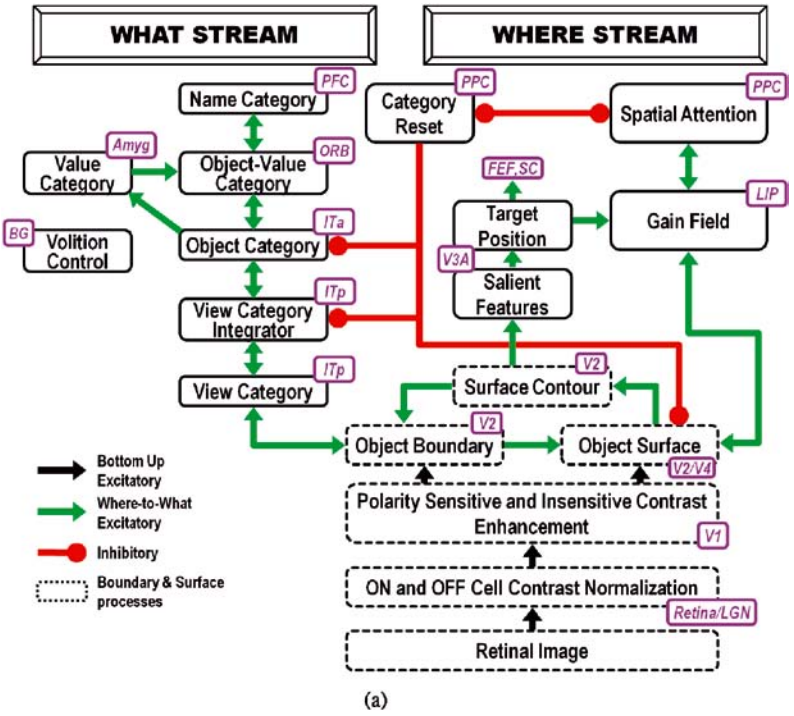


Figure 15. ARTSCAN Search macrocircuit. Invariant category learning and recognition. The dashed boxes indicate the boundary and surface processes. The green arrows represent the excitatory cortical signals from the Where cortical stream to the What cortical stream whereby invariant category learning and recognition, and reinforcement learning, occur. The red connections ending in circular disks indicate inhibitory connections. ITa: anterior part of inferotemporal cortex; ITp: posterior part of inferotemporal cortex; PPC: posterior parietal cortex; LIP: lateral intra-parietal cortex; LGN: lateral geniculate nucleus; ORB: orbitofrontal cortex; Amyg: amygdala; BG: basal ganglia; PFC: prefrontal cortex; FEF: frontal eye fields; SC: superior colliculus; V1 and V2: primary and secondary visual areas, V3 and V4: visual areas 3 and 4. (Reprinted with permission from Chang *et al*., 2014).

Downloaded from Brill.com04/06/2021 01:46:22AM
via free access

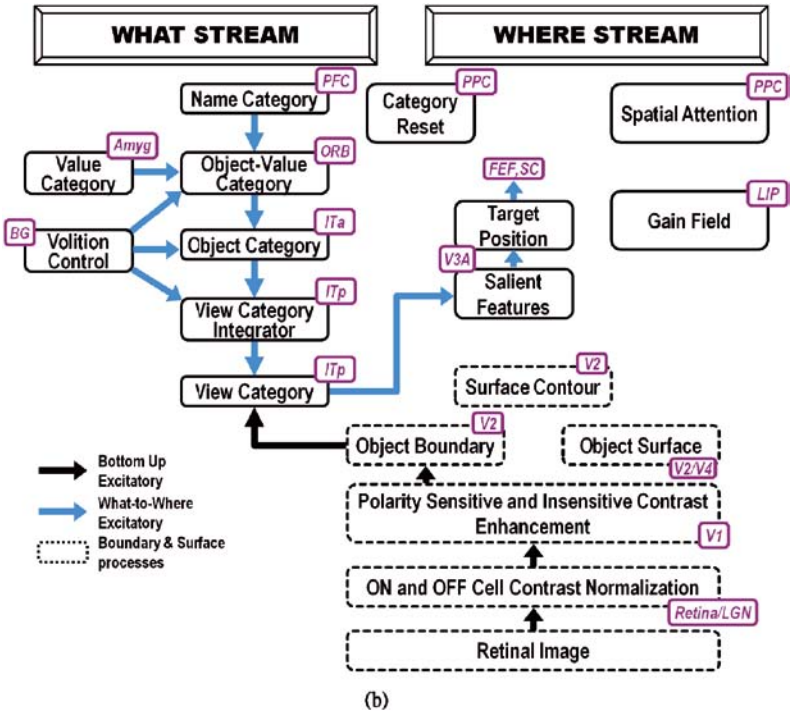


Figure 16. ARTSCAN Search macrocircuit for Where's Waldo search. Both a cognitive search driven by an object's name via the prefrontal cortex, or a value-driven search driven by the affective value of the object via the amygdala, can initiate a search for a desired object via (What cortical stream)-to-(Where cortical stream) interactions. The black arrows represent bottom-up excitatory input signals. The blue arrows represent top-down excitatory search signals. Abbreviations are the same as in the caption of Fig. 15. (Reprinted with permission from Chang *et al*., 2014).

attention, invariant object category learning, predictive remapping, reinforcement learning and motivation, and attentive visual search. 3D ARTSCAN Search proposes how spatial attention modulates invariant category learning in the manner that is reviewed in the next section. This analysis clarifies how spatial attention can focus on the different textures that make up a scene, as was discussed for the ARTSCENE model in Sect. 2.4. Within ARTSCAN Search, the ART model circuitry that was reviewed in Sects 2.5 and 2.6 is used to learn view-specific and position-specific categories that are bound together into invariant categories by the regulatory machinery of the more comprehensive 3D ARTSCAN Search circuitry.

3D ARTSCAN Search provides functional explanations and predictions of neurobiological data about interactions between multiple brain regions, including cortical areas V1, V2, V3a, V4, PPC, LIP, ITp, ITa, and PFC, as illustrated in Figs 15 and 16. 3D ARTSCAN Search is thus not just a neural model.

Downloaded from Brill.com04/06/2021 01:46:22AM
via free access

Rather, it is a neural *architecture* that tries to explain how a significant fraction of the visual brain works together to attend, learn, recognize, and search for valued objects in complicated scenes. With this kind of model in hand, it is possible to explain how a person viewing a painting from any viewpoint can understand not only its properties as a texture, as during recognition of its gist, but also what objects are represented within it.

2.9. Attentional Shrouds Modulate Invariant Category Learning

Many visual scientists have observed that spatial attention tends to fit itself to the shape of an attended object. Tyler and Kontsevich (1995) called such a form-fitting distribution of spatial attention an *attentional shroud*. The ARTSCAN model, on which the 3D ARTSCAN Search model built, proposed how an attentional shroud maintains its activity within the posterior parietal cortex, or PPC, of the dorsal, or Where, cortical stream during active scanning of an attended object's surface. ARTSCAN also predicted how shrouds regulate learning of an invariant object category in the following way (Fazl *et al.*, 2009; Grossberg, 2007, 2009).

When the eyes fixate on a particular view of an attended object, the brain can rapidly learn a view-specific category of that object using an ART circuit. To fix ideas, suppose that such a view-specific category is learned in the posterior inferotemporal cortex, or ITp (Fig. 15). The first view-specific category to be learned in ITp activates some cells in the anterior inferotemporal cortex, or ITa. These ITa cells will become an invariant object category as they are associated with multiple view-specific categories in ITp that represent different views of the object. An attentional shroud keeps the ITa cells active while multiple view-specific categories get associated with them through time. This happens as follows.

As the eyes move relative to the object's surface, they may fixate on different views of the object than the initial one. When this happens, the first view-specific category is inhibited by ART mismatch (Fig. 14b, c) thereby enabling a view-specific category of the new object view to be learned in ITp (Fig. 14d). When the first view-specific category is inhibited, the cells that it activated in ITa are not inhibited. They remain active so that the new view-specific category can be associated with them. As the eyes continue to scan the object, all the newly learned view-specific categories can continue to be associated with the active ITa cells. In this way, the ITa cells learn to fire in response to all the ITp categories that have been associated with it; that is, the ITa cells become a view-invariant object category.

Why did not the ITa cells shut down when the first view-specific category was inhibited, given that these ITp cells activated these ITa cells in the first place, and their inputs to ITa are eliminated when the first view-specific category is inhibited? The ARTSCAN model, and its 3D ARTSCAN Search

Downloaded from Brill.com04/06/2021 01:46:22AM
via free access

generalization, predicts that the shroud protects the ITa cells from getting reset, even while view-specific categories in ITp are reset, as the eyes explore an object. The shroud does this by inhibiting an ITa reset mechanism in the parietal cortex that would otherwise have inhibited ITa (Fig. 17). When the eyes move to attend a new object, this reset mechanisms is disinhibited, and can inhibit ITa.

At around the same time that the ARTSCAN model was published, Chiu and Yantis (2009) published data that provide strong support for prediction that

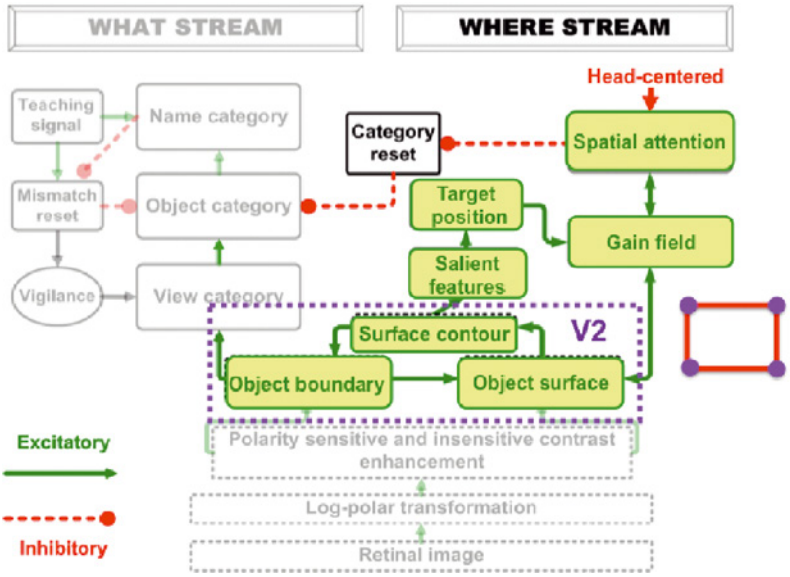


Figure 17. When spatial attention is maintained on an object's surface via a surface–shroud resonance, the shroud inhibits parietal Category Reset cells that would otherwise inhibit the invariant Object Category representation of that object that is being learned in ITa (see Fig. 15). When there is a spatial attention shift to another object, the shroud collapses and transiently disinhibits the parietal Category Reset cells. The Category Reset cells, in turn, inhibit the currently active invariant Object Category representation in ITa, thereby enabling an invariant Object Category representation of the newly attended object to be learned. While feedback between an Object Surface representation and its attentional shroud maintains Spatial Attention on that Object Surface, feedback is also occurring between V2 Object Surface and Object Boundary representations. The feedback signals from surfaces to boundaries are called Surface Contour signals. Surface Contour signals initiate figure–ground separation of objects in the scene from one another and from their backgrounds (see Fig. 44 and surrounding text). Surface Contour signals are more active at the positions of salient features due to the way that they are computed. In response to the rectangular image in this figure (see also Fig. 11), they are most active at the rectangle corners. They can therefore also be used as the target positions of eye movements with which to scan salient features of the attended object surface, such as these corners, and thereby to trigger the learning of multiple View Category representations in ITp that can be fused via associative learning into an invariant Object Category representation (see Fig. 15).

Downloaded from Brill.com04/06/2021 01:46:22AM
via free access

spatial attention in the posterior parietal cortex, or PPC, regulates the inhibitory reset of invariant object categories in ITa. Using rapid event-related MRI to record brain activity in humans, they found that a shift of spatial attention evokes a transient domain-independent signal in the medial superior parietal lobule that corresponds to a shift in categorization rules. In ARTSCAN, collapse of an attentional shroud (spatial attention shift) disinhibits the parietal reset mechanism (transient signal) that inhibits the previous view-invariant object category and allows activation of a new one (shift in categorization rules). The transient signal is *domain-independent* because all objects inhibit the parietal reset cells while their shrouds are active, and these reset cells can inhibit all invariant object category representations in ITa. The collapse of the shroud also enables the eyes to move to another surface, whereupon a new view-specific and view-invariant object category can be learned. The cycle can then repeat itself. Many other data support this prediction as well. See Grossberg (2013, 2017) for reviews, as well as the other articles that developed the 3D ARTSCAN Search architecture.

2.10. Surface–shroud and Feature–category Resonances: Conscious Seeing and Knowing

How can a viewer pay sustained attention to an interesting object, scene, or picture while scanning its interesting features? This translates into the mechanistic question: How does shroud activity persist during active scanning of an object, scene, or picture? Looking at an object generates a surface representation of the object in the prestriate visual cortex, including cortical area V4, as was discussed in Sect. 2.1. The surface representation, in turn, activates a shroud in the posterior parietal cortex, or PPC. Several different objects can simultaneously compete for spatial attention in PPC. All of them also send positive feedback signals back to their generative surfaces, thereby creating what is called a recurrent, or feedback, on-center off-surround network (Fig. 18). The on-center consists of the positive feedback signals between the surface in V4 and its shroud in PPC. The off-surround consists of the recurrent inhibitory signals within PPC whereby the shrouds compete with each other.

It has been known for many years how such a network can contrast-enhance, or choose, the representation that starts with the most activity (e.g., Grossberg, 1973, 1980). When this happens, the winning shroud focuses spatial attention on its surface (Fig. 18). The positive feedback loop sustains spatial attention on the object surface while it amplifies and synchronizes the activities of the attended cells, thereby leading to what is called a *surface–shroud resonance* between a surface representation (e.g., in cortical area V4) and spatial attention (e.g., in PPC).

The concept of a surface–shroud resonance also provides new insights into basic issues such as: How do we consciously see? How is what we

Downloaded from Brill.com04/06/2021 01:46:22AM
via free access

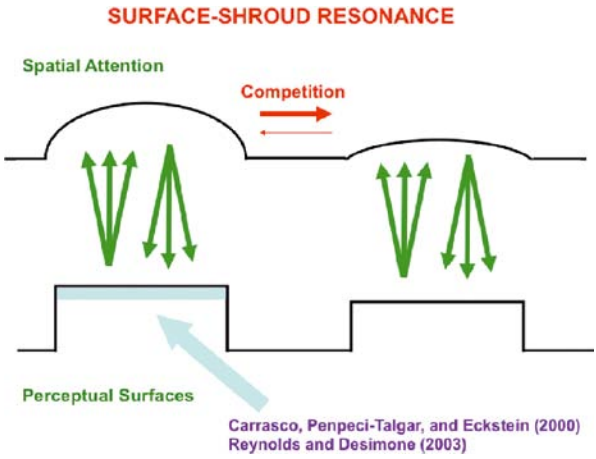


Figure 18. A cross section of the luminance of a simple Perceptual Surface representation, say in prestriate cortical area V4, contains a more luminous region, seen as a rectangle in cross section, and a less luminous region, also a rectangle in cross section. Each active surface sends bottom-up topographic excitatory signals (green upward-pointing arrows) to a Spatial Attention network in posterior parietal cortex (PPC). The PPC cells that receive inputs from the more luminous surface become more active than those which receive inputs from the less luminous surface. Cells in PPC compete among themselves (red arrows) and send topographic excitatory feedback signals back to their surfaces (green downward-pointing arrows). These interactions create a recurrent on-center off-surround network that is capable of contrast-enhancing the PPC activities, thereby creating a surface-fitting attentional shroud with which to focus spatial attention upon the more luminous surface. The top-down excitatory signals also increase the perceived contrast of the attended surface (enhanced surface activation in pale blue), as has been shown in both psychophysical experiments (e.g., Carrasco *et al.,* 2000) and neurophysiological experiments (e.g., Reynolds and Desimone, 2003).

see related to what we recognize? Grossberg (2009) proposed that we consciously see surface–shroud resonances; that is, we see the visual qualia of a surface when they are synchronized and amplified within a surface–shroud resonance. Such a resonance can propagate both top-down from V4 to lower cortical levels, such as V2 and V1, as well as bottom-up to higher cortical areas, such as the prefrontal cortex, or PFC, thereby synchronizing an entire cortical hierarchy.

The hypothesis that a surface–shroud resonance supports conscious seeing of visual qualia provides an answer to a basic question that was raised by two previous predictions. The first prediction was made when ART was first being introduced (e.g., Grossberg, 1980); namely, that "all conscious events are resonant events". The second prediction was made when the properties of boundaries and surfaces were articulated (e.g., Grossberg, 1987a, b, 1994); namely, that "all conscious percepts of visual qualia are surface percepts"; see Sect. 2.1. Combining these two predictions leads to the question: What kind of resonance supports conscious percepts of visual qualia? The prediction by

Downloaded from Brill.com04/06/2021 01:46:22AM
via free access

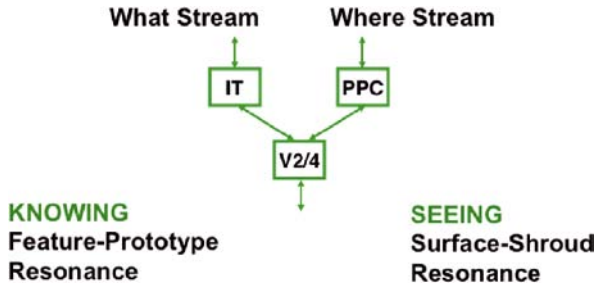


Figure 19. Coordinating conscious seeing and recognizing. A surface–shroud resonance for conscious seeing may be synchronized with a feature–prototype resonance for conscious knowing via shared circuits in prestriate cortical areas V2 and V4, despite the fact that the spatial attentional shroud is in the Where stream and the recognition category is in the What stream.

Grossberg (2009) is: a surface–shroud resonance. Grossberg (2017) explains many data about conscious and unconscious visual perception in normal individuals and clinical patients using surface–shroud resonances.

How do we recognize what we see? While a surface–shroud resonance is being activated between V4 and PPC, the brain is also activating known recognition categories, or learning new ones, via interactions between brain regions such as V2, V4, ITp, and ITa. These latter interactions cause a *feature–category resonance* (Fig. 13) to occur. Surface–shroud resonances for seeing and feature–category resonances for recognizing are linked via prestriate cortical areas like V4, with the surface–shroud resonances including Where cortical stream regions like PPC, and the feature–category resonances including What cortical stream regions like IT. When the two kinds of resonances get synchronized to each other via their shared representations in the visual cortex, we can simultaneously see *and* know things about an attended object (Fig. 19).

The above concepts clarify how an observer may become conscious of different views of a painting, and simultaneously know things about them, as the observer's spatial attention shifts from one part of the painting to the next.

2.11. Scene Understanding, Contextually-Cued Search, and Episodic Learning and Memory

The ARTSCENE model (Sect. 2.4) clarifies how the gist of a scene can be learned, and how gist may be refined by spatial attention shifts across space that learn finer textures of a scene (Grossberg and Huang, 2009). Indeed, each attention shift enables a newly attended region to be selected using its own attentional shroud, which then restricts learning to that region's texture category. The 3D ARTSCAN Search model clarifies, in addition, how we can learn to recognize and search for valued objects in such a scene.

Downloaded from Brill.com04/06/2021 01:46:22AM
via free access

Neither of these models proposes how, as the eyes scan a scene, they learn to accumulate contextual information about the kind of scene that it is. For example, when a refrigerator and a sink are viewed in sequence, an expectation of other kitchen appliances, like stoves and microwaves, may be primed, rather than of beds or beaches. If the refrigerator and sink are familiar, and in familiar positions relative to one another, then more definite expectations of particular kitchen appliances and their positions may be primed. The same is true for any scene, including a painting, leading to expectations whose confirmation or violation after a shift of attention may influence both our recognition, and our aesthetic appreciation, of it.

The ARTSCENE Search model (Fig. 20; Huang and Grossberg, 2010) goes beyond the ARTSCENE model to propose how object and spatial contexts about a scene, or picture, can be created as our eyes sequentially scan it, and then be used to more efficiently search for desired goal objects based upon this stored information. ARTSCENE Search has been used to quantitatively simulate many challenging data about what is called *contextual cueing* within the cognitive science literature; e.g., data of Brockmole *et al.* (2006), Chun

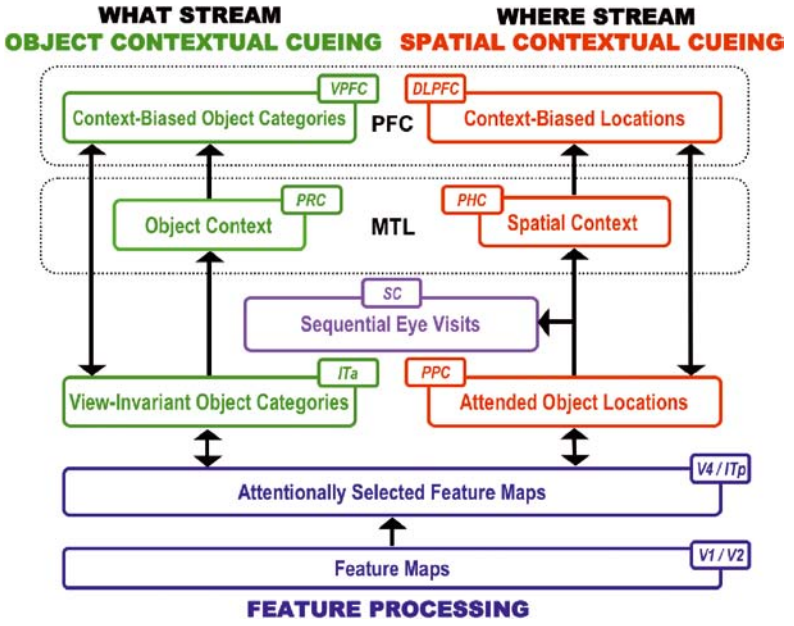


Figure 20. The ARTSTREAM Search model computes object contexts in the perirhinal cortex (PRC) and spatial contexts in the parahippocampal cortex (PHC). This contextual information enables PRC and PHC to interact with the ventral prefrontal cortex (VPFC) and dorsolateral prefrontal cortex (DLPFC) to compute object and spatial plans that can control object and spatial searches for desired goal objects. See text for details. (Reprinted with permission from Huang and Grossberg, 2010).

Downloaded from Brill.com04/06/2021 01:46:22AM
via free access

(2000), Chun and Jiang (1998), Jiang and Chun (2001), Jiang and Wagner (2004), Lleras and von Mühlenen (2004), and Olson and Chun (2002). An *object context* accumulates information about the sequence of objects that the eyes foveate as they scan a scene. Such sequential information is stored in an *object working memory*. The contents of this working memory can then be used to learn a special kind of recognition category that is called an *object list chunk*, or plan. Learning and activation of such a plan as a series of eye movements unfolds represents that particular object context. In much the same way, a *spatial context* accumulates information about the sequence of positions that the eyes foveate as they scan a scene. Such sequential information is stored in a *spatial working memory*. The contents of this working memory can then be used to learn a special kind of recognition category that is called a *spatial list chunk*, or plan. Object and spatial plans can be used to direct a search for additional objects that are expected after such an object and spatial context have been experienced, as well as to read out sensory expectations of these objects that prime us to expect to see them.

The signals from active object and spatial plans may be combined to 'vote' for the object that a viewer most probably expects to see, as well as the position where this object is expected to be seen. This kind of knowledge begins to achieve context-sensitive *scene understanding*. It helps us to understand a painting as an entity within which certain combinations of objects, textures, and colors occur where we expect them to be.

ARTSCENE Search clarifies how the brain does this using interactions between multiple brain regions. In addition to the brain regions that are modeled by the 3D ARTSCAN Search model, the ARTSCENE Search model also includes interactions of the perirhinal and parahippocampal cortices with the temporal, parietal, and prefrontal cortices (Fig. 20). Object contexts are computed in the perirhinal cortex, and spatial contexts are computed in the parahippocampal cortex. Their interactions with object and spatial plans in the prefrontal cortex read out visual expectations of what and where new objects, textures, and colors are expected to be seen.

With these mechanisms also available, a viewer of a painting can more intelligently explore different parts of a painting based upon its gist and other regions of the painting that have already been viewed.

3. How Boundaries Are Completed and Surfaces Filled-in with Brightness and Color

3.1. From Complementary Boundaries and Surfaces to Complementary Consistency

As summarized in Sect. 2, all visual percepts are built up from interactions between boundary and surface representations. These representations are

Downloaded from Brill.com04/06/2021 01:46:22AM
via free access

processed in parallel cortical processing streams (Fig. 9) that individually compute computationally complementary properties (Fig. 10). Interactions between these streams, across multiple processing stages, overcome their complementary deficiencies to compute effective representations of the world. Said in another way, these interactions convert computations that obey complementary laws into a consistent percept, hereby achieving the property of *complementary consistency*. In order to better understand important properties of many different kinds of paintings, additional information is needed about how boundaries are completed and surfaces filled-in.

3.2. Oriented Filtering, Spatial and Orientational Competition, and Bipole Grouping

The boundary signals that activate the brain may be incomplete either because they receive incomplete inputs from the world, as illustrated by the famous example of a Dalmatian in Snow (Fig. 3), or because they are processed by the retina across the blind spot or retinal veins (Fig. 8). Oriented filtering by simple cells and complex cells (Fig. 5) cannot complete these incomplete boundary fragments. However, output signals from complex cells input to a subsequent processing stage where boundary completion may be realized by a process called *bipole grouping* (Fig. 5). Bipole grouping cells can cooperate together to complete boundaries when their (almost) collinear and (almost) like-oriented neighbors are also activated. Bipole grouping hereby completes boundaries *inwardly* between pairs, or greater numbers, of (almost)-collinear and (almost)-like-oriented cells. Bipole grouping is thus an *oriented* process, and can occur between image contrasts with opposite contrast polarities, and is in this sense *insensitive to contrast polarity*, because bipole cells are activated by complex cells.

There is a vast experimental literature about properties of the perceptual grouping process of which bipole grouping forms a part. See Wagemans et al. (2012a, b) for a useful review.

The grouping properties of bipole cells were predicted and simulated in a neural network vision model by Cohen and Grossberg (1984), Grossberg (1984), and Grossberg and Mingolla (1985a, b). Cells with the predicted bipole properties were first reported to occur in cortical area V2 based upon neurophysiological experiments by von der Heydt *et al.* (1984). Laminar cortical models of how the brain sees (e.g., Grossberg, 1999, and Grossberg and Raizada, 2000) proposed, moreover, how recurrent long-range excitatory interactions between neurons in layer 2/3 of V2, supplemented by short-range inhibitory interneurons, can realize the bipole grouping property (Fig. 21), and thereby explain consistent anatomical data.

Downloaded from Brill.com04/06/2021 01:46:22AM
via free access

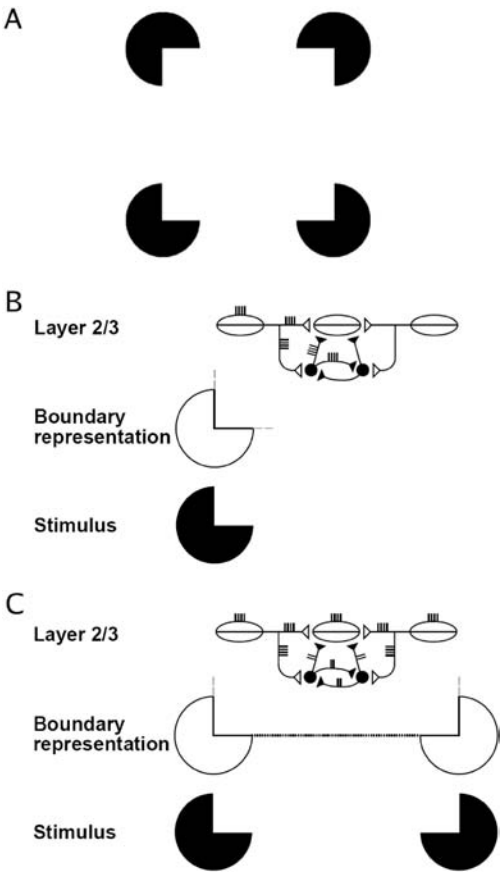


Figure 21. How bipole cells in layer 2/3 of cortical area V2 as modeled within the 3D LAMINART model complete boundaries between pacman inducers of a Kanizsa square. (A) Kanizsa square stimulus. (B) Response of a horizontally-oriented bipole cell to a horizontal edge of one pacman figure. The pacman-activated horizontally-tuned bipole cell sends excitatory signals (vertical spikes) along its horizontally-oriented long-range excitatory horizontal connections to a nearby bipole cell. The horizontal connections also excite short-range disynaptic inhibitory interneurons (more vertical spikes) that inhibit the target bipole cell. The inhibitory interneurons also inhibit each other (still more vertical spikes), thereby normalizing their total activity in response to different combinations of inputs. The excitation and inhibition that converge upon a bipole cell tend to cancel each other ('one-against-one'), thereby preventing individual inducers from causing widespread horizontal activation across the visual field. (C) When two collinear pacmen are presented, pairs of horizontally-tuned bipole cells on opposite sides of an intervening target bipole cell can summate their excitatory inputs at that cell, as they also activate their short-range inhibitory interneurons. The inhibitory interneurons again inhibit the target bipole cell, but also use recurrent interactions to inhibit each other. As a result of these recurrent inhibitory connections, the total inhibition on the bipole cell from all the active interneurons is normalized, and the bipole cell can fire ('two-against-one'). In this way, bipole cells can build illusory contours inwardly between pairs of inducers, but not outwardly from a single inducer. (Adapted with permission from Grossberg and Raizada, 2000).

Downloaded from Brill.com04/06/2021 01:46:22AM
via free access

Figures 5 and 21 note that boundary completion is oriented. How is this property achieved? To initiate boundary processing, the brain uses simple cells and complex cells, which are oriented contrast detectors in cortical area V1, as noted in Sect. 2. Orientationally-selective tuning of cells has limitations, however. By averaging input contrasts across their receptive fields, simple cells and complex cells enable the brain to avoid the need to use a large number of different specialized detectors to edges, or shading, or texture, and some other highly specific visual properties. Instead, a single simple cell and complex cell can respond to edges, shading, and texture, among other visual properties, that roughly share the orientation and spatial scale of their receptive fields. For example, in Fig. 22, a vertically oriented simple cell with a dark–light polarity can respond at the right edge of a vertical black bar on a white background. Such a simple cell can also respond to the contrast difference that is caused by two vertically oriented black dots on the white background. However, such a simple cell cannot respond to the white background itself, because the background has uniform luminance. Nor can a horizontally oriented simple cell respond to the vertical black bar.

This efficiency also leads to computational uncertainties. For example, the oriented receptive fields of simple cells and complex cells cannot respond at the ends of thin lines, or other high curvature contours, as illustrated by the computer simulation in Fig. 23 (left image). Such a hole in a boundary is called an *end gap*. Without further processing, the end gaps in the boundaries at line ends would allow brightness and color to flow out of every line end due to surface filling-in, leading to perceptually disastrous consequences.

To prevent this perceptual catastrophe, additional processing stages are needed subsequent to complex cells. Grossberg (1984) and Grossberg and Mingolla (1985a) predicted and modeled how this additional processing can be accomplished in two stages (Fig. 5): Complex cells are proposed to input to a subsequent processing stage via a *spatial competition*. Cells at this next stage are often called *hypercomplex* cells (Hubel and Wiesel, 1968). They receive inputs from the complex cells via an on-center off-surround network: Each complex cell excites the hypercomplex cell at its position with the same orientational preference, while inhibiting nearby hypercomplex cells with the same, or similar, orientational preferences. Due to the off-surround, the responses of hypercomplex cells are sensitive to the length of input lines, since line inputs that fall within a strong region of inhibition within the off-surround are suppressed. The cells that are affected by this on-center off-surround spatial competition are said to constitute the *first competitive stage*.

Many boundaries would still remain incomplete if boundary processing stopped with the first competitive stage. For example, line end boundaries

Downloaded from Brill.com04/06/2021 01:46:22AM
via free access

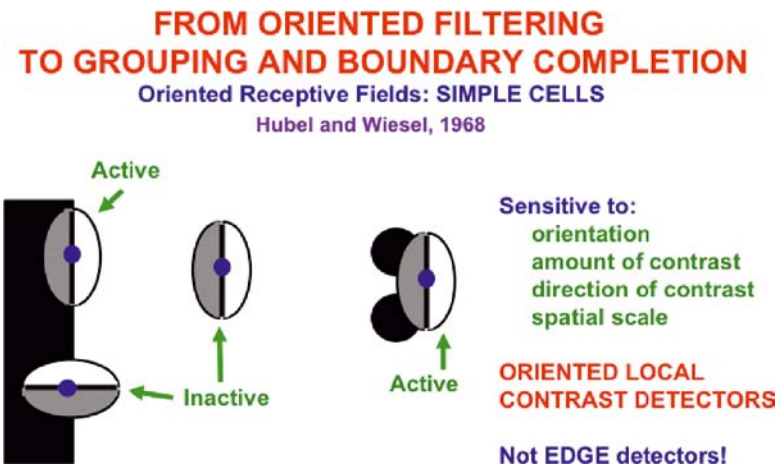


Figure 22. How simple cells that are sensitive to specific positions, contrast polarities, and orientations respond to contrastive stimuli. See text for details.

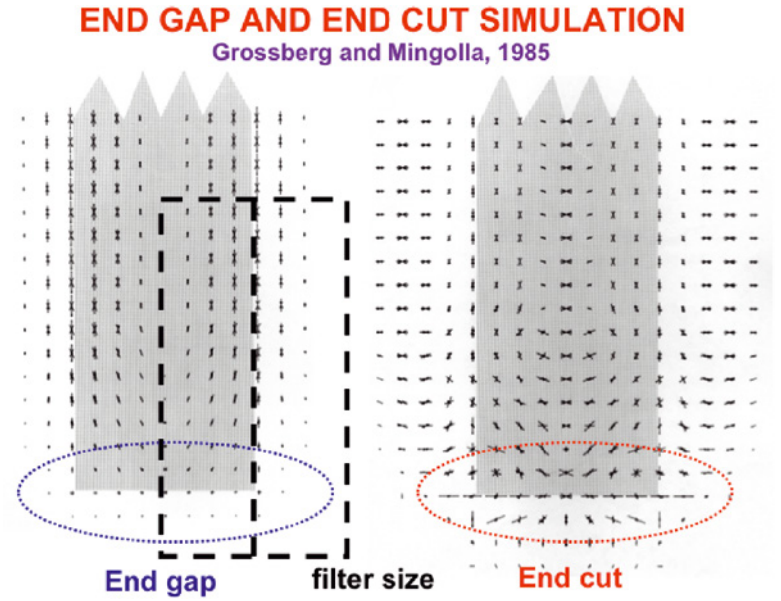


Figure 23. Computer simulations of the responses of simple cells at the end of a thin line (in gray). (Left) Simulated activities at different positions of simple cells with prescribed orientational tuning are represented by the lengths of lines at the same positions and with the same orientations. The largest responses occur in vertically oriented simple cells at the two sides of the line, but no responses occur at the end of the line. (Right) The first and second competitive stages (Fig. 5) create a boundary at the line end (end cut) that is positionally accurate but orientationally fuzzy. (Reprinted with permission from Grossberg and Mingolla, 1985b).

Downloaded from Brill.com04/06/2021 01:46:22AM
via free access

would still have a hole in them, and the two Kanizsa squares in Fig. 4 would continue to be seen and recognized as four spatially disjoint pacman figures.

The cells of the first competitive stage activate cells at a *second competitive stage*. These latter cells compete with other cells at the same position that have different orientational tuning (Fig. 5). The maximal mutual inhibition occurs between cells that prefer perpendicular orientations. Cells at the second competitive stage are also tonically active, which means that they are continuously activated by an internal energy source. However, their mutual inhibition holds down their activity when all orientations receive no external inputs. Suppose, however, that a cell in the second competitive stage lies just beyond the end of a line. Then it will be additionally inhibited by cells at the line end from the first competitive stage. Then its inhibition of other cells at its position in the second competitive stage will also be reduced, notably of cells that code the perpendicular orientation, and orientations close to it. These latter cells are hereby disinhibited, and can create a boundary at the line end. Such a boundary is called an *end cut* (Fig. 23, right). The spread of brightness and color signals out of every line end is prevented by their end cuts.

In summary, the two competitive stages compensate for the missing boundaries at the end of a vertical thin line as follows: The competition across the hypercomplex cells at the first competitive stage is *across position* and *within orientation*, whereas the competition across the hypercomplex cells at the second competitive stage is *within position* and *across orientation*. At the first competitive stage, active vertically oriented complex cells near the line end inhibit vertically oriented hypercomplex cells just beyond the line end. When these cells are inhibited, their inhibition of cells at their position that are tuned to other orientations is removed. The most inhibition is removed from cells that are preferentially tuned to the horizontal orientation. These horizontally oriented cells are hereby disinhibited, and create line ends, or end cuts, that complete the boundary at the line end, and are capable of containing the line's brightness and color within its borders.

Designers of letter fonts often attach a small line, or *serif*, to the end of a stroke in a letter. There are many possible historical explanations of serifs, or "Roman" typefaces, but one psychological benefit of them is to strengthen the boundary at a line end.

Although cells at the second competitive stage can complete the boundaries at line ends, they cannot complete boundaries due to incomplete textures such as the Dalmatian in Snow (Fig. 3), or across the blind spot and retinal veins (Fig. 8). Cells at the second competitive stage input to bipole cells that can carry out boundary completion (Fig. 5). Because cells that are tuned to several different orientations are activated at each line end at the second competitive stage (Fig. 23, right), bipole cells can group across space in any of several orientations, depending upon which of these orientations line up best across space and have the most support from inputs in the scene or picture.

Downloaded from Brill.com04/06/2021 01:46:22AM
via free access

3.3. Hierarchical Resolution of Uncertainty

These spatial and orientational competitive interactions are part of a process that is called *hierarchical resolution of uncertainty* (Grossberg, 1984, 1994; Grossberg and Mingolla, 1985a) whereby uncertainties caused by lower levels of boundary processing are compensated by processes at higher cortical levels.

Why is hierarchical resolution of uncertainty needed? Why does not the brain just define perfect edge detectors that do not fail to detect line ends? A basic reason for this has already been briefly noted; namely, simple cells are not just edge detectors. They are oriented local contrast detectors that can respond to edges, textures, shading, and depth differences. If the brain defined edge detectors instead, then it would also have to define many other types of specialized detectors. It would then be faced with the extremely challenging problem of figuring out how to put together the specialized information that is detected by all these different kinds of detectors. Such a fusion of information could be rendered impossible by the fact that many properties of scenes and paintings contain edges, textures, shading, and depth all overlaid in the same locations. Specialized detectors could not respond accurately to such an overlay of properties.

Instead, the brain starts to build boundaries with simple cells, which can respond to all of these different types of properties, albeit imperfectly, and then uses hierarchical resolution of uncertainty to compensate for where these detectors are insensitive to image data. It turns out that the interactions that accomplish hierarchical resolution of uncertainty also play many other useful roles in visual perception. Indeed, several hierarchical resolutions of uncertainty are used to generate the visual percepts that we consciously see. Three more of them are summarized below, one having to do with the process of boundary grouping, another with surface filling-in, and the third with figure–ground separation. The general conclusion is that the brain does not eliminate uncertainty too soon. It takes advantage of uncertainty until processing stages are reached at which uncertainty can profitably be drastically reduced or eliminated. Many artists have also understood this point, at least implicitly, with the Impressionists, such as Claude Monet (see Sect. 4.1), and other plein air painters being notable examples.

3.4. Neon Color Spreading: Coordinating End Gaps, End Cuts, and Fuzzy-to-Sharp Groupings

Many experimental data and percepts support the idea of hierarchical resolution of uncertainty and how it works. For example, evidence of how end cuts usually prevent spurious filling-in from occurring can be seen in some percepts due to the fact that, although the two competitive stages work most of the time, they do not contain spurious color spreading in response to all scenes

Downloaded from Brill.com04/06/2021 01:46:22AM
via free access

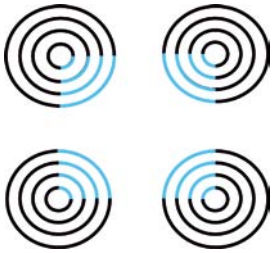


Figure 24. A neon color spreading stimulus. See text for an explanation of how the illusory square boundary is formed between the positions where the black and blue arcs touch, and how blue color spills out of the blue line ends to fill-in the emergent square with blue color.

and pictures. A famous example of color spreading from line ends is shown in Fig. 24. It is called *neon color spreading*. Neon color spreading illustrates key properties of color filling-in, in addition to how end gaps and end cuts are formed. The image in Fig. 24 consists of circular annuli, part of which are black and part blue. When viewing this figure, an illusory square can be seen that is filled with blue 'neon' color, even though the only blue in the image is in the concentric blue circular arcs.

Neon color spreading was reported in Varin (1971), who studied a "chromatic spreading" effect that was induced when viewing an image similar to the one in Figure 24. Van Tuijl (1975) independently introduced images that gave rise to percepts that he called "neon-like color spreading". To explain the percept in Figure 24, first note that the black and blue arcs in Fig. 24 both create boundaries in our brains. At the positions where the differently colored boundaries join, the boundaries caused by the black arcs cause end gaps to occur at the ends of the boundaries caused by the blue arcs, at positions where they abut the black arcs. For this to occur, the contrast of the black arcs with respect to the white background must be chosen larger than the contrast of the blue arcs with respect to the white background. The boundaries formed by these contrasts are contrast-sensitive, so that the boundaries formed by the black–white contrasts are stronger than those formed by the blue–white contrasts, and can thus inhibit these spatially abutting weaker boundaries using the spatial competition of the first competitive stage, just as they do in Fig. 23 (left).

Although the heuristics of these operations are simple to describe, to actually make them work requires that one uses the correct cellular laws and network interactions. This fact can be illustrated when we try to answer the following question: If spatial competition occurs along an entire line, then why does not the balanced inhibition along the entire interior of the line suppress it entirely? The entire line could have been inhibited if spatial competition just *subtracts* activity from its target cells. Instead, a form of competition is used that is called *shunting* competition, which fortunately is the kind of

Downloaded from Brill.com04/06/2021 01:46:22AM
via free access

competition that is obeyed by the membrane equations that define the properties of neurons. Shunting competition does not just subtract activity from inhibited cells. Instead, it also *divides* the activity of these cells. Shunting competition can be designed so that, when the strength of competition between nearby cells is the same, such as along a black arc, then the activities of these cells are reduced, but not eliminated, by an amount that is sensitive to the *relative* activities of all the neighboring cells. This property is called *contrast normalization*. However, at positions where the weaker blue arc inputs receive shunting competition from nearby black arc inputs, shunting competition can cause holes, or end gaps, in blue arc boundaries at positions that lie just beyond the ends of black arc boundaries.

The spatial competition that causes end gaps in response to a neon color spreading image can also cause end cuts when the first competitive stage inhibits cell activities at the second competitive stage, as in Fig. 23 (right). In particular, cells at the second competitive stage that are tuned to other orientations, notably the perpendicular orientation, are disinhibited and thereby create end cuts. Thus, both end gaps and end cuts can occur where the black arcs abut the less-contrastive blue arcs in a neon color spreading image.

Because end cuts occur at multiple line ends in response to the neon color spreading image, the end cuts activate bipole cells that can cooperate across space to complete a boundary with the shape of an illusory square.

Bipole grouping can occur even when the adjacent end cuts do not have the same orientational preference. Grouping can happen because these orientations are similar enough to be grouped by the bipole cell receptive fields. These receptive fields prefer to group collinear orientations, but can also group non-collinear orientations with a strength that decreases with the orientational difference from the preferred orientation and the physical distance from the bipole cell body (Grossberg and Mingolla, 1985b), as noted in Fig. 25. These receptive field properties can arise during development of the bipole cells as they are exposed to the statistics of many scenes and images. See Grossberg and Swaminathan (2004) and Grossberg and Williamson, (2001) for computer simulations of how bipole receptive fields can develop.

Due to this uncertainty in the bipole cell receptive fields, bipole cells can interact together to complete boundaries in response to collinear, perpendicular, or oblique inducers (Fig. 26). Each of these final boundary groupings typically appears to be spatially sharp, despite the uncertainty, or fuzziness, in the number of possible inducing orientations before a final grouping is chosen. This property of transforming an initially fuzzy orientational and positional grouping choice into a sharp final grouping is a second example of hierarchical resolution of uncertainty.

This particular hierarchical resolution solves the following problem: The probability that pairs of nearby inducers are perfectly aligned across space is

Downloaded from Brill.com04/06/2021 01:46:22AM
via free access

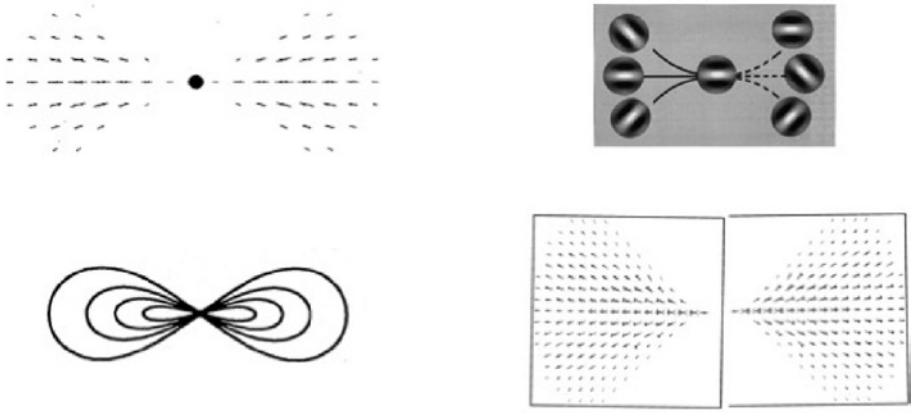


Figure 25. (Left, upper row) This image shows the receptive field of the bipole cell that was simulated by Grossberg and Mingolla (1985b). The black dot represents the position of the bipole cell body. The other markings represent the receptive field of the cell. The length and orientation of each line represents the relative strength of the bipole cell receptive field's reaction to an input at that position with that orientation. Colinear, like-oriented inputs are favored, but nearby orientations and positions can also influence bipole cell firing. (right, upper row) This image depicts the "association field" that was described by Field *et al.* (1993) as a result of their psychophysical experiments. As in the bipole cell receptive field, collinear inputs are favored, but nearby orientations and positions can also have an effect. (Left, lower row) This figure shows the version of bipole grouping modeled by Heitger and von der Heydt (1993). (Right, lower row) This figure shows the version of the bipole cell that was modeled by Williams and Jacobs (1997).

vanishingly small, especially when one considers that the design and placement of neurons is not perfect. Despite this limitation, it is important to be able to initiate grouping between them. The fuzzy bipole cell receptive fields enable this to occur (Fig. 25), and a band of grouping possibilities is initially generated (Fig. 27). Once grouping begins, however, cooperative and competitive bipole cell interactions choose the strongest grouping and inhibit the weaker ones, thereby converting a fuzzy initial grouping into a sharp final grouping. When the strongest inducers differ in orientation and position, the final boundary groupings may go through the locally preferred end cut orientations; namely, the orientations that are perpendicular to their line ends (Fig. 28, upper row). However, if there is no sufficiently simple grouping that can go through the perpendicular end cuts, then a grouping may form that goes through weaker end cut orientations (Fig. 28, lower row). The boundary completion process chooses the grouping that has the most evidence in support of it, using the same cooperative and competitive interactions that induce end gaps and end cuts. These interactions realize a kind of probabilistic hypothesis testing that chooses the best supported grouping from uncertain data.

Downloaded from Brill.com04/06/2021 01:46:22AM
via free access

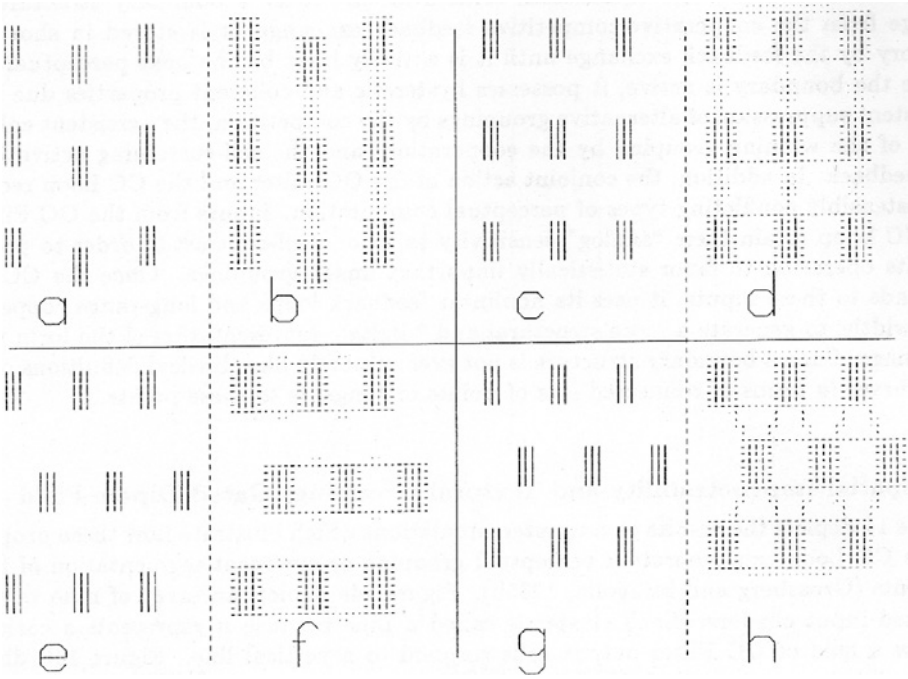


Figure 26. Gestalt grouping by bipole cells, notably how bipole cells can build illusory boundaries that are parallel, perpendicular, or diagonal in orientation relative to their inducing stimuli. (a, c, e, g) In each case, the inducing stimuli are three rows of three vertically oriented bars. The bars are moved relative to one another to illustrate how Gestalt grouping laws can lead to (b) vertical groupings, (d) vertical and horizontal groupings, (f) horizontal groupings, and (h) diagonal groupings. (Reprinted with permission from Grossberg and Mingolla, 1985b).

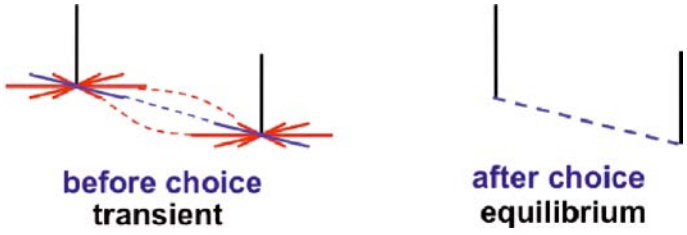


Figure 27. (Left) An initially fuzzy grouping can form in response to the multiple orientations that are created within end cuts; see Figure 23 (right). (Right) A sharp grouping, after the grouping process equilibrates, chooses the grouping with the most support from the totality of its inducers, while suppressing weaker groupings via short-range competitive interactions, as illustrated by Fig. 5.

Downloaded from Brill.com04/06/2021 01:46:22AM
via free access

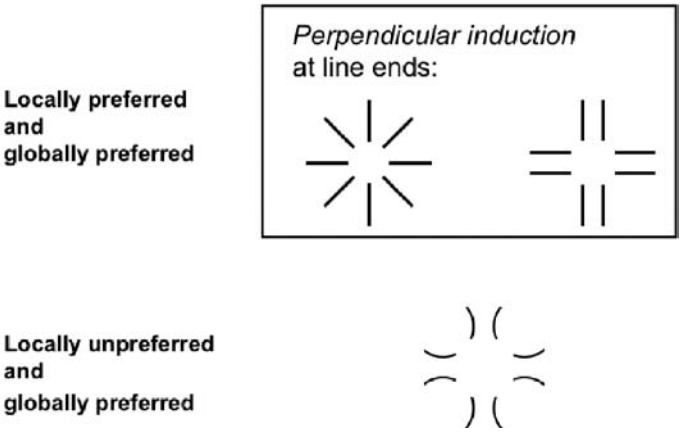


Figure 28. (Upper row) Two stimuli that lead to illusory contours that are perpendicular to their inducing line ends. Here the locally-preferred perpendicular inducer orientations win in the globally chosen grouping. (Lower row) Here the emergent square boundary generates constraints that can only be satisfied by non-perpendicular orientations within the end cuts at the inducing line ends. Thus this global grouping is not built on globally preferred perpendicular end cuts.

The simulations in Figs 23 and 26 illustrate another important property of boundary completion. Note that, in addition to the end cuts that enable line ends to group, there are also small horizontal boundaries that are created on both sides of the vertical boundary inducers in each input. However, these small horizontal boundaries do not create emergent horizontal boundaries between pairs of inputs along the entirety of their sides. If this did happen, then many spurious boundary groupings would occur. They do not occur because of a property that is called *spatial impenetrability* (Grossberg and Mingolla, 1985b). Spatial impenetrability is achieved by the fact that (e.g.) horizontally-oriented bipole cells are inhibited by vertically-oriented hypercomplex cells at the same positions (Fig. 5b). In the middle of a vertical line input, a combination of horizontal and vertical responses prevents horizontally-oriented bipole cells from "penetrating" these vertical inducers. Only at line ends are there enough like-oriented cell responses to be able to induce boundaries, whether collinear (Fig. 4) or almost perpendicular to (Fig. 28) the line's orientation.

In summary, the illusory contours that are illustrated in Figs 4, 24, 26, and 28 can be quite sharp, despite the fuzziness of the end cut orientations, because bipole cells interact via feedback with spatial and orientational competition to inhibit weaker cell responses (Figs 5 and 21). This sharpening of end cuts during illusory contour formation is another example of hierarchical resolution of uncertainty: The initial fuzziness is needed to initiate boundary grouping (Fig. 27) but risks a loss of acuity. The ensuing sharpening by

Downloaded from Brill.com04/06/2021 01:46:22AM
via free access

cooperative–competitive feedback interactions permits the final grouping to form without a loss of acuity.

When considering properties of this boundary choice, another question arises whose answer also requires that cellular and network computations be done in the correct way. This question asks how strong the boundary choice will be? There are scores of examples that illustrate the need for boundary strength to covary with the number, spatial separation, relative orientation, and contrast of the boundary inducers. These constraints are often called Gestalt grouping laws. One can summarize them heuristically by saying that the *degree of commitment—in this case, boundary strength—should covary with the amount of evidence*. The computational challenge is that cooperative-competitive feedback interactions that make a boundary choice could easily cause the winning cells to attain their maximal activities, and thus be insensitive to the amount of evidence. This could cause catastrophic problems if the amount of evidence is needed to determine adaptive choices at later processing stages. It has been shown that the *laminar circuits* of the visual cortex enable such choices to be made without losing analog sensitivity to the Gestalt grouping laws (e.g., Grossberg and Raizada, 2000; Grossberg *et al.*, 1997), a property that Grossberg has called *analog coherence*.

3.5. Neon Color Spreading: How Discounting the Illuminant Is Compensated by Filling-in

In response to the neon color spreading image in Fig. 24, blue color can flow out of the end gaps in the broken boundaries. This spreading, or filling-in, of blue color across space continues until the color hits the square that is made up of both pacman and illusory boundaries, which together prevent its further spread. Thus, both real and illusory boundaries act as barriers to the filling-in of brightness or color. In summary, the process of end cutting that completes line ends, and thus illustrates the first hierarchical resolution of uncertainty, can also create end gaps through which brightness or color can flow when a stronger boundary abuts a collinear weaker boundary. The end cuts can, in turn, cooperate via bipole grouping to complete an illusory boundary that can contain this flow of brightness or color, thereby illustrating a second hierarchical resolution of uncertainty.

Neon color spreading also provides a third example of hierarchical resolution of uncertainty. Recall from Sect. 2.1 that the process of discounting the illuminant computes feature contour signals that significantly eliminate effects of the illuminant, in order to compute properties of the object itself, such as its reflectances. Feature contour signals tend to occur along boundary contours (Fig. 11) because those are the positions where image contrasts may occur rapidly across space. Intuitively, one can think of feature contours as a kind

Downloaded from Brill.com04/06/2021 01:46:22AM
via free access

of color or brightness contour that is localized in space. In order to recover surface properties of an object, such as its color and brightness, in a way that is relatively insensitive to illumination changes, filling-in of feature contours occurs at a processing stage that is subsequent to discounting the illuminant (Fig. 12). Thus, the filling-in that is observed during neon color spreading illustrates a hierarchical resolution of uncertainty whereby feature contours are computed to discount the illuminant, and then the feature contours induce filling-in within the boundary contours to recover surface color and brightness.

In summary, the charming but seemingly unimportant percept of neon color spreading may be understood as the combined effect of three hierarchical resolutions of uncertainty, each of which is essential for the visual system to work well. The insights that we have gleaned about how neon color spreading occurs will be used below to help clarify effects that various artists have achieved in their paintings.

3.6. Seeing a 2D Image of a Shaded Ellipse as a 3D Surface: Boundary Webs

The examples to the present have focused on 2D properties of visual percepts. However, even in response to the Kanizsa square in Fig. 4, the completed square can consciously be seen 'in front of' four partially occluded discs whose unoccluded parts can be seen as pacman figures. This percept leads to the general question: How does a 2D picture give rise to a 3D percept? Were it not for this property, 2D paintings might never have been invented to represent aspects of the 3D world that humans experience every day. FACADE theory (e.g., Grossberg, 1994, 1997; Grossberg and McLoughlin, 1997), and its refinement and generalization by the 3D LAMINART model of how the laminar circuits of visual cortex see (e.g., Cao and Grossberg, 2005; Grossberg, 1999; Fang and Grossberg, 2009; Grossberg and Raizada, 2000; Grossberg and Swaminathan, 2004; Grossberg and Yazdanbakhsh, 2005), explain how the brain mechanisms that have evolved to see the 3D visual world can automatically generate 3D representations of 2D pictures. Aspects of how this proposed to happen will now be reviewed.

The 3D percept that is generated by a 2D picture of a shaded ellipse is one possible place to start this explanation (Fig. 29, left). Despite its simplicity, this picture raises fundamental issues about how vision works, issues that have influenced the art of many painters. One of the first issues concerns the difference between an edge detector and a boundary detector that is sensitive to edges, shading, texture, and depth. Recall that the brain does not use edge detectors if only to avoid having to process and fuse information from a proliferation of specialized detectors. Instead, oriented simple cells are used, and their uncertain computations are compensated by hierarchical resolutions of uncertainty (Fig. 5). If the brain did use edge detectors, then it could compute

Downloaded from Brill.com04/06/2021 01:46:22AM
via free access

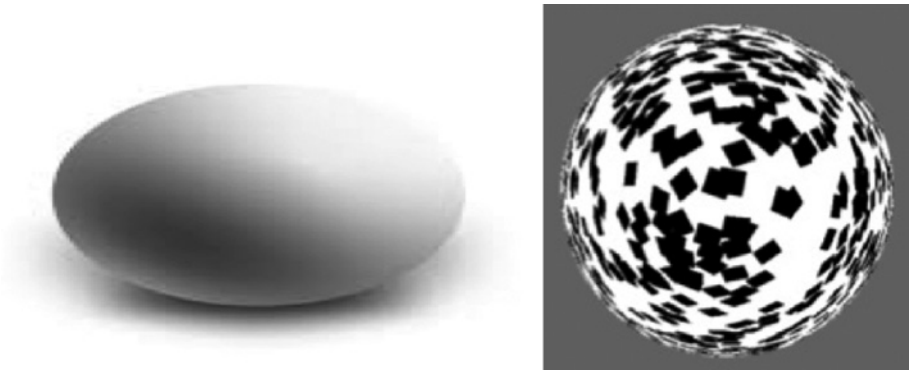


Figure 29. (Left) A shaded 2D ellipse generates a compelling rounded 3D shape percept because of the way in which multiple-scale boundary webs react to the shading gradients and capture the gray color within their form-sensitive compartments in depth-selective filling-in domains. (Right) A smooth rounded 3D shape percept is generated from a spatially discrete set of 2D black patches, again by using multiple-scale boundary webs and their interactions with depth-selective filling-in domains.

only the bounding edge of the ellipse. If this were the only boundary that formed, then the gray colors within the ellipse could all spread throughout its interior surface via filling-in to generate a flat percept of a uniformly gray ellipse, just as occurs during filling-in of a Kanizsa square (Figs 4 and 6). This is not, however, what is seen.

Instead, the brain computes a dense array of boundaries that configure themselves along the isophotes, or equal luminance positions, of the gradients of shading and texture in an image or scene. Such an array is called a *boundary web* (Grossberg, 1987a, b; Grossberg and Mingolla, 1987; Grossberg *et al*., 2007). Oriented simple cells of a given size can respond the elliptical form of the shaded gradient along these isophotes to generate such a boundary web. The web's boundaries parallel one another from the bounding contour of the ellipse inwards towards its interior. Each of the small compartments in the web can trap its local shade of gray and prevent it from spreading via filling-in to a larger region of the ellipse.

If this is indeed the case, then why is this boundary web not visible? There are two answers to this question. The first answer is that 'all boundaries are invisible'. The second answer is that the invisible boundary web reveals itself through the smooth gradient of gray color that it trap within its compartments to form a shaded surface percept.

Using simple cells with a single receptive field size, or spatial scale, is not sufficient to generate a percept of a depthful, or 3D, shaded ellipse. For this to happen, multiple simple cell spatial scales need to simultaneously respond

Downloaded from Brill.com04/06/2021 01:46:22AM
via free access

to the image. Other things being equal (which is not always the case), large boundary scales code 'near' and small scales code 'far', if only because objects of a given size caste larger images on the retinas as they are viewed at nearer distances. This property is often called the *size-disparity correlation* (Julesz and Schumer, 1981; Kulikowski, 1978; Richards and Kaye, 1974; Schor and Tyler, 1981; Schor and Wood, 1983; Schor *et al.*, 1984; Tyler, 1975, 1983). In particular, complex cells with larger receptive fields can binocularly fuse a broader range of binocular disparities than can cells with smaller receptive fields, and thus can represent a larger range of depths, with the largest disparities often coding the nearest depths.

In response to a picture of a 2D shaded ellipse, each scale of simple cells activates its own network of complex, hypercomplex, and bipole cells (Fig. 5), and each network of bipole cells responds differently to the gradient of gray shading in the ellipse: The smallest scale generates the narrowest band of boundaries in its web. This boundary web lies adjacent to the bounding contour of the ellipse. The largest scales generate the broadest bands of boundaries. Larger boundary webs can more deeply penetrate into the interior of the ellipse. Said in another way, other things being equal, smaller scale simple cells can fire more easily nearer to the bounding edge of the ellipse. As the spatial gradient of shading becomes more gradual with distance from the bounding edge, it becomes harder for smaller scales to respond to this gradient. Thus, other things being equal, larger scales tend to respond more as the distance from the bounding edge increases. As a result of the size-disparity correlation, the boundary webs nearer to the middle of the ellipse code a nearer depth than the ones near the bounding edge of the ellipse.

It should also be noted, however, that larger scales do not always code for nearer depths, due to the way in which multiple spatial scales interact with grouping properties. These exceptions will not be needed to make our main points here. They are explained in Grossberg (1994), along with a more detailed explanation of when and how larger spatial scales do code for nearer depths.

Each boundary scale, in turn, maximally activates a Filling-In DOmain, or FIDO, which controls filling-in of surface brightness and color at the corresponding depth (Fig. 12). In particular, each boundary web traps gray shading within itself at the corresponding FIDO. The spatial distribution of filled-in contrasts across all these depth-selective FIDOs represents the perceived 3D shape of the elliptical surface. Many details need to be carefully developed for this simple scheme to work properly, but the main idea should be clear.

These datails are mathematically modeled in the LIGHTSHAFT (LIGHTness-and-SHApe-From-Texture) model of Grossberg *et al.* (2007). Percepts such as the smooth 3D shape that is generated by the spatially discrete 2D texture in Fig. 29 (right) are explained by LIGHTSHAFT, whose explanation is supported by quantitative simulations of parametric psychophysical data

Downloaded from Brill.com04/06/2021 01:46:22AM
via free access

about such percepts. Here too one needs to ask: If different boundary web scales generate a depthful percept of surface form and lightness by differentially generating filling-in of the texture contrasts within them, then why are not these boundary webs visible? Again the answer is that invisible boundary webs reveal themselves by the way in which they selectively organize the filling-in of surface contrasts within their respective Filling-In Domains.

3.7. Chiaroscuro and Perspective

How multiple-scale boundary webs can create 3D percepts from a 2D image, as in response to the images in Fig. 29, also explains the effects achieved by the painterly technique of *chiaroscuro*. Leonardo da Vinci was one of the first painters to use the chiaroscuro technique to create bulging 3D percepts from 2D paintings.

Koenderink *et al*. (2015) have discussed the paintings of several other artists from the perspective that they derive depthful object percepts using a variety of 'relief articulation techniques' that can all be interpreted, and explained, by how they activate multiple-scale, and often coarse, boundary webs, even when the techniques do not respect basic physical constraints such as the location of light sources.

Vanderbosch *et al*. (2015) have additionally considered conditions under which changing a uniform background color may influence the perceived depthful relief of simple objects, such as shaded spherical images, that the backgrounds surround, and how these insights may be used to analyze paintings that employ various techniques to generate coarse boundary webs with which to trap their lightnesses and colors. A typical demonstration (their Fig. 2) compares three identically shaded gray spherical images against three different uniform backgrounds that are white, black, or gray. The white and black backgrounds both cause a fixed direction-of-contrast (from object to background) at every position along the bounding contour of the corresponding spherical image, albeit one that goes from dark-to-light when the background is white, and from light-to-dark when the background is black. In contrast, the gray background generates a light-to-dark direction-of-contrast at the top of its spherical image, and a dark-to-light direction-or-contrast at the bottom of its image, with intermediate contrasts in between. This third image raises the same issue as the image in Fig. 4 (lower row), wherein direction-of-contrast flips along the bounding contour of an object. Again, complex cells, by pooling over opposite contrast polarities (Fig. 5a), can generate a continuous boundary around the entire spherical image, except at the unique pair of positions where there is zero contrast between image and background, due to the reversal of relative contrast around the bounding contour. This infinitesimal break in the boundary may be completed by collinear bipole grouping

Downloaded from Brill.com04/06/2021 01:46:22AM
via free access

(Fig. 5b). These images also illustrate that the steepness of the boundary webs within the shaded spheres is greater in the case of the gray background, due to the contrast reversal around that sphere's bounding contour, and thus create a percept of greater relief, just as changes in the steepness of shading and texture gradients in the images of Fig. 29 control their perceived relief.

Koenderink *et al.* (2016) consider a different property of images around whose bounding contours relative contrast reverses continuously from dark-to-light to light-to-dark; e.g., their Fig. 1. They call the zero-contrast positions of such a contour a 'passage' and consider passages within the broader context that "many pictures are approximately piecewise uniform quilts…[Their] borders may occasionally get lost and sometimes pick up again, creating a '*passage*' that partly blends adjacent patches. This type of structure is widely discussed in treatises on painting technique" (p. 185). Its various painterly effects can also be explained using the mechanisms that are described in the current article.

The same sorts of multiple-scale boundary webs can be used to explain 3D percepts that are caused by *perspective* in a 2D picture, or the systematic reduction of spatial scale as a vanishing point on the horizon is approached. Perspective has been regularly used in paintings since the time of the Renaissance. The bigger scales, other things being equal, create a percept of a nearer surface, while the smaller scales create a percept of a surface that is farther away, as in the size-disparity correlation. The ambiguous depths that multiple-scale filters code is disambiguated to create percepts of relative depth by scale-to-depth and depth-to-scale maps, cooperative-competitive boundary interactions, and the depth-selective filling-in of surface representations under the control of the resultant boundaries, in much the same way as the LIGHTSHAFT model explains the percept of Fig. 29 (right).

3.8. Recognition Without Seeing, and Seeing to Reach

We are almost ready to focus more completely upon how various painters have exploited for their aesthetic goals different combinations of processes that control how our brains see. Before doing so, it is relevant to the goals of this article to ask: Why do we consciously see at all?

Many people, if they have thought about this basic issue at all, may conclude after reflection that 'we see things to recognize them'. These people include artists who have thought deeply about how they see the world. For example, Mann *et al.* (2016, p. 270) has written: "The main purpose of this conscious macular vision is to enable us to recognize," where *macular vision* is the kind of high acuity vision that is achieved using the fovea (Fig. 8). However, the fact that 'all boundaries are invisible' contradicts this claim, at least as a general explanation of why we consciously see, because these invisible boundaries enable us to consciously recognize many emergent structures

Downloaded from Brill.com04/06/2021 01:46:22AM
via free access

without seeing them; e.g., Fig. 2 (left), Figure 3, Fig. 4 (right, top row), and (see below) Fig. 41 (left).

Grossberg (2017) has proposed how a surface–shroud resonance may (Fig. 18) support conscious seeing when it is triggered between prestriate visual cortical area V4 and the posterior parietal cortex, or PPC, while propagating top-down to cortical areas such as V2 and V1, and bottom-up to cortical areas such as the prefrontal cortex, or PFC. This analysis proposes how such a surface–shroud resonance enables a conscious brain to control effective looking and reaching movements towards the unoccluded surfaces of overlapping objects (see Fig. 41, left, below). It is also proposed that conscious hearing is organized to enable effective vocalizing and speaking, and conscious feeling is organized to enable effective actions towards valued goal objects in the world. These conclusions are supported by explanations of psychological and neurobiological data from normal subjects and clinical patients that have no other explanations at present.

These conclusions do not deny that visible representations may also augment human abilities to recognize objects, and to enjoy the beauty of the world, after the evolutionary pressures for creating brain processes that could support conscious representations had already acted. It is interesting to contemplate from this perspective the following quote of Mann *et al.* (2017, pp. 274–275) about an experience that he had when painting after he became blind: "…I put ultramarine on a brush and painted the top right hand corner of the canvas and I had one of the most extraordinary sensations of my life. I saw the canvas go blue." It is of interest that Mann believed that we "see to recognize" even after he experienced the powerful link between seeing and acting.

4. Towards Neural Explanations of Painting Percepts

4.1. Claude Monet, Coarse Boundary Webs, and Uncertainty

Claude Monet, who lived from 1840 to 1926, was one of the founders of the French Impressionist movement, which derived its name from his painting called Impression, Soleil Levant (Impression, Sunrise). This painting was exhibited in 1874 in an exhibition that was organized by Monet and his associates as an alternative to the Salon de Paris. Monet's use of color significantly departed from the established Academic painting style that dominated the Salon de Paris. The surfaces of his paintings of natural scenes are built up from small brushstrokes. Using these discrete, but juxtaposed, color elements, Monet created complex surfaces that induce amodal boundary webs, which define the structures of object forms. These emergent forms can be seen and recognized even though the individual compartments from which they are composed are also often clearly visible when spatial attentional shrouds focus upon them.

Downloaded from Brill.com04/06/2021 01:46:22AM
via free access

Without a certain degree of orientational uncertainty embedded in the human visual system (Figs 23 and 25), long-range cooperative boundary completion and emergent segmentation would be nearly impossible in response to Monet's paintings. His paintings exemplify the paradoxical prediction that orientationally 'fuzzy' computations often lead to sharp segmentations (Fig. 27) because hardly any sharp edges are in them. Edges, forms, and the light that falls on them are perceived despite the fact that the paintings are irregular, noisy surfaces composed of varying densities of luminance and color patches.

Monet's brushstrokes induce boundary webs in which many of the compartments within a form are nearly equiluminant and there may be stronger boundary signals between forms. This combination facilitates color spreading within forms and better separation of brightness and color differences between forms. Figure 30 (upper row) shows four paintings from Monet's Rouen Cathedral series and their grayscale counterparts (Fig. 30, lower row). This famous series, created in the 1890s, includes over 30 paintings of similar views of the Rouen Cathedral at different times of day. The grayscale versions of these paintings demonstrate the near equiluminance of the brushstrokes within forms and places in which brightness and color differences significantly influence the groupings that differentiate between forms, including the differentiation between the cathedral and the sky. Monet took advantage of the statistical uncertainty that our visual system is designed to overcome to create work in which both color and luminance play a role in creating form, rather than luminance alone, as in the first and fourth images from left to right. Taking the color out of Monet's paintings removes some critical closed boundaries that allow us to segregate forms. It is because the boundary system pools both chromatic color and brightness signals at complex cells that boundaries can remain that may be due almost entirely to a color difference, without a corresponding achromatic brightness difference, as in the fourth image to the right in the upper row of Fig. 30.

Figure 31 contains images of two paintings from Monet's Rouen cathedral series that illustrate how different lighting of the same object can lead to dramatically different percepts, in part by causing different boundary webs. The left painting captures the lighting on the cathedral facade while the sun is setting, whereas the right painting does so in full sunlight. The right painting generates a more depthful percept than does the left one. This is due to multiple factors interacting together. One factor is that the colors in the left painting are more equiluminant, and the ensuing boundary webs coarser and more uniform. The strong gradients of light that are evident in the right painting, including effects of sharp architectural features and shadows, are not seen in the left painting. Note in addition that, due to the greater range of contrasts in the right painting, many more nearly horizontal boundaries are clearly seen to be occluded by vertical boundaries. Such occlusions often occur at visual

Downloaded from Brill.com04/06/2021 01:46:22AM
via free access

Figure 30. (Upper row) A set of four paintings by Monet of the Rouen cathedral at different times of day (all painted in 1894). (Lower row) Achromatic rendering of the four Monet paintings, illustrating where nearly equiluminant parts of the paintings occur, and thus which groupings depend primarily upon color differences, rather than luminance differences. See text for details.

Figure 31. Two paintings of the Rouen cathedral by Monet at different times of day (also painted in 1894). See text for details.

Downloaded from Brill.com04/06/2021 01:46:22AM
via free access

features that are called T-junctions. Such boundary occlusions can generate percepts of depth in the absence of any other visual cues. An explanation of how this may happen is given in Sect. 4.7 when figure–ground separation is discussed. As will be shown below, T-junctions can help to trigger figure–ground separation using basic properties of the boundary completion and surface filling-in processes, without the need for specialized T-junction detectors.

Kass *et al.* (2015) have discussed the Rouen cathedral series from the perspective of how prototypes are formed from a set of exemplars, or instances, of an object that humans are learning to recognize. To that end, the authors compare different instances in the Rouen Cathedral series with an average of these paintings (their Fig. 4). The authors note that these instances "would result in a more stable concept and robust mental representation of the cathedral in the minds of those viewing the series" (p. 145), and commented about how painting a series of instances influenced later developments in several artistic movements, including Conceptual Art. As noted in Sect. 2.5, however, a prototype of an object is not just an average of its exemplars. Rather, a prototype is a *critical feature pattern* that is discovered by incremental learning of a series of object exemplars that activate the same recognition category. This critical feature pattern also constitutes the feature set upon which attention focuses when viewing an exemplar of the object. Thus, although incremental viewing of a series of Rouen Cathedral exemplars may indeed result in learning 'a more stable concept' of the Cathedral, that concept is not just the average of the exemplars. Rather, the prototype helps to focus selective attention upon the features that viewers use to identify and appreciate the object after learning about it.

4.2. Boundary Webs from Color Field Paintings of Jules Olitski that "Hang Like a Cloud"

Nearly a century after the Impressionists did their work, Color Field painting began to flourish in the United States. This term is generally used to describe the canonical Color Field painters (e.g., Mark Rothko, Barnett Newman) who were part of the larger Abstract Expressionist movement and their successors, who painted in the 1960s and 1970s (e.g., Jules Olitski, Helen Frankenthaler). The Color Field painters were primarily concerned with exploring the pure emotional power of color. Thus, color—particularly contained within large, undisrupted surfaces—is a dominant element in these paintings. Jules Olitski created several such paintings when he lived and worked in New York City. By 1965, Olitski developed a technique in which he sprayed paint onto unprimed canvases and created some of his most famous paintings, which are accordingly called 'spray paintings' (Fig. 32).

Downloaded from Brill.com04/06/2021 01:46:22AM
via free access

Figure 32. Four paintings by Jules Olitski. *Julius and His Friends (1967)* (upper left), *Instant Loveland* (1968) (lower left), *Lysander-1* (1970) (upper right), *Comprehensive Dream* (1965) (lower right). See text for details.

Spraying paint onto a canvas in fine mists gave Olitski control over the color density on the canvas. In contrast to many of Monet's paintings, in which the viewer can discern a multitude of individual brushstrokes contained within a coarse boundary web, it is impossible to visually perceive discrete colored units within the fine boundary webs in Olitski's spray paintings, as is also the case in the example of a shaded ellipse (Fig. 29, left panel). Unlike the boundary webs for the shaded ellipse, however, the boundary webs that spread over the surface of Olitski's spray paintings present a sense of ambiguous depth to the viewer. It is easy to imagine that one is staring into a space filled with colored fog or into the sky during a sunset free of discrete clouds when looking at these paintings. Olitski intentionally created this effect, writing that: "When the conception of internal form is governed by edge, color … appears to remain on or above the surface. I think … of color as being seen in and throughout, not solely on, the surface" (Riggs, 1997). Others recognized this quality, as well. The art critic Rosalind Krauss wrote: "…the very seeing of the painting in all its literalness poses a question about where the surface is. To see Olitski's color means to see the surface itself as elusive and unaligned" (Riggs, 1997). Occasional sharp spatial discontinuities in some of these paintings awaken in the viewer an acknowledgement that

Downloaded from Brill.com04/06/2021 01:46:22AM
via free access

drawing plays little or no role in creating the surface percept that is induced by most of the painting.

The principles of FACADE theory and the 3D LAMINART model (Sect. 2.1) clarify why the structure of Olitski's spray paintings create the sense that one is looking into a mist of color within a three-dimensional space. These models highlight the importance of closed boundary contours to create percepts of surfaces that are restricted to specific depth planes (Figs 11 and 12). Within the mists of color in Olitski's paintings, sharp, closed boundary contours are not perceptible. Instead, very gradual chromatic and luminance gradients coexist, thereby enabling multiple scales to form boundary webs, and to thereby capture color shading on multiple depth-selective surfaces whose gradients often coexist at similar positions. Unlike the example of the shaded ellipse (Fig. 29, left), these gradients and their boundary webs do not vary in either orientation or scale in a systematic way across space. Thus, we do not perceive the shaded surfaces of these paintings to be clearly contained within one depth plane. In this way, Olitski was able to achieve color that is "seen in and throughout".

Another factor that may contribute to the percept of amorphous depth is the gradual brightness and color gradations on the surface. The property of *proximity-luminance covariance* (Dosher *et al.*, 1986; Egusa, 1983) describes the fact that increasing brightness can make a surface look closer. A similar property obtains when brighter Kanizsa squares look closer. Grossberg (2014) provides an explanation of this kind of effect using mechanisms that also help to ensure complementary consistency (Sect. 3.1) and figure–ground separation (see Sect. 4.8 below).

It remains to ask: Why does the web within the painting appear as a 'colored mist' in an ambiguous three-dimensional space that is separate from the rectangular boundary of the painting? Why does the web appear as mist rather than as a surface that is 'attached' to the rectangular canvas like the boundary web of the shaded ellipse in Fig. 29 (left) seems to be attached to its bounding contour? One factor is that the orientations of the boundary webs induced by the interior of the painting tend to be parallel to the isophotes of each surface. However, these isophotes are not parallel to, and indeed are incongruent with, in both orientation and scale, the rectangular boundary of the canvas. This is unlike the case of the shaded ellipse, whose boundary webs are consistent with the bounding contour of the ellipse. These incongruities in Olitski's paintings disrupt the nesting and attachment of the boundary web structures within the painting with the bounding contour of the rectangular frame of the painting, and helps to explain why the color in Olitski's spray paintings "hangs like a cloud, but does not lose its shape" (Anonymous, 2016).

The only visible boundary contours within Olitski's spray paintings are the edges of the canvas itself and the framing lines he introduced along the edges of some of the paintings, which the famous American critic Clement

Downloaded from Brill.com04/06/2021 01:46:22AM
via free access

Greenberg named *selvage* (Olitski, 1994) (Fig. 32, left). The word selvage can be used in many contexts (e.g., textiles, manufacturing), but generally refers to an edge added to a material to finish it, prevent it from fraying, or allow it to be handled. In contrast to the mists of color, the depth plane of the selvage in Olitski's paintings is unambiguous. The selvage is always perceived as being on the surface of canvas due to the sharp boundaries that contain it, and thereby can 'visually ground' the viewer in the depth plane of the picture's frame.

4.3. Watercolor Illusion, Jo Baer, and Mach Bands

Before turning to a discussion of figure–ground separation due to partially occluded boundaries, let us consider other examples of how boundary webs with multiple spatial scales can generate depthful percepts, even if there are no occluded boundaries. This is first illustrated using a display that induces the watercolor illusion (Fig. 33) that has been intensely studied by Baingio Pinna and his colleagues since the 1980s (http://www.scholarpedia.org/article/Watercolor_illusion). Variants of the watercolor illusion have also been featured in the works of visual artists, notably paintings of Jo Baer, since the 1960s. After explaining how the watercolor illusion, including its depthful figure–ground properties, can be elicited, some of Jo Baer's paintings that exhibit properties of the watercolor illusion will also be discussed.

Inspection of Fig. 33 induces the watercolor illusion percept because a more contrastive and undulating purple band abuts a less contrastive yellow band. The most obvious property of the illusion is the yellow color that fills the entire interior of the region surrounded by these boundaries. Pinna and Grossberg (2005) have explained this percept, and many of its variants, using

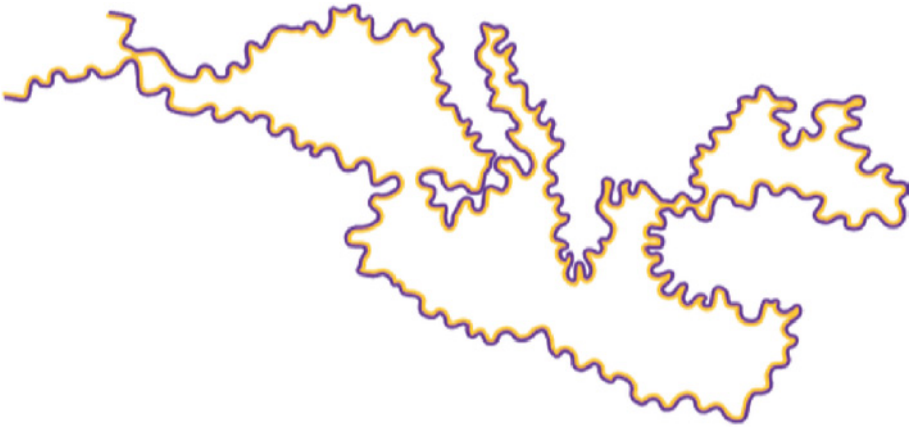


Figure 33. An example of the watercolor illusion due to Baingio Pinna. (Reprinted with permission from Pinna and Reeves, 2006.) See text for details.

Downloaded from Brill.com04/06/2021 01:46:22AM
via free access

properties of spatial competition, or the first competitive stage (Fig. 5), of the boundary contour system. In particular, the boundary that is formed between the purple band and the white background is stronger than the boundary that is formed between the yellow band and the white background. The boundary that is formed at the purple–yellow contour is also weaker than that between the purple–white contour, but stronger than that between the yellow–white contour. Due to spatial competition, the stronger boundaries inhibit the yellow–white boundary more than conversely, thereby weakening the yellow–white boundary and allowing yellow color to spill out beyond it to fill-in the entire surface within. The undulating shape of the contours creates longer boundaries per unit area through which color can spread, thereby intensifying the effect. Basic properties of perceptual grouping, notably of hierarchical resolution of uncertainty, hereby help to explain the watercolor illusion in this and many of its variations.

The watercolor illusion can also generate a 3D percept. Just as in response to a shaded ellipse, there is formed in response to the image in Fig. 33 a spatial array of successively weaker boundary strengths, from purple–white to purple–yellow to yellow°white, as the distance increases from the most contrastive purple–white edge of the display. These successively weaker boundary responses constitute a boundary web, albeit a more spatially discrete one than the one formed in response to a shaded ellipse. When this multiple-scale boundary web traps the corresponding colors (Fig. 12), it can generate a rounded appearance of the percept, with the filled-in yellow surface looking a little closer, using the same size–disparity correlation mechanisms that also play a role in creating chiaroscuro percepts.

Jo Baer demonstrated the watercolor illusion in some of her paintings from the 1960s in New York City. Baer was at that time interested in working with black, white, and color in a minimalist style; that is, without direct reference to anything outside the painting (Boersma, 1995). Her series of three paintings called *Primary Light Group: Red, Green, Blue*, shown in Fig. 34, exhibits watercolor properties, and Baer and others in the art world recognized the role of active perception that is induced by these paintings. For example, in a catalogue for a Baer exhibit that included paintings of like those in Fig. 34, David Elliot wrote: “The positioning of colour in this way … induced certain optical effects which were not part of the physical make-up of the painting. The existence of these effects was wholly dependent upon the perceptual responses of the viewer” (“Jo Baer: Stations of the Spectrum”, n.d.).

Baer studied biology at the University of Washington, Seattle and physiological psychology at the New School in New York, which equipped her to understand why the colors in her paintings were different from the colors that were perceived. In an interview published in 1995 in *BOMB Magazine* (Boersma, 1995), Baer said: “I was always curious why the color on the palette was

Downloaded from Brill.com04/06/2021 01:46:22AM
via free access

Figure 34. The three paintings by Joe Baer called *Primary Light Group: Red, Green, Blue* (1964–'65). See text for details.

different than the color on the painting. I knew what I wanted something to look like, and I found that the means to do it were so different than the end result. And then I was very pleased to discover the reasons why". Baer wrote about some of these reasons in an article for *Aspen Magazine* (1970–'71): "Tucked in between the white and black, the narrow color band gains a vast brightness due to…distinct edge effects working at the color interfaces… [these]…brightness contrast effects push the color band…higher into luminosity through a physiological neural phenomenon called *Mach bands*" (Baer, 1970–'71). Baer wrote that she "intuitively fashioned" the paintings in the style of those shown in Fig. 34 with her knowledge of Mach bands and retinal physiology data "somewhere in mind" (Baer, 1970–'71), as reflected in the title of the painting shown in Fig. 34, which includes the word "light" in the title, and her description of these paintings as containing "primary colors of light" ("Jo Baer, primary light group: red, green, blue", n.d.).

The Mach bands to which Baer refers is an optical illusion that enhances the contrast between abutting regions composed of slightly different luminances, just as the black band can enhance the contrast of the less luminous color band, and the color band can influence the perceived contrast of the white interior.

However, it is not local contrast effects like Mach Bands that create the watercolor properties of Baer's paintings. Rather, it is a combination of spatial competition within the boundary stream followed by surface filling-in within the surface stream (Fig. 9). Indeed, when viewing the paintings in Fig. 34, the white central area of the painting on the left appears to have a red/pink tint, in the middle painting appears to have a green tint, and in the painting on the right appears to have a blue/purple tint. The same spatial competition that helps to explain the watercolor illusion explains these percepts as well because the luminance contrast between the dark band and the white wall is larger than between the colored band and the white interior of the painting. As a result,

Downloaded from Brill.com04/06/2021 01:46:22AM
via free access

the boundary between the dark band and the white wall is stronger than the boundary between the colored band and the white interior. Spatial competition between the stronger, white–black boundary and the weaker color–white boundary weakens the color–white boundary, leading to color spreading, or surface filling-in, within the interior white surface of each painting. When viewing several of these paintings side-by-side in the same room, as in Fig. 34, the effects of color spreading are even more striking.

4.4. Lightness Anchoring and Self-Luminosity

Some paintings, such as those of Ross Bleckner that are discussed in Sect. 4.6, generate percepts of self-luminosity. How are self-luminous percepts generated just using paint? This section and the next one describe two different neural mechanisms whereby self-luminosity can be induced. Both of these mechanisms are exploited in many of Bleckner's paintings.

As discussed in Sects 2.1 and 3.5, reliable object colors may be perceived under widely different lighting conditions because the visual system can 'discount the illuminant' to compute surface properties based on changes in reflectance across boundaries. As noted in Sects 3.4 and 3.5 during the discussion of how boundary interactions can lead to neon color spreading, on-center off-surround networks whose cells obey shunting dynamics are capable of contrast normalization. Contrast normalization can also occur when these networks process the luminances and colors within a scene or picture. Contrast normalization leads to discounting of the illuminant by enabling cell activities to compute reflectances, or the *relative* amounts of light reflected from a surface to the eyes within each wavelength of light. This process does not, however, compute the *absolute* perceived brightnesses of a scene. A more sophisticated way of saying this is that this process does not use the *full dynamic range* of the cells that estimate surface brightnesses and colors. How does the brain use the full dynamic range of cells to compute estimates of the absolute brightnesses of scenes and of specific objects within them?

There is a rich and often paradoxical experimental literature about how the brain sees the absolute brightnesses of scenes and the objects within them. This literature provides examples of self-luminosity among the many percepts that can be generated through this process. Other famous effects include: the Gelb effect, whereby a black surface can look white when it is intensely illuminated; and the Area effect that is often demonstrated by placing a viewing subject's head within a dome that is divided into two regions. When the highest luminance area occupies more than half of the visual field, it appears white while the darker part looks gray. When the darker luminance area occupies more than half of the visual field, it tends to look increasingly white, while the lighter area appears to be self-luminous. Alan Gilchrist and his colleagues (e.g., Gilchrist *et al*., 1999) have been leaders in demonstrating such effects,

Downloaded from Brill.com04/06/2021 01:46:22AM
via free access

which they attribute to a process of *anchoring*; namely, the process whereby the brain uses its full dynamic range to compute brightnesses and colors. Great predecessors in the study of how anchoring may work include Helmholtz (1866) and Wallach (1948, 1976). Indeed, Wallach proposed the highest-luminance-as-white (HLAW) rules whereby the perceptual quality 'white' is assigned to the highest luminance in a scene, and gray values of less luminous surface regions are assigned relative to the white standard. Although this rule works in many situations, there are many other situations in which it leads to the wrong answer.

The anchored Filling-In Lightness Model (aFILM) of Grossberg and Hong (2006) and Hong and Grossberg (2004) describes an alternative neural model of how the full dynamic range of neurons is used in processing the absolute lightness of scenes, and along the way explains and simulates all of the experimental anchoring effects that are reviewed in Gilchrist *et al*. (1999), including a mechanistic explanation of how percepts of self-luminosity occur. aFILM revises the Wallach HLAW rule by using a *blurred-highest-contrast-as-white* (BHCAW) rule that overcomes problems with the Wallach rule and, when embedded in suitable model processing stages (Fig. 35), can explain and simulate many anchoring data.

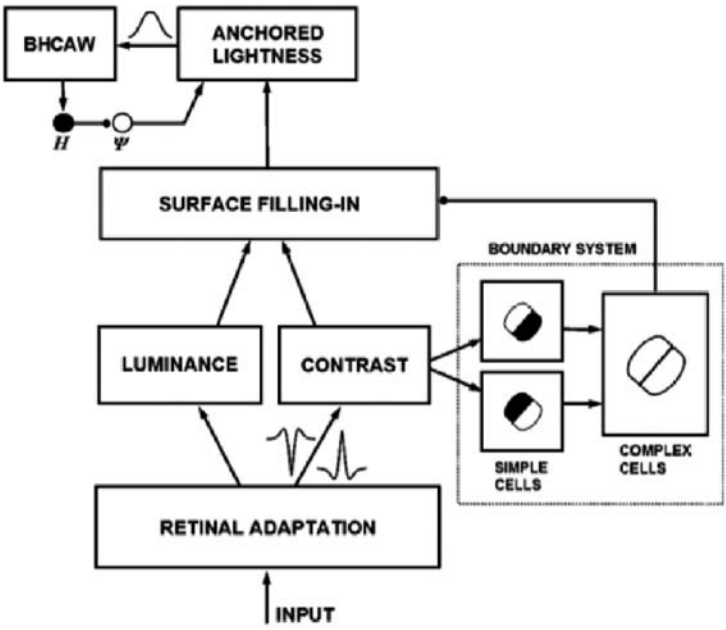


Figure 35. Macrocircuit of the aFILM model for lightness anchoring The final BHCAW stage of the model causes the lightness anchoring that can lead to percepts of self-luminance, among other striking effects. See text for details. (Reprinted with permission from Grossberg and Hong, 2006).

Downloaded from Brill.com04/06/2021 01:46:22AM
via free access

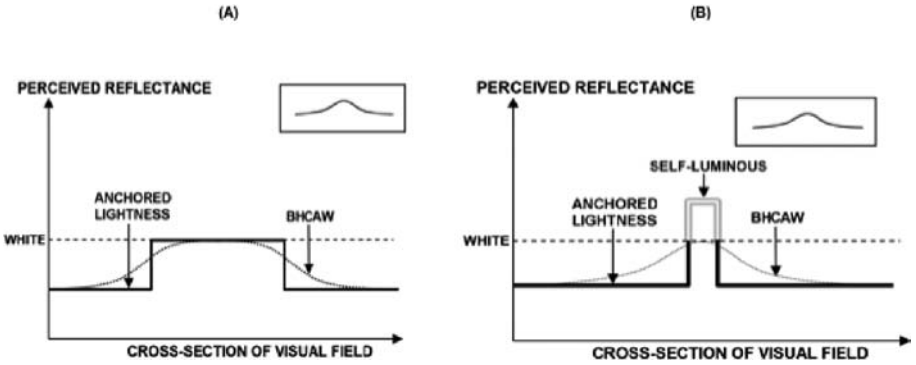


Figure 36. How the aFILM model explains self-luminous percepts using a blurred-highest-contrast-as-white (BHCAW) kernel. Whether self-luminosity is perceived or not due to the model mechanism depends upon whether the scenic region is as big or bigger than the blurring kernel, as in (A), or smaller than the blurring kernel, as in (B). See text for details. (Reprinted with permission from Grossberg and Hong, 2006).

Figure 36A illustrates how the BHCAW rule anchors whiteness and, with it, the gray shades in a scene, whereas Fig. 36B proposes how a percept of self-luminosity may be generated in response to different scenic conditions. The aFILM model proposes that a scene is anchored to the highest luminance that is computed after it is processed through a blurring kernel. If the area of highest luminance in the scene is the same size or larger than the blurring kernel, as occurs in Fig. 36A, then this region is perceived as white. However, if the area of highest luminance in the scene is smaller than the blurring kernel, as occurs in Fig. 36B, then the scene is again anchored to the blurred highest luminance as white, but this corresponds to a luminance that is lower than the highest luminance in the scene. This allows the highest luminance in the scene to be perceived as being 'brighter than white', or self-luminous.

Self-luminosity is one of many percepts that the highest-luminance-as-white (HLAW) model cannot explain. This is because self-luminous regions are typically perceived as having a lightness value that is higher than white, but the HLAW model sets the highest luminance in a stimulus to equal white.

4.5. Boundary Webs, Glare, Double Brilliant Illusion, and Gloss

The spatial context in which a luminance occurs within a scene can also create a percept of self-luminosity, even without the intervention of the BHCAW rule, notably when a region of prescribed luminance is surrounded by a suitable form-selective luminance gradient. This can be seen in examples such as the glare effect (Zavagno, 1999; Zavagno *et al*., 2004), the double brilliant illusion (Bressan, 2001), and percepts of gloss (Beck and Prazdny, 1981), which are illustrated in Fig. 37.

Downloaded from Brill.com04/06/2021 01:46:22AM
via free access

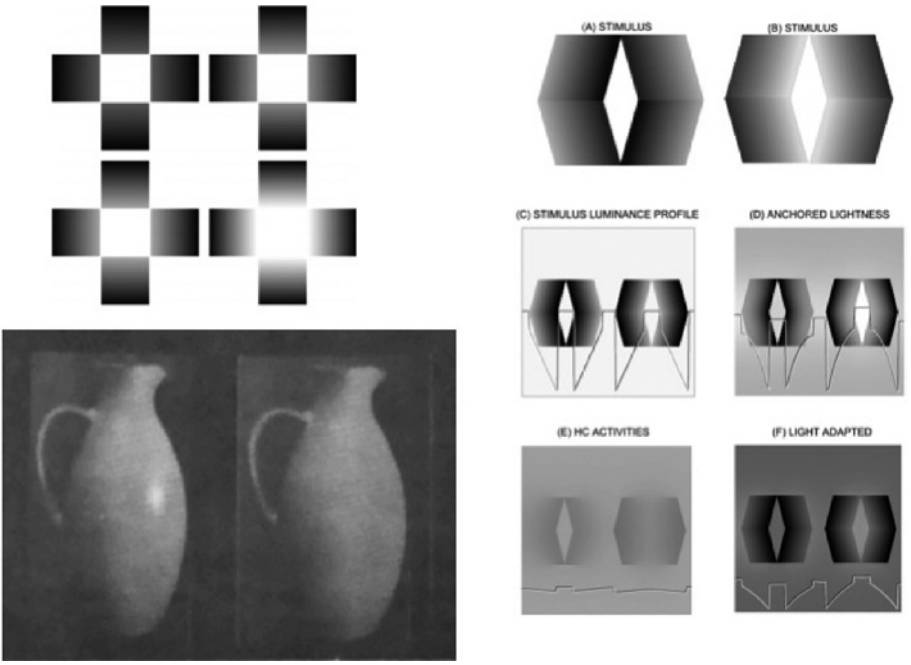


Figure 37. Stimuli that induce a percept of glare (left, upper row), of the double brilliant illusion (upper row, right column), and of gloss due to attachment of a highlight to a matte vase (left, lower row). Model computer simulation of the double brilliant illusion (right, lower row). (Figures reprinted with permission from Zavagno, 1999; Bressan, 2001; Beck and Prazdny, 1981; and Grossberg and Hong, 2006, in the reported order).

In the glare effect (Fig. 37, left, upper row), the interior white region of all four figures is the same, but an increasingly strong luminance gradient surrounds this region as one proceeds from the left figure in the first row to the right figure in the second row. The percept in response to the strongest gradient in the lower right image is one of a glowing self-luminosity, with lesser effects occurring in response to the weaker luminance gradients. This effect may be explained as the result of the boundary webs that are generated in response to the luminance gradients and how they control the filling-in of lightness within themselves and abutting regions. In particular, such a boundary web is like a continuous version of the boundary web that enables color spreading to occur out of the weakest boundary in the watercolor illusion (Fig. 33) or Joe Baer's paintings (Fig. 34). Due to these boundary webs, more lightness can spread into the central square as the steepness of the boundary gradient strengths increases.

Koenderink *et al.* (2016) describe the glare percept that is generated by Fig. 37 (left, upper row) as being "whiter than white"; see their Fig. 5 (left). This description may cause confusion from a mechanistic point of view, because

Downloaded from Brill.com04/06/2021 01:46:22AM
via free access

those self-luminous percepts that may be thought of as 'whiter than white' due to the BHCAW mechanism (Fig. 36, right) seem to be caused by different mechanisms than those which create the glare effect.

A similar way of thinking helps to explain the double brilliant illusion that is illustrated in Fig. 37 (right, upper row), along with a schematic in Fig. 37 (right, bottom two rows) of how the aFILM model explains and simulates it using mathematically rigorous embodiments of these concepts.

Finally, consider the gloss effect that is perceived in response to Fig. 37 (left, lower row). Here two almost identical vases are shown. In the left vase, a highlight has been added to a local region of the vase surface. Remarkably, this local change makes the entire vase look glossy compared with the matte vase to the right from which it was constructed. This glossy percept has been explained (Grossberg and Mingolla, 1987) by the way in which the boundary web that is induced by the highlight is assimilated into the boundary web that is generated by the rest of the vase's form. The lightness within the highlight can hereby spread to abutting areas of the vase, rendering its appearance glossy. As with the case of the shaded ellipse, this boundary web is invisible and is seen only through the gradients of surface lightnesses that it traps within its compartments. In support of this explanation, Beck and Prazdny (1981) showed that changing the gradual luminance gradient of the highlight to a spatially abrupt one, or changing its relative orientation with respect to the rest of the vase, can eliminate the glossy percept.

4.6. Ross Bleckner's Self-Luminous Paintings

Ross Bleckner, who has worked as a painter in New York for the past several decades, has described his paintings as attempts to "eke out of a formal code a maximum amount of light" (Rankin, 1987). Bleckner does this, not by painting large surface areas with high reflectances or bright colors, but rather by creating compositions of small, star-like, circular regions that are perceived as self-luminous (Fig. 38). At least three interacting properties of the paintings in Fig. 38 contribute to their self-luminosity: high luminance areas relative to a dark background, their small size, and the smooth luminance gradient that surrounds many of them. The large luminance difference enhances the effects of brightness contrast; the small size can create self-luminosity via the BHCAW rule (Fig. 36); and the luminance gradient can enhance self-luminosity using boundary webs, much as in in percepts like the glare and double-brilliant percepts that are induced by Fig. 37. When the points of light surrounded by a boundary web are magnified, as in the center of Fig. 38, they still appear to be highly self-luminous, despite their larger size. This percept illustrates the independence of the BHCAW rule and the self-luminosity

Downloaded from Brill.com04/06/2021 01:46:22AM
via free access

effects of boundary webs, as in Fig. 37, since the boundary web is still present even when the figure is too large for the BHCAW rule to cause self-luminosity.

The images in Fig. 39, a painting (right) and a photo (left) by Bleckner, further illustrate the power of surrounding a high-luminance point with a dense boundary web. In the image on the left, the large yellow circular regions appear to be self-luminous. In contrast, the yellow circular regions in the painting on the right are surrounded by sharp, black contours that isolate them from being enhanced by spreading color from surrounding regions, since sharp boundaries act as filling-in barriers, as in the percept of neon color spreading (Fig. 24). Even the smallest high-luminance points in the painting in Fig. 39 are thus not as self-luminous as the points within Fig. 38, which receive the added benefits of self-luminosity due to boundary webs.

4.7. Gene Davis's Depth from Proximity-Luminance Covariance and Color Assimilation

Depth can be conveyed through different means. Binocular disparity between left and right eye views of an object in depth is one important cue. Size is another important factor, if only because of the size–disparity correlation (Sect. 3.6), which summarizes the property that larger objects tend to appear closer, while smaller objects tend to appear farther away. Shading can also create a three-dimensional percept by generating multiple-scale form-sensitive boundary webs, as another manifestation of the size-disparity correlation, with examples in the works of Monet (Figs 30 and 31), Olitski (Fig. 32), and Bleckner (Figs 38 and 39) of how very different painterly percepts can be generated by different boundary web configurations and how they organize the filling-in of surface brightness and color.

The paintings of Gene Davis provide additional examples of how boundary groupings can influence painterly percepts that can induce percepts of different relative depths. Although Gene Davis was a journalist, he began to paint seriously in Washington D.C. in 1950. His vertical stripe compositions (Fig. 40) are the paintings for which he is most well known and they also comprise the vast majority of his life's work. These paintings are interesting stimuli from the perspective of scene stratification because many of them do not contain size differences, shading, or recognizable objects, although some stripe paintings did contain stripes with shading and/or physical size differences. The painting in Fig. 40 (left, upper row) is called *Black Popcorn*. Despite the lack of explicit depth cues, the 'brightest'stripes—namely, the four intermingled yellow and pale blue stripes located approximately one-third of the length of the canvas from the left—appear to be in the nearest depth plane. In contrast, in *Flamingo* (Fig. 40, right, upper row), both individual stripes and stripe groupings can be perceived as nearer or further in depth,

Downloaded from Brill.com04/06/2021 01:46:22AM
via free access

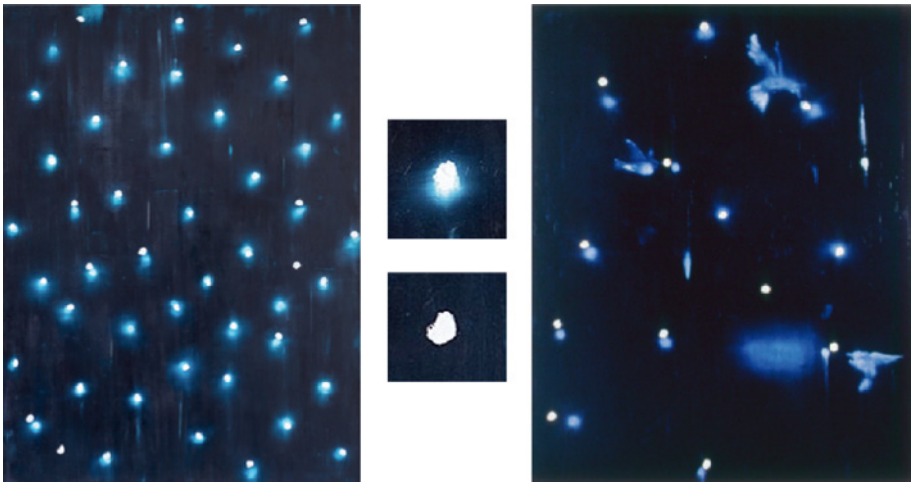


Figure 38. Two paintings by Ross Bleckner that exhibit self-luminosity. (left) *Galaxy Painting* (1993). (right) *Galaxy with Birds* (1993). (center) details from *Galaxy Painting*. See text for details.

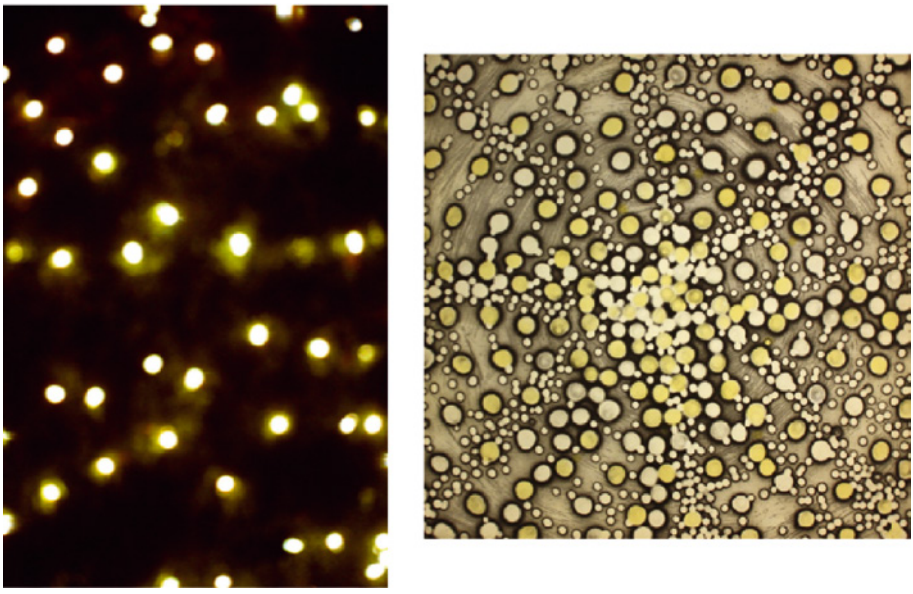


Figure 39. A painting (right, *Insertion Sequence*, 2002) and a photograph (left, *Untitled)* by Ross Bleckner that illustrate how luminance gradients, or the absence thereof, may or may not support self-luminous percepts. See text for details.

Downloaded from Brill.com04/06/2021 01:46:22AM
via free access

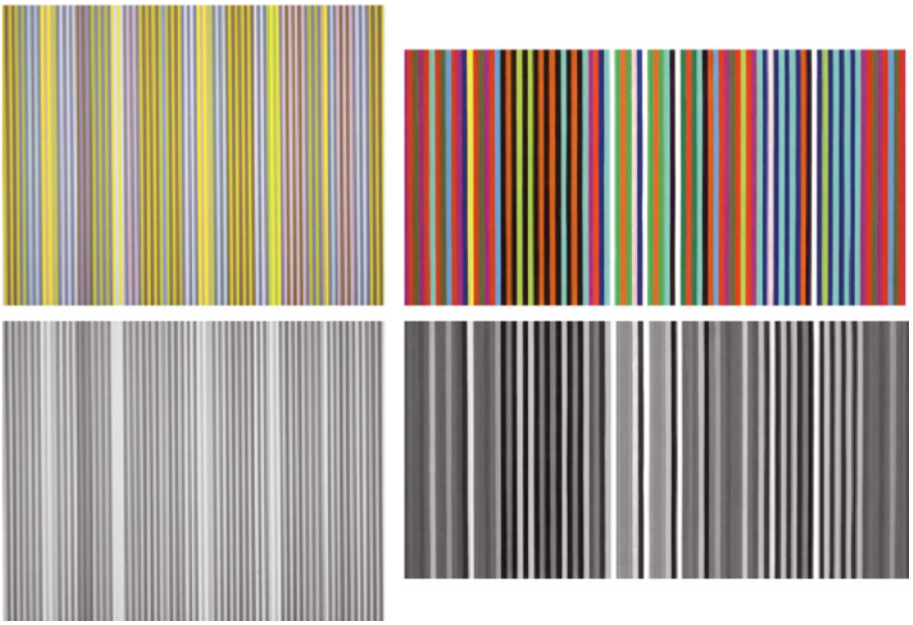


Figure 40. Two paintings by Gene Davis, *Black Popcorn* (left, upper row), and *Flamingo* (right, upper row), and their achromatic versions (lower row). The different combinations of colors and luminances in the two paintings influence their perceptual groupings and color filling-in dynamics and, with them, the attentional shrouds that they induce, along with percepts of nearer vs. further regions of the paintings. See text for details.

again with brighter stripes appearing closer. Some of the neural mechanisms that influence these depth percepts are now discussed.

Color similarities and/or near-equiluminance between stripes in these paintings can influence whether spatial attention is drawn to individual stripes or groups of stripes. The achromatic versions of the two paintings more clearly show regions where color assimilation is facilitated, notably the much smaller differences in luminance of the stripes within *Black Popcorn* (Fig. 40, left, lower row) than of *Flamingo* (Fig. 40, right, lower row). The boundaries between stripes that have low luminance contrast are weak, thereby facilitating increased surface filling-in across several stripes. Such color assimilation calls to mind the watercolor illusion (Fig. 33) and the paintings of Jo Baer (Fig. 34). As a result of weak boundaries and color assimilation, an attentional shroud (Figs 17 and 18) that is larger than one stripe is more easily formed during viewing of *Black Popcorn*, thereby enabling viewers to perceive multiple stripes as one object under the same shroud.

In *Flamingo*, individual bright stripes tend to be isolated between individual or groups of dimmer stripes and the luminance and color of adjacent stripes

Downloaded from Brill.com04/06/2021 01:46:22AM
via free access

vary more across space compared to those in *Black Popcorn*. As discussed in the context of multiple-scale boundary webs and proximity–luminance covariance in Sects 3.6 and 4.2, respectively, both relative size and relative brightness influence depth percepts. Larger objects tend to appear closer to the viewer because they more strongly activate larger spatial scales in the visual system. Because of this, the grouping of bright stripes into larger units results in larger filters being sensitive to them, which can cause them to appear closer to the viewer because of the size–disparity correlation. Their relative brightness can, in addition, cooperate with this effect to make them appear closer because of proximity–luminance covariance. These properties help to explain why the four-stripe unit in the left part of *Black Popcorn* appears to be nearer than the individual bright white stripes in *Flamingo*, despite the fact that the latter stripes appear to be brighter than any individual stripe in *Black Popcorn*. These *Flamingo* stripes can benefit from proximity–luminance covariance, but not from the size–disparity correlation. Thus, despite the apparent simplicity of the iterated vertical stripes in the paintings of Gene Davis, the interactive effects of relative boundary strength, color assimilation, and spatial attention influence their percepts of relative depth using interactions between the neural processes that create both the size-disparity correlation and proximity-luminance covariance.

4.8. Figure–Ground Separation: Seeing and Recognizing Partially Occluded Objects

Many of the percepts that have already been discussed depend upon boundaries that can be completed directly from their image contrasts. When boundaries overlap as in 2D pictures of partially occluded objects, as they do vividly in the Monet painting in Fig. 31 (right), then additional cues to depth become available and influence what is consciously seen and recognized. Figure 41 (left) illustrates this fact in a simple example. This figure is composed of three abutting rectangles in a 2D picture. It is hard, however, not to perceive it as a

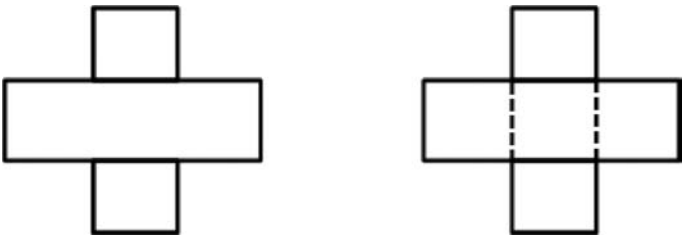


Figure 41. (Left) Three abutting rectangles create a compelling percept of a horizontal rectangle that partially occludes a further vertical rectangle. (Right) The text explains how the unoccluded regions of the vertical rectangle can be amodally completed behind the horizontal rectangle to create a percept of a partially occluded vertical rectangle. See text for details.

Downloaded from Brill.com04/06/2021 01:46:22AM
via free access

large horizontal rectangle that partially occludes a large vertical rectangle that lies 'behind' the horizontal rectangle (Fig. 41, right). Examples such as these provide useful probes to clarify how brain designs for 3D vision in the natural world have enabled 2D pictures to generate representations of the 3D world, and hereby created a foundation for all pictorial art, movies, TV, and all other cultural advances that depend upon this possibility.

The implications of this process for paintings are thus of enormous importance. Answering the following question therefore cannot be avoided: How does the simple 2D picture in Fig. 41 (left) give rise to a 3D percept? It may at first be thought that this is due to what is learned by many experiences with partially occluded objects. However, Kanizsa (1979) has shown with many compelling examples (e.g., Fig. 2, right, and Fig. 4, lower row) that many perceptual properties persist even when viewing unfamiliar pictures. In fact, percept in Fig. 41 (left) follows directly from basic properties of boundary grouping that have already been reviewed. This fact helps to answer the question: How might such a seemingly sophisticated process like figure–ground perception have arisen during evolution? FACADE theory and the 3D LAMINART model clarify how the key properties of figure–ground perception arise from basic properties of perceptual grouping and complementary consistency (Sect. 3.1).

To get started, consider one of the T-junctions in Fig. 41 (left); that is, a place where a horizontal boundary (the top of the T) is intersected by a vertical boundary (its stem), as in the top left image of Fig. 42a. The top middle image of Fig. 42a depicts the long-range oriented cooperation that activates bipole grouping cells, and the shorter-range competition that inhibits nearby bipole cells as part of the grouping selection process (also see Fig. 21). Consider how such grouping cells respond when they are centered where the top and stem join. We first consider bipole cells that have the orientational preference of the top (horizontal) of the T, and then bipole cells that have the orientational preference of the stem (vertical) of the T. As in the discussion of neon color spreading (Sects 3.4 and 3.5), a bipole cell that has the same orientation as the top receives excitatory inputs to both sides of its receptive field, and can thus strongly inhibit nearby bipole cells that respond preferentially to different orientations. A bipole cell that has the same vertical orientation as the stem gets excitatory inputs from only one side of its receptive field, so can either not respond at all, or can respond at best weakly. As a result, just as in the example of neon color spreading, an end gap is created in the stem boundary near where it intersects the top boundary, as in the top right image of Fig. 42a. Color can then spread across the end gap to both sides of this boundary, as it does during neon color spreading. This process enables figure–ground perception to begin when it interacts with processes leading to complementary consistency (Sect. 3.1), in a manner that is now explained.

Downloaded from Brill.com04/06/2021 01:46:22AM
via free access

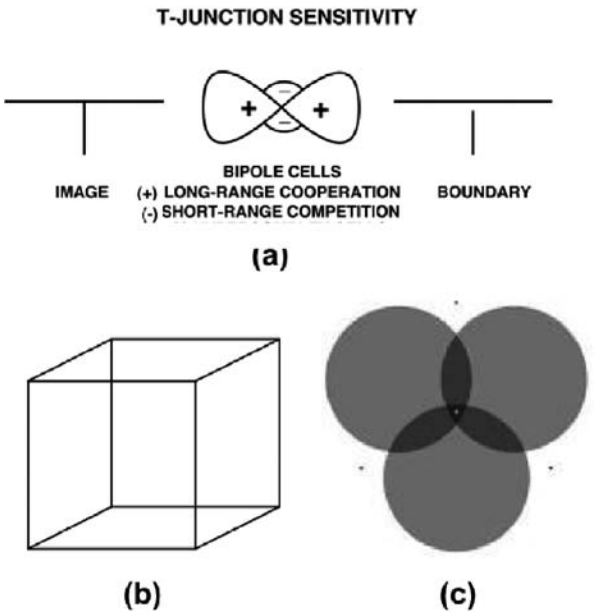


Figure 42. (a) These three images describe how, in response to a T-junction in an image, bipole grouping can cause an end gap to occur at the end of the stem boundary. This occurs because the horizontally-oriented bipole cell bodies near where the top of the T touches its stem get excitatory inputs from both branches of their receptive fields (middle image). In contrast, the vertically-oriented bipole cells at this intersection get excitatory inputs from only one branch of their receptive fields. As a result, short-range competition is much stronger from the vertical bipoles to the horizontal bipoles than conversely, and thereby causes an end gap to form in the stem boundary near where the top and stem of the T intersect (right image). (b) The Necker cube can be perceived as a 3D cube in two different ways that oscillate bistably through time. (c) When spatial attention is focused upon each circular disk, it appears darker and nearer. (Figure (c) is reprinted from Tse, 2005, with permission).

4.9. Complementary Consistency: Surface Contours and Eye Movements

Multiple boundary and surface representations are needed to represent a 3D scene during normal vision. Each of these boundary and surface representations can selectively respond to a different range of depths from an observer (Figure 12). The form- and depth-sensitive filling in of all these surface representations, or Filling-In-DOmains (FIDOs), gives rise to the 3D percept. In particular, FACADE theory predicts how binocular boundary signals are topographically projected, from where they form in layer 2/3 of the interstripes of cortical area V2, to the monocular surface FIDOs within the thin stripes of cortical area V2 (Fig. 43). These boundaries act as *filling-in generators* that initiate filling-in of surface brightness and color at positions where the boundary contour and feature contour signals are positionally aligned. After

Downloaded from Brill.com04/06/2021 01:46:22AM
via free access

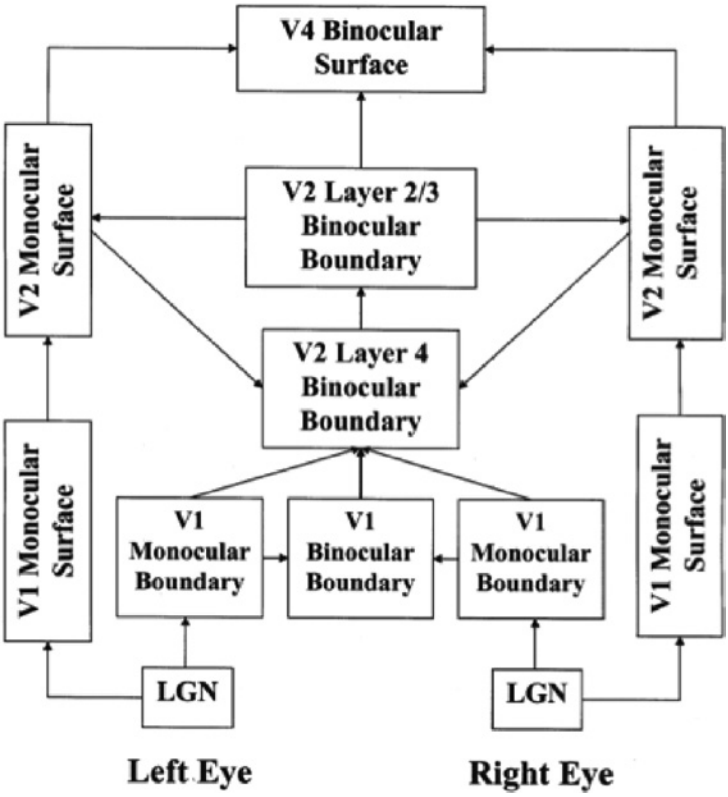


Figure 43. Macrocircuit of the 3D LAMINART model of 3D vision and figure–ground perception. See text for details. (Reprinted with permission from Cao and Grossberg, 2005).

filling-in is initiated, boundaries also act as *filling-in barriers* that prevent the filling-in of brightness and color from crossing object boundaries (Grossberg, 1994), as they do in all the percepts that have already been described.

If a boundary at a given depth is closed (Fig. 11, upper row), then it can contain the filling-in of an object's feature contour signals within it (Fig. 11, lower row, left). If the boundary at a different depth has a sufficiently big gap in it, then surface brightness and color can spread through the gap and surround the boundary on both sides, thereby equalizing the contrasts on both sides of the boundary (Fig. 11, lower row, right). Only a closed boundary can contribute to the final visible 3D percept. This last fact helps to explain how end gaps, such as those that occur during neon color spreading in Fig. 24, and in response to the abutting rectangles in Figs 41 and 42a, can contribute to figure–ground separation.

How do closed boundaries help to form a visible 3D percept? How does this process also help to ensure complementary consistency, and to thereby, as a

Downloaded from Brill.com04/06/2021 01:46:22AM
via free access

surprising consequence, contribute to figure–ground separation? In addition to the *boundary-to-surface* interactions that act as filling-in generators and barriers (Fig. 12), there are also *surface-to-boundary* feedback interactions from filled-in surfaces in the V2 thin stripes to the corresponding depth-selective boundaries in the V2 interstripes (Fig. 44). This feedback is carried out by *surface contour* signals. They are called surface contour signals because they are generated at the bounding contours of surface regions that fill-in brightness or color within a closed boundary (e.g., Fig. 11, lower row, left). As a result, the positions of surface contours and of the closed boundaries that surround the region are the same (Fig. 44).

Surface contours form at these positions because the outputs from the filled-in surface regions are generated by a *contrast-sensitive* on-center off-surround network. The inhibitory connections of this network's off-surround act *across position* and *within depth*, and thus within a FIDO, to generate output signals only at positions where the filled-in contrasts change rapidly across space. These are precisely the positions where boundaries block the further spread of the filling-in process. Surface contour signals are not, however, generated at boundary positions near a big enough boundary gap, such as an end gap, because brightnesses and colors can then spread across the gap and equalize on both sides of the boundary, thereby causing zero contrast, which generates no output from the contrast-sensitive network (e.g., Fig. 11, lower row, right).

Surface contour output signals generate feedback signals to the boundary representations that induced them (Fig. 44). These feedback signals are delivered to the boundary representations by another on-center off-surround network. The inhibitory surface-to-boundary connections of this network act *within position* and *across depth* (Fig. 45): The on-center signals strengthen the boundaries that generated the successfully filled-in surfaces. They can do this because surface contours occur at the same positions as the boundaries that contain surface filling-in. The off-surround signals inhibit redundant boundaries at the same positions but farther depths. This inhibitory process is called *boundary pruning*.

Surface contour signals achieve complementary consistency by strengthening consistent boundaries and pruning redundant boundaries. The inhibited inconsistent boundaries can then contribute to neither seeing nor recognition in the final percept. Only boundaries and surfaces that can contribute to seeing and/or recognition remain. In addition, inhibiting redundant boundaries prevents them from causing recognition of irrelevant contour fragments (see Fang and Grossberg (2009)).

As noted above, surface contour signals are generated by a contrast-sensitive on-center off-surround network. Their inhibitory signals are therefore weaker at high curvature points, such as the corners of the rectangle depicted in

Downloaded from Brill.com04/06/2021 01:46:22AM
via free access

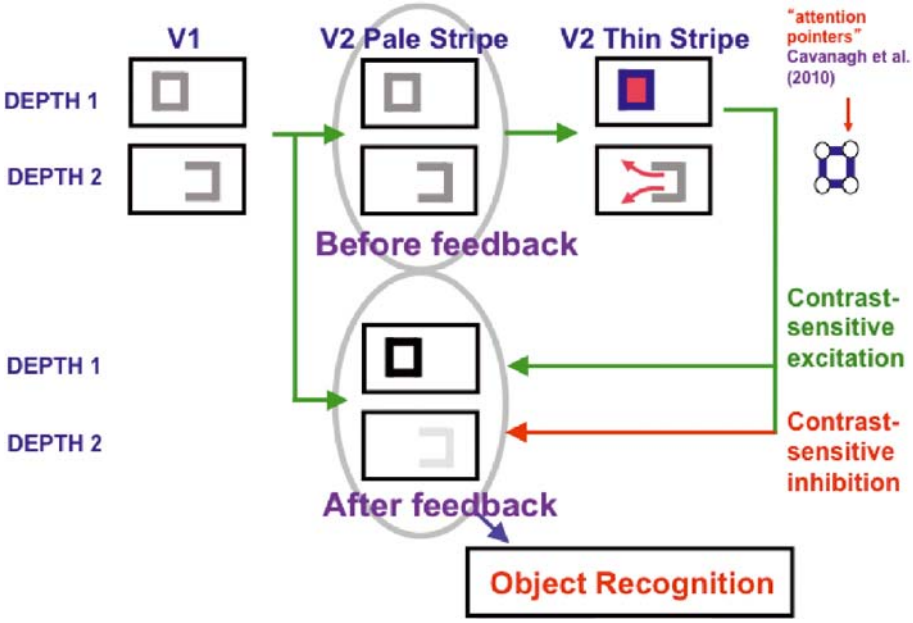


Figure 44. This figure depicts how surface contour surface-to-boundary feedback signals are generated from the model's V2 Thin Stripes to its V2 Pale Stripes. Surface contours that are generated at the bounding contours of filled-in surfaces that are surrounded by closed boundaries. In this figure, they occur at Depth 1 of the V2 Thin Stripes. Surfaces that are surrounded by closed boundaries are filled-in in the manner depicted in Fig. 11 (left). These surface contour feedback signals strengthen the V2 Pale Stripe boundaries that caused them at the same depth (green arrows to Depth 1 boundaries), while inhibiting boundaries at the same positions but further depths (red arrows to Depth 2 boundaries). Because surface contours are generated by contrast-sensitive on-center off-surround networks, they are strongest near high curvature positions, which can be used to direct attention—that is, to serve as 'attention pointers' (Cavanagh *et al*., 2010). Attention pointers play several roles in the model (Fazl *et al.*, 2009): They can control predictive remapping signals that maintain a stable head-centered representation of a shroud as the eyes movement, thereby helping to assure visual stability. They can also direct spatial attention and eye movements to scan the salient features on an object surface, and thereby enable multiple view-specific categorical representations of the object to be learned, and linked by associative learning with an emerging invariant object categorical representation. After incomplete boundaries at the farther Depth 2 are eliminated by using surface contour inhibitory signals from the filled-in surface at Depth 1, other boundaries of partially occluded objects, as in Fig. 41 (left), can be amodally completed behind them, as in Fig. 41 (right). This fact illustrates how surface contours help to initiate figure–ground separation and the completion of partially occluded object representations.

Downloaded from Brill.com04/06/2021 01:46:22AM
via free access

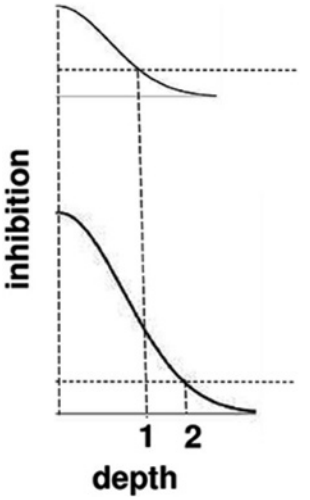


Figure 45. Surface contour inhibitory signals, such as the red arrows in Fig. 44, decrease in strength as the depth difference between the filled-in surface and the boundaries increases. In other words, there is an off-surround within position and across depth that becomes weaker as the depth difference increases. The strength of inhibition of boundaries across depth is depicted with a cross section of the inhibitory off-surround across depth that is caused by surface contour outputs. The top curve shows the inhibitory signals in response to a less bright stimulus. The bottom curve shows the inhibitory signals in response to a more bright stimulus. The numerals 1 and 2 indicate one of the depths where the two sets of inhibitory signals are equal. This illustrates how the brighter stimulus can inhibit boundaries at more depths between that of the stimulus and its inducers, thereby making the brighter stimulus stand out more in depth. This property helps to explain why brighter Kanizsa squares look closer (Bradley and Dumais, 1984; Kanizsa, 1955, 1974; Purghé and Coren, 1992), other things being equal. (Reprinted with permission from Grossberg, 2014).

Figure 17. Such high curvature points are often the positions of salient features of an object. As noted in Figure 17, these salient feature positions are chosen one at a time, in order of signal strength, to compute target positions that control where the eyes will move on an attended surface.

The ARTSCAN model hereby clarifies how several important processes during visual perception and recognition are coordinated: focusing spatial attention on an object of interest, enabling invariant object category learning and recognition of the object, and control of scanning eye movements to salient features, and thus different views, of an object while it is attended, so that invariant object categories can be learned.

4.10. From boundary pruning to figure–ground separation

Eliminating redundant boundaries at farther depths via boundary pruning also allows figure–ground perception to begin. For example, in response to the

Downloaded from Brill.com04/06/2021 01:46:22AM
via free access

image of three rectangles in Figure 41 (left panel), there are four T-junctions where end gaps will form. Only the horizontal rectangle is then surrounded by a closed boundary, so only the bounding contours of the horizontal rectangle will generate surface contour feedback signals. These surface contours use near-to-far inhibition to eliminate redundant copies of the boundaries of the horizontal rectangle at farther depths, using the network in Figure 44. The two pairs of collinear vertical boundaries due to the T stems then have no obstructing horizontal boundaries between them at farther depths. They can therefore use bipole grouping cells to complete a vertical boundary between them (Figure 41, right panel), hereby forming two complete vertical boundaries that are part of a complete vertically oriented rectangle. Because this boundary completion occurs at farther depths, the completed vertical rectangle lies "behind" the horizontal rectangle.

Why are the occluded parts of the vertical rectangle's boundary invisible? This is easy to explain, because *all* boundaries are invisible! The harder part is to explain why only the *unoccluded* surface regions of opaque objects are visible. These visible surface percepts are predicted to occur in cortical area V4, whereas cortical area V2 is proposed to generate amodal recognition of the occluded parts of the scene. A complete explanation of how this distinction between amodal recognition by V2 and modal seeing and recognition by V4 goes beyond the scope of the present article. See Grossberg (1997, 2017) for a heuristic explanation, and Grossberg and Yazdanbakhsh (2005) and Kelly and Grossberg (2000) for computer simulations of figure–ground percepts that support this explanation.

We have hereby seen how basic mechanisms of perceptual grouping and complementary consistency can give rise to properties of figure–ground perception, notably 3D percepts of partially-occluded objects, in response to 2D pictures. These mechanisms also help to explain temporally bistable 3D percepts, such as the 3D Necker cube percepts that oscillate in response to the 2D picture in Fig. 42b (Necker, 1832; for explanation and computer simulations, see Grossberg and Swaminathan, 2004), and the percept wherein attending to one disk is a display of three overlapping disks (Fig. 42c) makes that disk look nearer and brighter (Tse, 2005; for explanation and computer simulations, see Grossberg and Yazdanbakhsh, 2005), among many other percepts. Now let us return to more examples of how particular artists have used these properties to create 3D percepts in response to viewing 2D paintings that include figure–ground manipulations.

4.11. From Boundary Pruning to Proximity-Luminance Covariance

The boundary pruning mechanisms that are summarized in Fig. 44 also help to explain properties of proximity–luminance covariance as part of the

Downloaded from Brill.com04/06/2021 01:46:22AM
via free access

figure–ground separation process. This is because of the way that the surface contour inhibitory signals in Fig. 44 are delivered to the boundary system (Grossberg, 2014, 2017). These inhibitory signals are part of an off-surround network whose strength decreases with the distance from the source cell. This 'distance' translates into a depth difference (Fig. 45). The strength of the inhibitory surface contour signal thus *decreases* as the depth difference *increases* between the surface that generates the signal and its recipient boundaries.

A brighter surface generates a larger surface contour signal. This larger surface contour signal causes more inhibition to occur at every depth that the off-surround can inhibit. By increasing these inhibitory signals, the depth difference increases between the figural surface and the depth of the background surface that can generate boundaries after the inhibition acts. Hence, brighter surfaces look closer than their backgrounds.

4.12. Frank Stella, Occlusions, Kineticism, and Surface–Shroud Resonances

The luminance and color structure within a painting affects its figural grouping and stratification, and this in turn affects the formation of attentional shrouds that determine where viewers look at it. Section 4.7 discussed how larger stripe groupings within Davis's paintings result in the formation of surface–shroud resonances (Fig. 18) on nearer depth planes. This section discusses how different combinations of surface–shroud resonances may form in response to multiple groupings that include partial occlusions in response to Frank Stella's Protractor paintings. The paintings in Stella's Protractor series are completely abstract, brightly colored, and contain a number of interwoven and/or overlapping figures. The series is called the Protractor series because the paintings in this series are based on the semi-circular form of a protractor. Figure 46 shows *Firuzabad* (top) and *Khurasan Gate (Variation) I* (bottom). In describing *Firuzabad* and other similarly-structured paintings, Stella wrote: "I was looking for a really symmetrical base and the protractor image provided that for me. Firuzabad is a good example of looking for stability and trying to create as much instability as possible. 'Cause those things are like bicycle wheels spinning around'" ("*Frank Stella and the art of the protractor*"; Anonymous, 2000).

Stella also describes being inspired by decorative art in Iran in which patterns "doubled back on themselves" (Engberg, 2005). Furthermore, others have described these paintings as having an "engaging kineticism" (Engberg, 2005). The paintings containing interwoven patterns within Stella's Protractor series are visually dynamic and create a sense of movement in the viewer.

FACADE theory together with ARTSCAN model principles can explain the sense of visual movement and rhythm within many of Stella's Protractor series paintings. In *Firuzabad*, the circular form of the outer protractor shapes mirrors the shapes of the interlocking rings within the painting. The T-junctions

Downloaded from Brill.com04/06/2021 01:46:22AM
via free access

Figure 46. Two paintings by Frank Stella: (Upper row) *Firuzabad* (1970). (Lower row) *Khurasan Gate (Variation) I* (1969). See text for details.

where the surfaces of differently-colored segments meet signal which segments weave in front of other segments. FACADE theory predicts that this is the result of the end gaps that form at T-junctions due to the long-range cooperation and short-range competition properties of boundary grouping using bipole cells (Sect. 4.8). End gaps support boundary completion of the nearer figure and create holes in the boundaries of the figure that is 'behind' the nearer figure. This boundary discontinuity prevents filling in of the 'more distant' figure in the nearest FIDO, just as it does in response to the image in Fig. 41. As a result of boundary pruning, surface pruning, and surface-to-boundary feedback (Fig. 44), the boundary of the 'more distant' figure is amodally completed and its surface filled-in in a FIDO that is behind the nearer figure. Only the non-occluded part of each figure is consciously seen, but whole protractor figures may nonetheless be amodally recognized.

The center region of *Firuzabad* contains a number of interwoven segments, some of which are recognized as whole protractors due to amodal completion, and others as smaller segments that are nested in the overall circular shape of the canvas and/or within other protractors. Importantly, no one segment in this center portion of *Firuzabad* is 'on top' of all the others. Instead, each segment weaves under and over other segments.

Similarly, in *Khurasan Gate (Variation) I*, no one segment, or protractor, in the entire painting lies completely on top of all the others, even though each distinctly colored protractor segment has exposed, visible regions that are sufficient to allow the viewer to understand the whole shape of the protractor. As a result, these structures are perceived as overlapping segments in

Downloaded from Brill.com04/06/2021 01:46:22AM
via free access

different depth planes. Importantly, amodal boundary completion and surface filling-in of distinct segment colors help us divide the painting into separate, interlocking protractor forms that are tightly linked in a larger structure. In particular, amodal boundary completion allows perceptual completion of protractor boundaries that lay behind other forms, while consistent segment colors within amodally-completed protractor surfaces allows disconnected segments to emerge as unified protractor forms. These boundary and surface completion processes enable attentional shrouds to spread across a whole protractor form, whether or not the entire form is consciously visible. Once an attentional shroud can form on a given protractor, a surface–shroud resonance regulates eye movements to this particular protractor (see Fig. 17), and guides these movements in and out of different depth planes because no one protractor lies on top of all the others. The ability of shrouds to form on the entire surface of figures, even when portions of the figure are occluded, combined with the interwoven nature of surface depths within the paintings shown in Fig. 46 contribute to the 'engaging kineticism' and simultaneous stability and instability within Stella's Protractor series paintings.

An attentional shroud that adheres to the surface of one figure inevitably collapses as a result of *inhibition of return*, which can be due either to inhibition of recently visited fixation positions, or to activity-dependent habituation of the surface–shroud resonance itself (Chang *et al.,* 2014; Fazl *et al.*, 2009). When a shroud collapses, spatial attention is *disengaged* from that object surface, thereby freeing it to become *engaged* by a new surface of interest (cf. Posner, 1980). By directing attentional flow through different depth planes and across different segmentations, Frank Stella achieves a sense of visual movement in paintings that are based on static interleaved patterns.

4.13. "All Boundaries Are Invisible" and Painting Directly in Color: Monet, Matisse, Hawthorne, and Hensche

Monet directly painted the visible colors in a scene as he viewed it, rather than constructing a scene by drawing its boundaries and then filling them in. Support for this idea lies within Monet's advice: "When you go out to paint, try to forget what objects you have before you, a tree, a house, a field, or whatever. Merely think, here is a little square of blue, here an oblong of pink … paint it just as it looks to you, the exact color and shape, until it gives your own naïve impression of the scene before you" (Perry, 1927, p. 120). This approach to painting has been vigorously pursued by many painters.

For example, Matisse wrote about "the external conflict between drawing and color…Instead of drawing an outline and filling in the color…I am drawing directly in color" (Matisse, 1947/1992). Matisse realized instinctively that, if he painted directly with appropriately shaped color patches, these patches

Downloaded from Brill.com04/06/2021 01:46:22AM
via free access

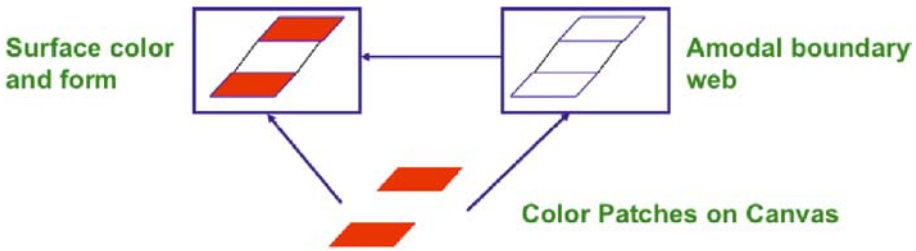


Figure 47. When Matisse or other artists 'paint in color' without using explicitly drawn edges that could darken the painting's perceived colors, these color patches generate amodal boundary webs that capture the colors and trigger surface filling-in to create representations of surface color and form that can be interpreted as parts of a scene by viewers. (Reprinted with permission from Grossberg, 2008).

would induce the formation of amodal boundaries within the brain of the viewer. These boundaries, in turn, would capture the inducing colors to form the surface representations of color and form that a viewer would use to understand the painting (Fig. 47).

How does "drawing directly in color" change how a painting looks? If instead of "drawing directly in color", Matisse did "draw an outline" around his surfaces, and did so in a dark color, then these outlines could darken the surface colors of the entire scene via the process of *color assimilation*. Color assimilation occurs due to mechanisms whereby boundaries control the filling-in of surface color. Grossberg and Todorovic (1988) and Kelly and Grossberg (2000) provide explanations and simulations of situations in which assimilation can occur, as well as of situations where brightness constancy, contrast, contrast constancy and many other visual phenomena can occur. Matisse's method of "drawing directly in color" enabled him to create bright, glowing surface colors, without encountering the darkening and other distorting effects of visibly drawn outlines. Both Matisse and Monet hereby achieved their aesthetic goals by exploiting an intuitive understanding of the fact that 'all boundaries are invisible'.

Many other artists have also struggled with how to represent object surfaces without drawing explicit outlines, lines, or curves around them, including plein air painters of the Cape Cod school of art. Charles Hawthorne (1938/1960), who founded this school of painting, wrote: "Beauty in art is the delicious notes of color one against the other…all we have to do is to get the color notes in their proper relation" (p. 18). "…put down spots of color…the outline and size of each spot of color against every other spot of color it touches, is the only kind of drawing you need bother about…Let color make form—do not make form and color it. Forget about drawing…" (pp. 25, 26).

Downloaded from Brill.com04/06/2021 01:46:22AM
via free access

Henry Hensche, Hawthorne's most famous student and a respected painter and teacher in his own right (Hensche, 1988), further developed these concepts and also noted Monet's key role in pioneering them: "When Monet came along…he revolutionized the 'art of seeing.' …it was the method of expressing light in color, and not value, to allow the key of nature to show clearly… The landscape helped Monet determine how color expressing the light key was the first ingredient in a painting, not drawing" (Robichaux, 1997, p. 27). "The untrained eye is fooled to think he sees forms by the model edges, not with color…Fool the eye into seeing form without edges" (p. 31). "Every form change must be a color change" (p. 33).

These and many other artists hereby developed their theories of painting to accommodate their intuitive understanding of the aesthetic implications of the fact that 'all boundaries are invisible'.

4.14. Back to Monet: Gist, and Global-To-Local Spatial Attention

Monet's painting process reflects how humans may initially perceive global information about a scene, such as its gist, before focusing attention upon its finer details. As discussed in Sect. 2.4, ARTSCENE models how gist may be computed first as a large-scale texture category. Indeed, rapid and accurate classification of natural scenes can be achieved using gist alone. Shifting attention via attentional shrouds to classify finer scenic textures helps to refine scene classification hypotheses.

The principle that scene gist is the first available information in a scene is mirrored by Monet's aesthetic goal to preserve the first glance, or first impression, of a scene. Descriptions of Monet's painting process show that he started with distributed patches of local contrast and slowly added to this contrast until long-range cooperation between his brushstrokes could more easily and unambiguously occur. Lilla Cabot Perry describes a canvas that Monet had painted only once: "it was covered with strokes about an inch apart and a quarter of an inch thick, out to the very edge of the canvas" (Perry, 1927, p. 120). Perry also noted: "[Monet] held that the first real look at the *motif* was likely to be the truest and most unprejudiced one, and said that the first painting should cover as much of the canvas as possible, no matter how roughly, so as to determine at the outset the tonality of the whole" (p. 120). Perry also describes one that was painted twice: "the strokes were nearer together and the subject began to emerge more clearly" (p. 120). Monet hereby instinctively carried out the kind of global-to-local process that humans use to understand scenes, gradually building detail with more and more brushstrokes, but never explicitly drawing fine scenic structures.

For example, in the paintings of the Rouen cathedral shown in Figs 30 and 31, the viewer can infer that there are sculptural elements on the facade of the

Downloaded from Brill.com04/06/2021 01:46:22AM
via free access

Figure 48. Three paintings by Monet in which the existence of trees and their leaves (surrounded by white squares) is contextually disambiguated, but not defined completely by delineating their local features. (Upper left) *Grainstacks in the Morning, Snow Effect* (1891) and detail. (Bottom left) *Mount Kolsaas, Rose Reflection* (1895) and detail. (Right) *Morning on the Seine, near Giverny* (1896) and detail. See text for details.

cathedral, but the fine structure of these elements is not explicitly present in the paintings. Similarly, in the small patches of Monet's paintings within the white squares shown in Fig. 48, one cannot tell from the patches alone that they all contain trees. However, in the context of each complete scene, it is easier to discern that these patches contain leaves on trees within the landscapes. By leaving out fine, scenic details in his paintings, Monet encourages viewers to experience the more global properties of a scene using larger-scale attentional shrouds. This reflects the information acquired during the 'first real look' at a scene and Monet's desire to preserve the freshness of this first impression.

4.15. Graffiti Artists and Mooney Faces

When a painting is not rendered on a smooth surface such as a canvas, additional constraints may influence an artist's technique. This occurs, for example, in the work of graffiti artists, such as Banksy (2005), whose paintings are often

Downloaded from Brill.com04/06/2021 01:46:22AM
via free access

made on walls. As Rubin (2015, p. 1) has noted: "analysis of a large corpus of work by the graffiti artist Banksy suggests that the type and condition of the background wall significantly affected his artistic choices. To minimize on-site production time, Banksy renders his famous subjects (e.g., the rat) by applying single-color paint over pre-fabricated stencils. When the wall is smooth, Banksy leaves the regions previously covered by the stencil unpainted, relying on observers' perception to segregate figural regions from the (identically colored) background. But when the wall is patterned with large-scale luminance edges—e.g., due to bricks—Banksy takes the extra time to fill in unpainted figural regions with another color."

An example of a rat wall painting by Banksy is shown in Fig. 49 (left, upper row). Rubin (2015, p. 2) goes on to discuss these paintings as examples of a "surface completion" process "whereby a whole surface is perceived when only fragments of its bounding contour are present in the image. The best known examples of surface completion are those of flat (2D) shapes bounded by illusory contours, such as Kanizsa's figures…" Rubin (2015) illustrates such surface completion processes in response to "two-tone images obtained by binarizing the luminance levels of pictures of real-world objects or scenes," as illustrated by 'Mooney faces' (Mooney, 1957) which depict heads and faces in black and white where only salient shadows or highlights are shown, as they would appear in strongly lighted photographs (Fig. 49, second panel, upper row). These stimuli have been used to study the development of 'closure' in schoolchildren; namely "the perception of an object or event which is not completely or immediately represented" (Mooney, 1957, p. 219). Rubin (2015) identifies closure with the process of surface completion. Rubin (2015) also discusses how surface completion may work by modifying Mooney faces with a lattice much like the bricks on a wall (Fig. 49, right, upper row) and by discussing how such backgrounds may influence the formation of Kanizsa square percepts (Fig. 49, lower row).

Despite the heuristic appeal of an analysis in terms of 'surface completion', it does not explain how the brain responds to these images. Rather, one needs to invoke how the computationally complementary processes of boundary completion and surface filling-in (Fig. 10) individually work and interact. Consider, for example, the Mooney face image in Fig. 49 (second panel, upper row). To facilitate recognition of this face, an illusory contour forms between the chin of the face at the picture's bottom right and the cheek of the face at the picture's middle right. This illusory contour thus proceeds obliquely upwards and to the right from the chin to the cheek. Once formed, the illusory contour helps us to recognize the face by organizing the surface filling-in process, notably to separate the white of the face from the white of the background, just as it does when we recognize the Dalmatian in Snow (Fig. 3).

Downloaded from Brill.com04/06/2021 01:46:22AM
via free access

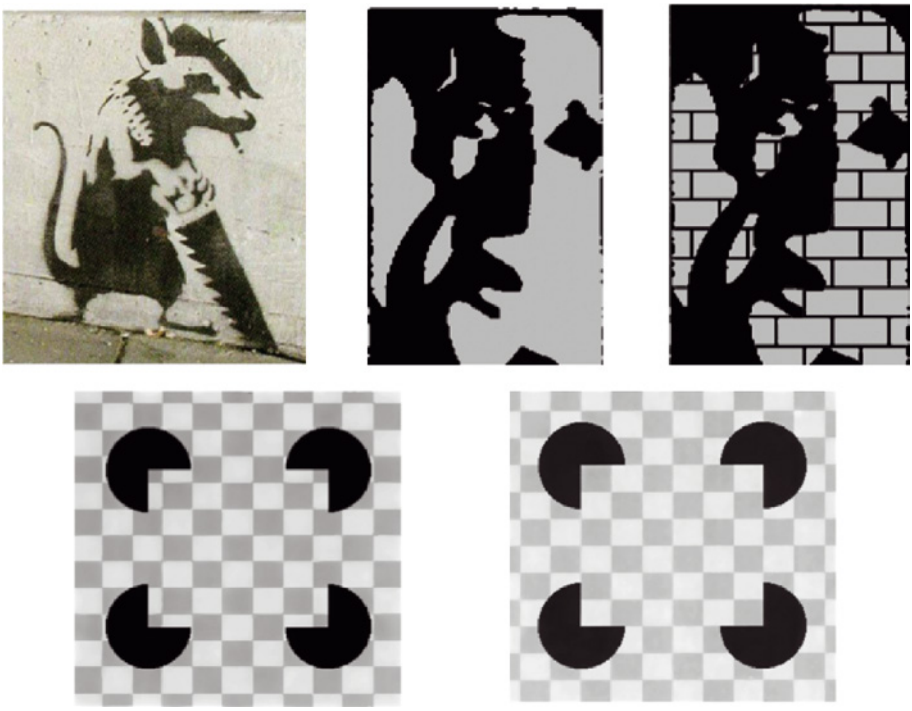


Figure 49. (Left, upper row) Photo of a graffiti painting by Banksy. (Middle, upper row) Face #13 from Mooney (1957). (Right, upper row) Mooney face #13 with brick pattern added (Reprinted with permission from Rubin, 2015). (Lower row) Background squares either support (right) or interfere with (left) the formation of a Kanizsa square. (Adapted from Ramachandran *et al*., 1994).

How does the brickwork pattern in Fig. 49 (upper row, right panel) interfere with this percept? This interference is due to the way in which at least two neural mechanisms react to the bricks:

The first mechanism activates the property of spatial impenetrability that was described in Sect. 3.4. In the present example, horizontally-oriented hypercomplex cells that are activated by the brick horizontal edges inhibit the (almost)-vertically-oriented bipole cells that would otherwise create the illusory contour between the chin and the cheek of the face (Fig. 5b). Because this illusory contour cannot form, it cannot separate the face from its background during the surface filling-in of white in the right half of the percept. In particular, this problem is triggered by the *boundary* system, not the *surface* system, as the analysis of Rubin (2015) might lead one to believe. Surface filling-in ('surface completion') is the result, not the cause, of this property.

Downloaded from Brill.com04/06/2021 01:46:22AM
via free access

The second mechanism causes the amodal completion of the horizontal boundaries of the bricks 'behind' the black shapes in the image, in the same way that the vertical boundaries in response to the image in Fig. 41 are completed behind the horizontal rectangle there. All the white parts of the face are therefore relegated to the background of the percept, and are thereby prevented from acting as part of the facial figure. This process is again triggered by the boundary system, with consequences for the course of subsequent surface filling-in.

A similar analysis can be immediately applied to the two Kanizsa square images in Fig. 49 (lower row). In response to the image shown in the left panel, spatial impenetrability causes the horizontal boundaries of the background squares to interfere with vertical boundary completion between the vertically-oriented collinear pacman edges by vertically-oriented bipole cells. Likewise, the vertical boundaries of the background squares interfere with horizontal boundary completion between the collinear horizontally-oriented pacman edges by vertically-oriented bipole cells. In response to the image shown in the right panel, in contrast, the vertical boundaries of the background squares are collinear with the vertical pacman inducers, and the horizontal boundaries of the background squares are collinear with the horizontal pacman inducers, thereby supporting formation of the Kanizsa square boundaries by bipole grouping. Finer aspects of these percepts, such as why the Kanizsa squares in Fig. 49 (lower row, right) and Fig. 4 (upper row) appear to be in front of four partially occluded circular disks, can be understood by reading how FACADE theory explains many figure–ground percepts of this type as a result of feedback interactions between the complementary boundary and surface cortical streams (Fig. 9) (e.g., Grossberg, 1997, 2014).

5. Concluding Remarks

This article illustrates how different artists have intuitively understood and exploited different combinations of brain processes to achieve their aesthetic goals. The article reviews how neural models have clarified how these processes work, and used this understanding to shed light on the aesthetic effects that are achieved by specific paintings and painterly methods of Jo Baer, Banksy, Ross Bleckner, Gene Davis, Charles Hawthorne, Henry Hensche, Henri Matisse, Claude Monet, Jules Olitski, and Frank Stella. These processes range from discounting the illuminant and lightness anchoring, to boundary and texture grouping and classification, through filling-in of surface brightness and color, to the allocation of spatial attention and eye movement control. The article also clarifies the role of surface–shroud resonances in supporting conscious visual percepts, including percepts of paintings, and

Downloaded from Brill.com04/06/2021 01:46:22AM
via free access

how such resonances control where viewers attend paintings, and how the attention shifts and eye movements that are regulated during such resonances can be used to better understand and appreciate paintings. In this way, the article clarifies how humans consciously see paintings, while also shedding light on how paintings illuminate how humans see.

References

Ahissar, M. and Hochstein, S. (1993). Attentional control of early perceptual learning, *Proc. Natl Acad. Sci. USA* **90**, 5718–5722.

Ahissar, M. and Hochstein, S. (1997). Task difficulty and the specificity of perceptual learning, *Nature* **387** (6631), 401–406.

Anonymous (2000). *Frank Stella and the art of the protractor* [Video file]. Retrieved from https://www.sfmoma.org/watch/frank-stella-and-the-art-of-the-protractor/. Accessed December 26, 2016.

Anonymous (2016). "*Lysander-1*". Retrieved from https://www.guggenheim.org/artwork/3324. Accessed December 26, 2016.

Baer, J. (n.d.). "*Primary Light Group: Red, Green, Blue, 1964-65*". Retrieved from http://www.moma.org/collection/works/79825. Accessed December 26, 2016.

Baer, J. (n.d.). "*Stations of the Spectrum (Primary)*". Retrieved from http://www.tate.org.uk/art/artworks/baer-stations-of-the-spectrum-primary-t03110/text-catalogue-entry. Accessed December 26, 2016.

Baer, J. (1970). Art & vision: Mach bands, *Aspen Mag.* Winter 1970*,* 8.

Banksy (2005). Wall and piece. Century, London, UK.

Beck, J. and Prazdny, S. (1981). Highlights and the perception of glossiness, *Percept. Psychophys.* **30**, 407–410.

Bhatt, R., Carpenter, G. and Grossberg, S. (2007). Texture segregation by visual cortex: Perceptual grouping, attention, and learning. *Vis. Res.* **47**, 3173–3211.

Boersma, L. (1995). Jo Baer, *BOMB Mag.* **53** (Fall). Retrieved from http://bombmagazine.org/article/1888/jo-baer. Accessed December 26, 2016.

Bradley, D. R. and Dumais, S.T. (1984). The effects of illumination level and retinal size on the depth stratification of subjective contour figures, *Perception* **13**, 155–164.

Bressan, P. (2001). Explaining lightness illusions, *Perception* **30**, 1031–1046.

Brockmole, J. R., Castelhano, M. S. and Henderson, J. M. (2006). Contextual cueing in naturalistic scenes: Global and local contexts, *J. Exp. Psychol. Learn. Mem. Cogn.* **32**, 699–706.

Brooks, P. J., Tomasello, M., Dodson, K., and Lewis, L. B. (1999). Young children's overgeneralizations with fixed transitivity verbs, *Child Dev.,* **70**, 1325–1337.

Bullier, J., Hupé, J. M., James, A. and Girard, P. (1996). Functional interactions between areas V1 and V2 in the monkey, *J. Physiol. (Paris)* **90**, 217–220.

Buschman, T. J. and Miller, E. K. (2007). Top-down versus bottom-up control of attention in the prefrontal and posterior parietal cortices, *Science* **315**, 1860–1862.

Cao, Y., and Grossberg, S. (2005). A laminar cortical model of stereopsis and 3D surface perception: Closure and da Vinci stereopsis*, Spat. Vis.* **18**, 515–578.

Downloaded from Brill.com04/06/2021 01:46:22AM
via free access

Cao, Y. and Grossberg, S. (2012). Stereopsis and 3D surface perception by spiking neurons in laminar cortical circuits: A method of converting neural rate models into spiking models, *Neural Netw.* **26**, 75–98.

Cao, Y., Grossberg, S. and Markowitz, J. (2011). How does the brain rapidly learn and reorganize view- and positionally-invariant object representations in inferior temporal cortex? *Neural Netw.* **24**, 1050–1061.

Caputo, G. and Guerra, S. (1998). Attentional selection by distractor suppression, *Vis. Res.* **38**, 669–689.

Carpenter, G. A. and Grossberg, S. (1987). A massively parallel architecture for a self-organizing neural pattern recognition machine, Comput. Vis. Graph. Image Process. **37**, 54–115.

Carpenter, G. A. and Grossberg, S. (Eds.) (1991). *Pattern Recognition by Self-Organizing Neural Networks*. MIT Press, Cambridge, MA, USA.

Carrasco, M., Penpeci-Talgar, C. and Eckstein, M. (2000). Spatial covert attention increases contrast sensitivity across the CSF: Support for signal enhancement, *Vis. Res.* **40**, 1203–1215.

Cavanagh, P. (2005). The artist as neuroscientist, *Nature* **434**, 301–307.

Cavanagh, P., Hunt, A. R., Afraz, A. and Rolfs, M. (2010). Visual stability based on remapping of attention pointers, *Trends Cogn. Sci.* **14**, 147–153.

Chang, H.-C., Grossberg, S. and Cao, Y. (2014) Where's Waldo? How perceptual cognitive, and emotional brain processes cooperate during learning to categorize and find desired objects in a cluttered scene, *Front. Integr. Neurosci.* **8**, 43. doi: 10.3389/fnint.2014.0043.

Chiu, Y. C. and Yantis, S. (2009). A domain-independent source of cognitive control for task sets: shifting spatial attention and switching categorization rules, *J. Neurosci.* **29**, 3930–3938.

Chun, M. M. (2000). Contextual cueing of visual attention, *Trends Cogn. Sci.* **4**, 170–178.

Chun, M. M. and Jiang, Y. (1998). Contextual cueing: Implicit learning and memory of visual context guides spatial attention, *Cogn. Psychol.* **36**, 28–71.

Cohen, M. A. and Grossberg, S. (1984). Neural dynamics of brightness perception: Features, boundaries, diffusion, and resonance, *Percept. Psychophys.* **36**, 428–456.

Conway, B. R. and Rehding, A. (2013). Neuroaesthetics and the trouble with beauty, *PLoS Biol.* **11**, e1001504. doi:10.1371/journal.pbio.1001504.

Desimone, R. (1998). Visual attention mediated by biased competition in extrastriate visual cortex. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **353**, 1245–1255.

DeYoe, E. A. and Van Essen, D. C. (1988). Concurrent processing streams in monkey visual cortex, *Trends Neurosci.* **11**, 219–226.

Dosher, B. A., Sperling, G. and Wurst, S. A. (1986). Tradeoffs between stereopsis and proximity luminance covariance as determinants of perceived 3D structure, *Vis. Res.* **26**, 973–990.

Downing, C. J. (1988). Expectancy and visual-spatial attention: Effects on perceptual quality, *J. Exp. Psychol. Hum. Percept. Perform.* **14**, 188–202.

Egusa, G. (1983). Effects of brightness, hue, and saturation on perceived depth between adjacent regions in the visual field, *Perception* **12**, 167–175.

Engberg, S. (2005). Frank Stella, in: *Bits & Pieces Put Together to Present a Semblance of a Whole: Walker Art Center Collections*, J. Rothfuss and E. Carpenter (Eds), pp. 530–533, Walker Art Center, Minneapolis, MN, USA.

Engel, A. K., Fries, P. and Singer, W. (2001). Dynamic predictions: Oscillations and synchrony in top-down processing, *Nat. Rev. Neurosci.* **2**, 704–716.

Downloaded from Brill.com04/06/2021 01:46:22AM
via free access

Fang, L. and Grossberg, S. (2009). From stereogram to surface: How the brain sees the world in depth, *Spat. Vis.* **22**, 45–82.

Fazl, A., Grossberg, S. and Mingolla, E. (2009). View-invariant object category learning, recognition, and search: How spatial and object attention are coordinated using surface-based attentional shrouds, *Cogn. Psychol.* **58**, 1–48.

Field, D. J., Hayes, A. and Hess, R. F. (1993). Contour integration by the human visual system: Evidence for a local "association field", *Vis. Res.* **33**, 173–193.

Foley, N. C., Grossberg, S. and Mingolla, E. (2012). Neural dynamics of object-based multifocal visual spatial attention and priming: Object cueing, useful-field-of-view, and crowding, *Cogn. Psychol.* **65**, 77–117.

Friedman, A. (1979). Framing pictures: the role of knowledge in automatized encoding and memory for gist, *J. Exp. Psychol. Gen.* **108**, 316–355.

Gao, E. and Suga, N. (1998). Experience-dependent corticofugal adjustment of midbrain frequency map in bat auditory system, *Proc. Natl Acad. Sci. USA* **95**, 12663–12670.

Gilchrist, A. L., Kossyfidis, C., Bonato, F., Agostini, T., Cataliotti, J., Li, X., Spehar, B., Annan, V. and Economou, E. (1999). An anchoring theory of lightness perception, *Psychol. Rev.* **106**, 795–834.

Gove, A., Grossberg, S. and Mingolla, E. (1995). Brightness perception, illusory contours, and corticogeniculate feedback, *Vis. Neurosci.* **12**, 1027–1052.

Gregoriou, G. G., Gotts, S. J., Zhou, H. and Desimone, R. (2009). High-frequency, long-range coupling between prefrontal and visual cortex during attention, *Science* **324**, 1207–1210.

Grossberg, S. (1973). Contour enhancement, short-term memory, and constancies in reverberating neural networks, *Stud. Appl. Math.* **52**, 213–257.

Grossberg, S. (1976). Adaptive pattern classification and universal recoding, II: Feedback, expectation, olfaction, and illusions, *Biol. Cybernet.* **23**, 187–202.

Grossberg, S. (1980). How does a brain build a cognitive code? *Psychol. Rev.* 87, 1–51.

Grossberg, S. (1984). Outline of a theory of brightness, color, and form perception, in: *Trends in mathematical psychology*, E. Degreef and J. van Buggenhaut (Eds), pp. 59–85, North-Holland, Amsterdam, Netherlands.

Grossberg, S. (1987a). Cortical dynamics of three-dimensional form, color, and brightness perception, I: Monocular theory, *Percept. Psychophys.* **41**, 87–116.

Grossberg, S. (1987b). Cortical dynamics of three-dimensional form, color, and brightness perception, II: Binocular theory, *Percept. Psychophys.* **41**, 117–158.

Grossberg, S. (1994). 3-D vision and figure–ground separation by visual cortex, *Percept. Psychophys.* **55,** 48–120.

Grossberg, S. (1997). Cortical dynamics of three-dimensional figure–ground perception of two-dimensional figures, *Psychol. Rev.* **104**, 618–658.

Grossberg, S. (1999). How does the cerebral cortex work? Learning, attention and grouping by the laminar circuits of visual cortex, *Spat. Vis.* **12,** 163–186.

Grossberg, S. (2000). The complementary brain: Unifying brain dynamics and modularity, *Trends Cogn. Sci.* **4,** 233–246.

Grossberg, S. (2007). Towards a unified theory of neocortex: Laminar cortical circuits for vision and cognition, in: *Computational Neuroscience: From Neurons to Theory and Back Again*, P. Cisek, T. Drew and J. Kalaska (Eds), pp. 79–104, Elsevier, Amsterdam, Netherlands.

Grossberg, S. (2008). The art of seeing and painting, *Spat. Vis.* **21**, 463–486.

Downloaded from Brill.com04/06/2021 01:46:22AM
via free access

Grossberg, S. (2009). Cortical and subcortical predictive dynamics and learning during perception, cognition, emotion and action, *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **364**, 1223–1234.

Grossberg, S. (2013). Adaptive Resonance Theory: How a brain learns to consciously attend, learn, and recognize a changing world*, Neural Netw.* **37**, 1–47.

Grossberg, S. (2014). How visual illusions illuminate complementary brain processes: Illusory depth from brightness and apparent motion of illusory contours, *Front. Hum. Neurosci.* **8**, 854. doi: 10.3389/fnhum.2014.00854.

Grossberg, S. (2016). Cortical dynamics of figure–ground separation in response to 2D pictures and 3D scenes: How V2 combines border ownership, stereoscopic cues, and Gestalt grouping rules, *Front. Psychol.* **6**, 2054. doi: 10.3389/fpsyg.2015.02054.

Grossberg, S. (2017). Towards solving the hard problem of consciousness: The varieties of brain resonances and the conscious experiences that they support*, Neural Netw*. **87**, 38–95.

Grossberg, S. (in press). The visual world as illusion: The ones we know and the ones we don't, in: *Oxford Compendium of Visual Illusions,* A. Shapiro and D. Todorovic (Eds), Ch. 7, Oxford University Press, Oxford, United Kingdom.

Grossberg, S. and Hong, S. (2006). A neural model of surface perception: Lightness, anchoring, and filling-in*, Spat. Vis.* **19**, 263–321.

Grossberg, S. and Huang, T.-R. (2009). ARTSCENE: A neural system for natural scene classification, *J. Vis.* **9**, 6. doi:10.1167/ 9.4.6.

Grossberg, S. and McLoughlin, N. (1997). Cortical dynamics of 3-D surface perception: Binocular and half-occluded scenic images*, Neural Netw.* **10**, 1583–1605.

Grossberg, S. and Mingolla, E. (1985a). Neural dynamics of form perception: Boundary completion, illusory figures, and neon color spreading, *Psychol. Rev.* **92**, 173–211.

Grossberg, S. and Mingolla, E. (1985b). Neural dynamics of perceptual grouping: Textures, boundaries, and emergent segmentations, *Percept. Psychophys.* **38**, 141–171.

Grossberg, S. and Mingolla, E. (1987). Neural dynamics of surface perception: Boundary webs, illuminants, and shape-from-shading, *Comput. Vis. Graph. Image Process.* **37**, 116–165.

Grossberg, S. and Raizada, R. (2000). Contrast-sensitive perceptual grouping and object-based attention in the laminar circuits of primary visual cortex*, Vis. Res.* **40,** 1413–1432.

Grossberg, S. and Seidman, D. (2006). Neural dynamics of autistic behaviors: Cognitive, emotional, and timing substrates, *Psychol. Rev.* **113**, 483–525.

Grossberg, S. and Swaminathan, G. (2004). A laminar cortical model for 3D perception of slanted and curved surfaces and of 2D images: Development, attention and bistability*, Vis. Res.* **44**, 1147–1187.

Grossberg, S. and Todorovic, D. (1988). Neural dynamics of 1-D and 2-D brightness perception: A unified model of classical and recent phenomena, *Percept. Psychophys.* **43**, 241–277.

Grossberg, S. and Versace, M. (2008). Spikes, synchrony, and attentive learning by laminar thalamocortical circuits, *Brain Res.* **1218**, 278–312.

Grossberg, G. and Williamson, J.R. (2001). A neural model of how horizontal and interlaminar connections of visual cortex develop into adult circuits that carry out perceptual groupings and learning, *Cereb. Cortex* **11,** 37–58.

Grossberg, S. and Yazdanbakhsh, A. (2005). Laminar cortical dynamics of 3D surface perception: Stratification, transparency, and neon color spreading*, Vis. Res.* **45**, 1725–1743.

Downloaded from Brill.com04/06/2021 01:46:22AM
via free access

Grossberg, S., Mingolla, E. and Ross, W. D. (1997). Visual brain and visual perception: How does the cortex do perceptual grouping? *Trends Neurosci.* **20**, 106–111.

Grossberg, S., Kuhlmann, L. and Mingolla, E. (2007). A neural model of 3D shape-from-| texture: Multiple-scale filtering, boundary grouping, and surface filling-in*, Vis. Res.* **47**, 634–672.

Grossberg, S., Markowitz, J. and Cao, Y. (2011). On the road to invariant recognition: Explaining tradeoff and morph properties of cells in inferotemporal cortex using multiple-scale task-sensitive attentive learning*, Neural Netw.* **24**, 1036–1049.

Grossberg, S., Srinivasan, K. and Yazdanbakhsh, A. (2014). Binocular fusion and invariant category learning due to predictive remapping during scanning of a depthful scene with eye movements, *Front. Psychol.* **5**, 1457. doi: 10.3389/fpsyg.2014.01457.

Harnad, S. (1990). The symbol grounding problem, *Phys. B* **42***,* 335–346.

Hawthorne, C. W. (1938/1960). *Hawthorne on Painting*. Dover, Mineola, NY, USA.

Heitger, F. and von der Heydt, R. (1993). A computational model of neural contour processing: Figure–ground segregation and illusory contours, *Proc. 4th Int. Conf. Comput. Vis.*, Berlin, Germany, pp. 32–40.

Helmholtz, H. von, (1866). *Helmholtz's Treatise on Physiological Optics*. Optical Society of America, New York, NY, USA.

Hensche, H. (1988). *The Art of Seeing and Painting*. Portier Gorman, Thibodaux, LA, USA.

Hong, S. and Grossberg, S. (2004). A neuromorphic model for achromatic and chromatic surface representation of natural images*, Neural Netw.* **17**, 787–808.

Huang, T.-R. and Grossberg, S. (2010). Cortical dynamics of contextually cued attentive visual learning and search: Spatial and object evidence accumulation, *Psychol. Rev.* **117,** 1080–1112.

Hubel, D. H. and Wiesel, T. N. (1968). Receptive fields and functional architecture of monkey striate cortex, *J. Physiol.* **195**, 215–243.

Hupé, J. M., James, A. C., Girard, D. C. and Bullier, J. (1997). Feedback connections from V2 modulate intrinsic connectivity within V1, *Abstr. Soc. Neurosci.* **23**, 1031.

Intraub, H. (1999). Understanding and remembering briefly glimpsed pictures: implications for visual scanning and memory, in: *Fleeting Memories: Cognition of Brief Visual Stimuli,* V. Coltheart (Ed.), pp. 47–70, MIT Press, Cambridge, MA, USA.

Ito, M., Westheimer, G. and Gilbert, C. D. (1998). Attention and perceptual learning modulate contextual influences on visual perception, *Neuron* **20**, 1191–1197.

Jiang, Y. and Chun, M. M. (2001). Selective attention modulates implicit learning, *Q. J. Exp. Psychol.* **54A**, 1105–1124.

Jiang, Y. and Wagner, L. C. (2004). What is learned in spatial contextual cueing: Configuration or individual locations? *Percept.Psychophys.* **66**, 454–463.

Julesz, B. and Schumer, R. A. (1981). Early visual perception, *Annu. Rev. Psychol.* **32**, 575–627.

Kanizsa, G. (1955). Margini quasi-percettivi in campi con stimulazione omogenea, *Rev. Psychol.* **49**, 7–30.

Kanizsa, G. (1974). Contours without gradients or cognitive contours, *It. J. Psychol*. **9**, 93–113.

Kanizsa, G. (1979). *Organization in Vision: Essays on Gestalt Perception*, Praeger, New York, NY, USA.

Kass, J., Harland, B. and Donnelly, N. (2015). Abstracting the set: Monet's cathedrals and stable mental concepts from serial pictorial artworks, *Art Percept.* **3**, 139–150.

Downloaded from Brill.com04/06/2021 01:46:22AM
via free access

Kastner, S. and Ungerleider, L. G. (2001). The neural basis of biased competition in human visual cortex, *Neuropsychologia* **39**, 1263–1276.

Kelly, F. J. and Grossberg, S. (2000). Neural dynamics of 3-D surface perception: Figure–ground separation and lightness perception, *Percept. Psychophys*. **62**, 1596–1619.

Koenderink, J. J., Van Doorn, A., Albertazzi, L. and Wagemans, J. (2015). Relief articulation techniques, *Art Percept.* **3**, 151–171.

Koenderink, J. J., Van Doorn, A., Pinna, B. and Wagemans, J. (2016). Boundaries, transitions and passage, *Art Percept*. **4**, 185–204.

Krupa, D. J., Ghazanfar, A. A. and Nicolelis, M. A. L. (1999). Immediate thalamic sensory plasticity depends on corticothalamic feedback, *Proc. Natl Acad. Sci.* **96**, 8200–8205.

Kulikowski, J. J. (1978). Limit of single vision in stereopsis depends on contour sharpness, *Nature* **275**, 126–127.

Lamme, V. A. F. (2006). Towards a true neural stance on consciousness, *Trends Cogn. Sci.* **10**, 494–501.

Land, E. H. (1964). The retinex, *Am. Sci.* **52**, 247–253, 255–264.

Land, E. H. (1977). The retinex theory of color vision, *Sci. Am*. **237**, 108–128.

Land, E. H. and McCann, J. J. (1971). Lightness and retinex theory, *J. Opt. Soc. Am.* **61**, 1–11.

Llinas, R., Ribary, U., Contreras, D. and Pedroarena, C. (1998). The neuronal basis for consciousness, *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **353**, 1841–1849.

Livingstone, M. (2002). *Vision and Art: The biology of Seeing,* Henry N. Abrams, Inc., New York, NY, USA.

Lleras, A. and von Mühlenen, A. (2004). Spatial context and top-down strategies in visual search, *Spat. Vis.* **17**, 465–482.

Lu, Z.-L. and Dosher, B. A. (2004). Perceptual learning retunes the perceptual template in foveal orientation identification, *J. Vis.* **4**, 5. doi:10.1167/4.1.5

Luck, S. J., Chelazzi, L., Hillyard, S. A. and Desimone, R. (1997). Neural mechanisms of spatial selective attention in areas V1, V2, and V4 of macaque visual cortex, *J. Neurophysiol.* **77**, 24–42.

Mann, S., edited by Mann, P. and Pepperell, R. (2016). Perceptual systems, an inexhaustible reservoir of information and the importance of art, *Art Percept.* **4**, *265–279.*

Matisse, H. (1947/1992). Jazz, George Braziller, Scranton, PA, USA.

Mooney, C. M. (1957). Age in the development of closure ability in children, *Can. J. Psychol.* **11**, 219–226.

Mounts, J. R. W. (2000). Evidence for suppressive mechanisms in attentional selection: Feature singletons produce inhibitory surrounds, *Percept. Psychophys.* **62**, 969–983.

Mumford, D. (1992). On the computational architecture of the neocortex. II. The role of corticocortical loops, *Biol. Cybernet.* **66**, 241–251.

Necker, L. A. (1832). Observations on some remarkable optical phaenomena seen in Switzerland; and on an optical phaenomenon which occurs on viewing a figure of a crystal or geometrical solid, *London Edinburgh Philos. Mag. J. Sci.* **1**, 329–337.

Olitski, J. (1994). Clement Greenberg in my studio, *Am. Art,* **8**, 125–129.

Oliva, A. (2005). Gist of the scene, in: *Neurobiology of Attention*, L. Itti, G. Rees and J. K. Tsotsos (Eds), pp. 251–257, Elsevier Academic Press, Burlington, MA, USA.

Oliva, A. and Torralba, A. (2001). Modeling the shape of the scene: A holistic representation of the spatial envelope, *Int. J. Comput. Vis.* **42**, 145–175.

Olson, I. R. and Chun, M. M. (2002). Perceptual constraints on implicit learning of spatial context, *Vis. Cogn.* **9**, 273–302.

Downloaded from Brill.com04/06/2021 01:46:22AM
via free access

Palma, J., Versace, M. and Grossberg, S. (2012a). After-hyperpolarization currents and acetylcholine control sigmoid transfer functions in a spiking cortical model, *J. Comput. Neurosci.* **32**, 253–280.

Palma, J., Grossberg, S. and Versace, M. (2012b). Persistence and storage of activity patterns in spiking recurrent cortical networks: Modulation of sigmoid signals by after-hyperpolarization currents and acetylcholine, *Front. Comput. Neurosci.* **6**, 42. doi: 10.3389.fncom.2012.00042.

Parker, J. L. and Dostrovsky, J. O. (1999). Cortical involvement in the induction, but not expression, of thalamic plasticity, *J. Neurosci.* **19**, 8623–8629.

Perlovsky, L. I. (2010). Intersections of mathematical, cognitive, and aesthetic theories of mind, *Psychol. Aesthet. Creat. Arts* **4**, 11–17.

Perry, L. C. (1927). Reminiscences of Claude Monet from 1889 to 1909, *Am. Mag. Art* **18**, 119–126.

Perry, E. K., Lee, M. L. W., Martin-Ruiz, C. M., Court, J. A., Volsen, S. G., Merrit, J., Folly, E., Iversen, P. E., Bauman, M. L., Perry, R. H. and Wenk, G. L. (2001). Cholinergic activity in autism: Abnormalities in the cerebral cortex and basal forebrain, *Am. J. Psychiat.* **158**, 1058–1066.

Pinna, B. and Grossberg, S. (2005). The watercolor illusion and neon color spreading: A unified analysis of new cases and neural mechanisms. *J. Opt. Soc. Am. A* **22**, 2207–2221.

Pinna, B. and Reeves, A. (2006). Lighting, backlighting and watercolor illusions and the laws of figurality, *Spat. Vis.* **19**, 341–373.

Posner, M. (1980). Orienting of attention, *Q. J. Exp. Psychol.* **32**, 3–25.

Pollen, D. A. (1999). On the neural correlates of visual perception, *Cereb. Cortex* **9**, 4–19.

Potter, M. C. (1976). Short-term conceptual memory for pictures, *J. Exp. Psychol. Hum. Learn. Mem.* **2**, 509–522.

Potter, M. C. and Levy, E. I. (1969). Recognition memory for a rapid sequence of pictures, *J. Exp. Psychol.* **81**, 10–15.

Purghé, F. and Coren, S. (1992). Amodal completion, depth stratification, and illusory figures: A test of Kanizsa's explanation, *Perception* **21**, 325–335.

Ramachandran, V. S., Ruskin, D., Cobb, S., Rogers-Ramachandran, D. and Tyler, C.W. (1994). On the perception of illusory contours, *Vis. Res*. **34,** 3145–3152.

Rankin, A. (1987). Ross Bleckner. *BOMB Mag.* **19**, Spring.

Rao, R. P. N. and Ballard, D. H. (1999). Predictive coding in the visual cortex: A functional interpretation of some extra-classical receptive field effects, *Nat. Neurosci.* **2**, 79–87.

Reynolds, J. H. and Desimone, R. (2003). Interacting roles of attention and visual salience in V4, *Neuron* **37**, 853–863.

Reynolds, J., Chelazzi, L. and Desimone, R. (1999). Competitive mechanisms subserve attention in macaque areas V2 and V4, *J. Neurosci.* **19**, 1736–1753.

Richards, W. and Kaye, M. G. (1974). Local versus global stereopsis: Two mechanisms, *Vis. Res*. **14**, 1345–1347.

Riggs, T. (1997). Instand Loveland, 1968. Retrieved from http://www.tate.org.uk/art/artworks/olitski-instant-loveland-t07244/text-summary. Accessed December 26, 2016.

Robichaux, J. W. (1997). *Hensche on Painting*, Dover, Mineola, NY, USA.

Roelfsema, P. R., Lamme, V. A. F. and Spekreijse, H. (1998). Object-based attention in the primary visual cortex of the macaque monkey, *Nature* **395**, 376–381.

Rubin, N. (2015). Banksy's graffiti art reveals insights about perceptual surface completion, *Art Percept.* **3**, 1–17.

Downloaded from Brill.com04/06/2021 01:46:22AM
via free access

Schor, C. M. and Tyler, C. W. (1981). Spatio-temporal properties of Panum's fusional area, *Vis. Res.* **21**, 683–692.

Schor, C. M. and Wood, 1. (1983). Disparity range for local stereopsis as a function of luminance spatial frequency, *Vis. Res.* **23**, 1649–1654.

Schor, C. M., Wood, I. and Ogawa, J. (1984). Binocular sensory fusion is limited by spatial resolution, *Vis. Res.* **24**, 661–665.

Sillito, A. M., Jones, H. E., Gerstein, G. L. and West, D. C. (1994). Feature-linked synchronization of thalamic relay cell firing induced by feedback from the visual cortex, *Nature* **369**, 479–482.

Singer, W. (1998). Consciousness and the structure of neuronal representations, *Philos. Trans. R. Soc. B Biol. Sci.* **353**, 1829–1840.

Somers, D. C., Dale, A. M., Seiffert, A. E. and Tootell, R. B. (1999). Functional MRI reveals spatially specific attentional modulation in human primary visual cortex, *Proc. Natl Acad. Sci. USA* **96**, 1663–1668.

Steinman, B. A., Steinman, S. B. and Lehmkuhle, S. (1995). Visual attention mechanisms show a center-surround organization, *Vis. Res.* **35**, 1859–1869.

Thorell, L. G., De Valois, L. G. and Albrecht, D. G. (1984). Spatial mapping of monkey V1 cells with pure color and luminance stimuli, *Vis. Res.* **24**, 751–769.

Tomasello, M. and Herron, C. (1988). Down the garden path: Inducing and correcting overgeneralization errors in the foreign language classroom, *Appl. Psycholinguist.* **9**, 237–246.

Tse, P. U. (2005). Voluntary attention modulates the brightness of overlapping transparent surfaces, *Vis. Res.* **45**, 1095–1098.

Tyler, C. W. (1975). Spatial organization of binocular disparity sensitivity, *Vis. Res.* **15**, 583–590.

Tyler, C. W. (1983). Sensory processing of binocular disparity, in: C. M. Schor and K. J. Cuiffreda (Eds), Vergence Eye Movements, pp. 199–295, Butterworths, Boston, MA, USA.

Tyler, C. W. and Kontsevich, L. L. (1995). Mechanisms of stereoscopic processing: Stereoattention and surface perception in depth reconstruction, *Perception* **24**, 127–153.

Vanderbosch, M. E., Van Doorn, A. J., Koenderink, J. J. and Te Pas, S. F. (2015). Edge-based shading as a depth cue in painting, *Art Percept.* **3**, 173–189.

Vanduffel, W., Tootell, R. B. and Orban, G. A. (2000). Attention-dependent suppression of meta-bolic activity in the early stages of the macaque visual system, *Cereb. Cortex* **10**, 109–126.

Van Tuijl, H. F. J. M. (1975). A new visual illusion: Neonlike color spreading and complementary color induction between subjective contours, *Acta Psychol.* **39**, 441–445.

Varin, D. (1971). Fenomini di contrasto e diffusione chromatica nell organizzazone spaziale del campo percettivo, *Rev. Psychol.* **65**, 101–128.

von der Heydt, R., Peterhans, E. and Baumgartner, G. (1984). Illusory contours and cortical neuron responses, *Science* **224**, 1260–1262.

Wagemans, J., Elder, J. H., Kubovy, M., Palmer, S. E., Peterson, M. A., Singh, M. and von der Heydt, R. (2012a). A century of Gestalt psychology in visual perception I. Perceptual grouping and figure–ground organization, *Psychol. Bull.* **138**, 1172–1217.

Wagemans, J., Feldman, J., Gepshtein, S., Kimchi, R., Pomerantz, J. R., van der Helm, P. A. and van Leeuwen, C. (2012b). A century of Gestalt psychology in visual perception II. Conceptual and theoretical foundations, *Psychol. Bull.* **138**, 1218–1252.

Downloaded from Brill.com04/06/2021 01:46:22AM
via free access

Wallach, H. (1948). Brightness constancy and the nature of achromatic colors, *J. Exp. Psychol.* **38**, 310–324.

Wallach, H. (1976). *On perception*, Quadrangle/The New Your Times Book Co., New York, NY, USA.

Williams, L. R. and Jacobs, D. W. (1997). Stochastic completion fields: A neural model of illusory contour shape and salience, *Neural Comput.* **9**, 837–858.

Zavagno, D. (1999). Some new luminance-gradient effects, *Perception* **28**, 835–838.

Zavagno, D., Annan, V. and Caputo, G. (2004). The problem of being White: Texting the highest luminance rule, *Vision* **16**, 149–159.

Zeki, S. (1999). *Inner Vision: An Exploration of Art and the Brain*. Oxford University Press, Oxford, UK.

Downloaded from Brill.com04/06/2021 01:46:22AM
via free access