Chapter 7

The Visual World as Illusion

The Ones We Know and the Ones We Don't

Stephen Grossberg

EXPECTATION, IMAGINATION, AND ILLUSION

When you open your eyes in the morning, you usually see what you *expect* to see. Often it will be your bedroom, with things where you left them before you went to sleep. What if you opened your eyes and unexpectedly found yourself in a steaming tropical jungle? Why do we have *expectations* about what is about to happen to us? Why do we get *surprised* when something unexpected happens to us? More generally, why are we Intentional Beings who are always projecting our expectations into the future? How does having such expectations help us to fantasize and plan events that have not yet occurred? Without this ability, all creative thought would be impossible, and we could not imagine different possible futures for ourselves or our hopes and fears for them.

What is the difference between having a fantasy and experiencing what is really there? What is the difference between illusion and reality? What goes wrong when we lose control over our fantasies and hallucinate objects and events that are not really there? Given that vivid hallucinations are possible, especially in mental disorders like schizophrenia, how can we ever be sure that an experience is really happening and is not just a particularly vivid hallucination? If there a fundamental difference between reality, fantasy, and illusion, then what is it?

SEEING, RECOGNIZING, AND CONSCIOUSNESS

Once we mention illusion, it is hard not to talk about consciousness. Our consciousness of visual events is particularly vivid to us because we get so much information about the world from our ability to see. But why do we bother to see? One attractive hypothesis is that we see objects in order to recognize them, so that we can then act upon them to test our expectations or fulfill our desires. But do we really need to *see* something in order to *recognize* it? The answer is no. For example, when we look at the image in Figure I.7-1a, we can all recognize that there is a vertical line, or boundary, of some sort that is interpolated between the array of horizontal black lines. But we do not *see* this line, because it does not have a different brightness or color from its background. Nor does it separate the image into regions of different brightness or color. We *recognize* that the line is there, but it is invisible, or not seen. Such a percept is called *amodal* in this chapter. A *modal* percept is one that does carry some visible brightness or color difference. This use of the term *amodal* generalizes the more traditional use by authors such as Albert Michotte, Gabio Metelli, and Gaetano Kanizsa because the mechanistic analysis of how boundaries and surfaces interact, which is summarized later in this chapter, supports a more general terminology in which perceptual boundaries are formed without any corresponding visible surface qualia.

Simple images like this turn our naive views about the function of seeing upside down. For example, one introspectively appealing answer to the question "Why do we see?" is that "We see things in order to recognize them." But we can recognize the vertical boundary of an offset grating without seeing it. The offset grating is a counterexample to the hypothesis that seeing is needed for recognizing. But if we do not always need to see things in order to recognize them, then why do we bother to see at all? Once we start making observations like this, we realize that things are not what they seem and that we need a different way of thinking in order to understand how we consciously see the world.

ILLUSION AND REALITY

The percept that we experience when viewing Figure I.7-1 shows that we can recognize, and be conscious of, events that are invisible. Many people, including distinguished philosophers, have written about visual consciousness as if it includes only consciousness-of-qualia, or of visible sensations. Figure I.7-1 shows that this is a wrong, or at least overly simple, idea.

This percept of an invisible vertical boundary is also of interest because it is a *visual illusion*—it is not "really there" in the image. Why can we see or recognize illusions? How much of what we see is an illusion? Many Eastern thinkers, such as in Buddhism, have spoken about the world-as-illusion. How correct are they? Do they think we see illusions for the right reasons? Moreover, if our consciousness can include even such bizarre experiences as invisible illusions, then why do we not we get hopelessly confused between illusion and reality all the time? Is there some adaptive property of our brains that implies that we can be conscious of

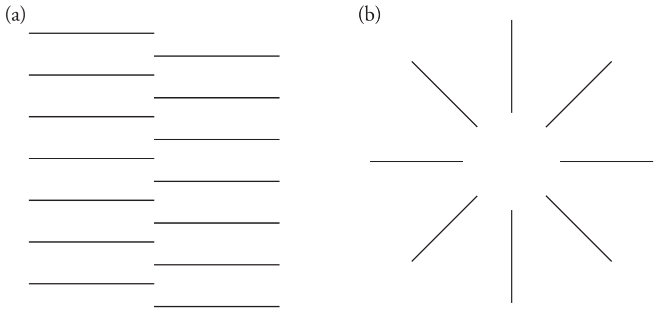


Figure I.7-1. (a) Offset grating. The vertical boundary that emerges between the horizontal black edges can be recognized but not seen. (b) The Ehrenstein illusion. The ends of the black lines induce an illusory circular boundary as well as "brightness buttons" that fill in within the circle to create a percept of enhanced brightness relative to the level of filled-in brightness exterior to the circle.

invisible illusions, without preventing us from experiencing "reality" most of the time? If so, what is it? In fact, in all cases that have been analyzed by our neural theory of how the brain sees, visual illusions arise from adaptive properties of our brains that are needed for survival. More generally, many of the percepts that we believe to be "real" and that we use to survive, are, strictly speaking, visual illusions that are actively constructed from the visual information that we receive from the world through our retinas.

ART AND MOVIES ARE SEEN USING BOUNDARIES AND SURFACES

Before explaining this, let me emphasize that the visual arts and the entertainment industry depend greatly upon visual illusions. Whenever we look at a photograph, painting, movie screen, or computer monitor, we respond to a two-dimensional (2D) picture by constructing a three-dimensional (3D) percept—that is, the picture is experienced as a scene in depth. The human urge to represent the world using 2D pictures dates back at least to Paleolithic times. Artists from ancient to modern times have sought to master how a few lines or color patches on a flat surface can induce mental representations of multiple objects in a scene, notably of occluding objects in front of the objects that they partially occlude. Neural models of seeing have clarified how our ability to experience the world through a picture, or series of pictures in a movie, derives from more basic properties of how we see the world in depth (Grossberg, 1994, 1997).

Figure I.7-2 illustrates the claim that percepts derived from pictures are often illusions. The lower left image in Figure I.7-2 shows three rectangular shapes that abut one another. Our percept of this image irresistibly suggests a different interpretation, however. We perceive a horizontal bar lying in front of a partially occluded vertical bar that is amodally completed behind it. The alternative percept of two rectangles abutting a horizontal bar is much less frequent, although it provides a more literal description of the image.

Even this simple image raises deep questions: Why do the horizontal edges, or boundaries, that are shared by two abutting rectangles seem to "belong" to the horizontal rectangle? When these shared horizontal boundaries are attributed to the horizontal bar, the vertical boundaries of the two rectangles are somehow freed from interference by these horizontal boundaries. The vertical boundaries can then cooperate to induce amodal vertical boundaries "behind" the horizontal bar. These amodal boundaries connect the vertical boundaries of the two rectangles to form a percept of a partially occluded vertical bar. But how can a 2D picture generate a percept in which anything is *behind* anything else? How do two dimensions create a percept in three dimensions? Moreover, how do these perceptual properties relate to classical mathematical concepts of geometry? Whatever its origins, this sort of phenomenon is clearly essential for our ability to recognize the world through pictures.

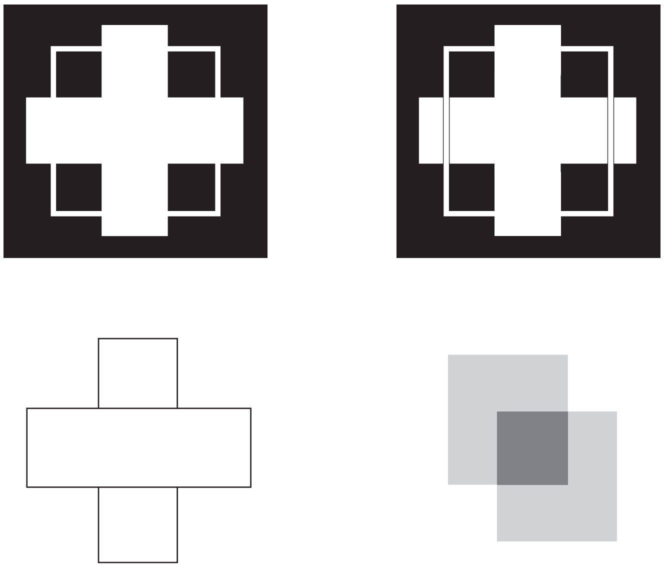


Figure I.7-2. Kanizsa stratification figures are in the top row. The left figure usually looks like a white cross in front of a partially occluded white square, but can switch to look like the square in front of a partially occluded cross. The white region that is shared by the square and the cross is "captured" by the surface form that appears to be in front. In the right figure, complete vertical lines of the square in front of the cross leads to an unlikely percept of a cross that interleaves a square in depth. Bayesian theorists claim that what we see is likelihood statistics. However, this figure causes the highly unlikely percept of a cross both in front of and behind the square. The left image in the second row contains three abutting rectangles but is perceived as a horizontal bar that partially occludes a vertical bar just behind it in depth. The occluded vertical surface is invisible, or amodal, but can nonetheless be recognized. The right image in the second row is also composed of three regions. It generates a bistable percept of a transparent square in front of a background square. (Top row, left column figure is from "Nuove Ricerche Sperimentali Sulla Totalizzazione Percettiva," by G. Petter, 1956, Rivista di Psicologia. 50. Copyright 1956 by Giunti: Gruppo Editoriale. Adapted with permission. Top row, right column figure is from "Seeing and Thinking," by G. Kanizsa, 1985, Acta Psycologia, 59. Copyright 1985 by Elsevier Science. Adapted with permission.)

The two images in the top row of Figure I.7-2 provide another example of how our brains can interpret 2D pictures as events in three dimensions. These images were introduced by the great Italian psychologist Gaetano Kanizsa, who was one of the most insightful contributors to visual perception until his death in 1993. They are called *stratification* images because they generate a percept of an occluding object in front of an occluded object. Thus, although the left image is just a white region on a black background, it generates a percept either of a white cross in front of a white outline square, or of a white outline square in front of a white cross. The former percept usually occurs, but the percept can intermittently switch between these two interpretations. When a percept switches like this between two different interpretations, it is said to be *bistable*.

Stratification percepts raise many questions: How does the boundary of the cross get completed across the white regions where it intersects the square? How does the white color within the completed cross boundary get "captured" into a cross-shaped surface that is enclosed by this boundary? How is this cross assigned to a depth closer than that of the square? How does the rest of the white region get assigned to the square at this more distant depth? How do the incomplete square boundaries get amodally completed "behind" the cross to facilitate recognition of the square? Why do not the completed square boundaries capture the white color within them where they intersect the cross? When the percept switches, so that the square appears in front of the cross, why do the completed square boundaries succeed in capturing the white color that previously was attributed to the cross, whereas the completed cross boundaries are merely amodally completed behind the square? Why is the percept bistable? And why does the cross win over the square more often than not?

One might at first think that such percepts arise because of prior learning about familiar objects in the world, like crosses and squares, as the large psychophysical literature on perceptual learning illustrates. Although visual experiences do influence how the brain develops and learns, something different is going on when we perceive images such as those in Figure I.7-2. For example, in the upper right image of Figure I.7-2, two pairs of vertical black lines create boundaries that help to capture the intervening white color into a vertical surface of the square. This vertical region seems to lie in front of the horizontal bars of the cross. This effect influences the percept only locally, however. When we look to the middle of the picture, we see that the vertical bars of the cross are still completed in front of the square. The vertical bars of the cross appear to lie in front of the cross, even though the horizontal bars of the cross appear to lie in back of the cross. Kanizsa used this image to argue that knowledge about the world could not determine such percepts because, surely, our many experiences with crosses and squares would, with high probability, place the cross behind the square *everywhere* once it is forced to be behind the vertical bars of the square. This is not, however, what we perceive. We perceive, instead, the highly unlikely situation in which the cross appears to be deformed to lie both in front and in back of the square. The famous graphic artist M. C. Escher used such unlikely properties to create fascinating images that look sensible locally but are globally impossible.

These 2D images in Figure I.7-2 all generate 3D percepts, and the upper left image shows that such percepts can be bistable, with different 3D perceptual interpretations of the same picture alternating through time. The thickness of the limbs of the cross relative to those of the square are responsible for the cross boundaries appearing in front of the square boundaries most of the time. In response to the lower right image in Figure I.7-2, again, two 3D percepts can alternate through time. In this case, the nearer square looks transparent because one can see the farther square behind it. This percept is due to the relative luminances of the various parts of the image, since changing these luminances can create either a 2D opaque percept or a 3D transparent percept that is not bistable. Thus, either geometrical properties of an image, as in the upper left image, or contrastive properties of the image, as in the lower right image, can cause 3D bistable percepts to be seen.

All of these percepts, and many more like them, have been explained and simulated using the form-and-color-and-depth (FACADE) theory of how the visual cortex accomplishes 3D vision and figure–ground perception, and its refinement as the 3D LAMINART theory of how the laminar circuits of visual cortex carry out these perceptual functions (Cao & Grossberg, 2005, 2012; Fang & Grossberg, 2009; Grossberg, 1987, 1994, 1997; Grossberg & Howe, 2003; Grossberg & McLoughlin, 1997; Grossberg & Pessoa, 1998; Grossberg & Swaminathan, 2004; Grossberg & Yazdanbakhsh, 2005; Grossberg, Yazadanbakhsh, Cao, & Swaminathan, 2008; Kelly & Grossberg, 2000). In particular, Grossberg (1997) predicted how interactions between figural geometry and contrast can change what we see, Kelly and Grossberg simulated various forms of these percepts, and Grossberg and Yazdanbakhsh explained and simulated how both opaque and transparent percepts can occur in response to different combinations of contrasts in the lower right figure.

These figures illustrate a controversy about how we see and recognize the world that has continued to the present time. The alternative view—that we see what we expect to see because of prior learning—also has a lot of evidence to support it. Hermann von Helmholtz, one of the greatest scientists of the 19th century, proposed that we see using *unconscious inferences*, or learned expectations. Helmholtz was a precursor of the use of Bayesian statistics by some vision scientists today. In contrast, Kanizsa provided hundreds of brilliant counterexamples to Helmholtz's hypothesis using images, such as the ones in Figure I.7-3 that violate expectations. Kanizsa emphasized the power of *bottom-up* visual processes, such as filtering, grouping, and filling-in that act directly on visual scenes and images at early stages of brain processing, whereas Helmholtz called attention to the impact of *top-down* processes, such as learned expectations, that are activated by processes deeper in the brain and are matched against these more peripheral processes.

Helmholtz and Kaniza emphasized important, but complementary, types of visual processes. Visual illusions can be created by both bottom-up and top-down processes. A more comprehensive understanding has arisen from neural models that show how bottom-up and top-down processes work together to generate an attentive consensus, or resonance, between bottom-up signals that are derived from the world and what we expect to see. How such a consensus arises also helps to explain how we can rapidly learn about the world throughout life and become conscious of events in it without being forced to unselectively forget already learned knowledge as we try to learn

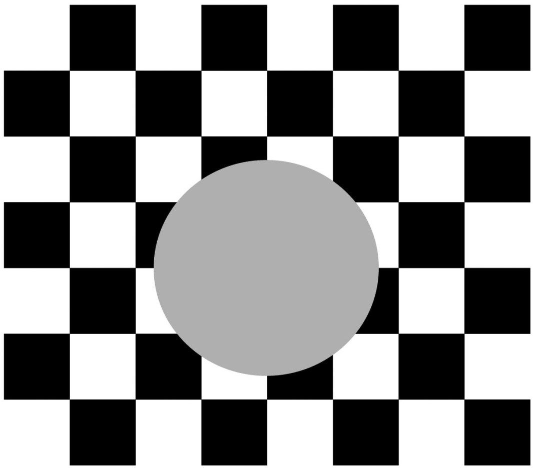


Figure I.7-3. Two gray disks are superimposed on a black-and-white checkerboard grating. The gray disk to the upper left seems to partially occlude a white cross, whereas the disk to the lower right seems to occlude a black cross. Due to the regularity of the checkerboard and our prior experiences with checkerboards, a Bayesian account would have predicted that a black square is occluded at the upper left and a white square at the lower right, but this is not what is perceived.

new things (Cao, Grossberg, & Markowitz, 2011; Carpenter & Grossberg, 1991; Fazl, Grossberg, & Mingolla, 2009; Grossberg, 1980, 2007; Grossberg, Markowitz, & Cao, 2011; Raizada & Grossberg, 2003). Helmholtz and Kanizsa could not fully make these connections because they did not have some of the critical intuitive concepts and none of the mathematics needed to express them clearly.

WHY DO WE BOTHER TO SEE?

Earlier we introduced the distinction between seeing and recognizing, which is classically sometimes described as the distinction between *seeing* and *knowing*. Figure I.7-1a shows that we can consciously recognize the offset grating even though we cannot see it. The image in Figure I.7-1b consists of a series of concentric black lines that face the same central point on the paper. Our percept of this figure is one in which we both see *and* recognize a circle that abuts the ends of the black lines. We *see* this circle because it separates the brighter white disk within the circle from the darker background that lies outside it. This circle is a visual illusion, since neither the circle nor the bright disk within it is part of the picture itself. It is called the Ehrenstein illusion in honor of the German perceptual scientist Walter Ehrenstein, who studied it.

We next propose answers to the following questions:

- Why are there so many visual illusions?
- How does illusion differ from reality?
- Why do illusory percepts not cause us a lot of trouble in successfully adapting to the "real" world?

Illusions provide vital clues about how the brain successfully adapts to the real world, and many percepts and other experiences that we believe to be real are, in actuality, mentally constructed illusions that just happen to match the structure of the world. How this happens clarifies why the debate about the difference between illusion and reality has gone on for so long, why many cultures have emphasized mystical states of mind that exploit the constructive properties of the mind, and why the long-held proposal of artificial intelligence that brains may be modeled by the architecture of a Von Neumann computer is wrong. The next section suggests that our percept of the offset grating in Figure I.7-1a reflects a general fact about how we perceive the "real" world.

ALL BOUNDARIES ARE INVISIBLE

The percept of a vertical boundary in the offset grating shows us that "some boundaries are invisible." In fact, "*all* boundaries are invisible," in the sense that the boundary system does not directly generate a percept of visible qualia such as brightnesses and colors (Cohen & Grossberg, 1984; Grossberg, 1984; Grossberg & Mingolla, 1985a, 1985b). The following discussion explains what perceptual boundaries are, how the brain computes them, and how boundaries reflect a general principle of brain design.

First, consider Figure I.7-3. This famous image was introduced by Kanizsa to argue against the position that all seeing is based on knowledge-based hypothesis testing. In particular, the two gray disks are perceived to partially occlude a white cross and a black cross, rather than the checkerboard squares that would be expected from experience. Figure I.7-3 can also be used to understand a different issue. Let your eye wander along the circular boundary of one of the two gray disks. At successive moments, the boundary may abut a white square, then a black one, then a white one, and so on. Where the square is white, the *contrast* between the background and the disk goes from white to gray, that is, from a lighter to a darker shade of gray. Where the square is black, the contrast goes from black to gray, or from a darker to a lighter shade of gray. As your eye traverses the boundary, the contrast keeps flipping between light to dark and dark to light. Despite these contrast reversals, we perceive a single continuous boundary surrounding the gray disk. The boundary is not broken into two sets of four pieces, depending on whether the direction of contrast goes from light to dark or from dark to light. This would have been the result if all that happened in perceiving the circular boundary was for cells that are sensitive to a light-to-dark contrast, or to a dark-to-light contrast, but not both, responded along the circle. Instead, this percept illustrates that the brain pools, or adds, signals from cells that are sensitive to both of these contrast polarities, at every position, in order to detect the boundaries of objects, irrespective of their direction of contrast. The same thing may be said about the checkerboard background of the image: We can recognize long vertical boundaries from the top to the bottom of the image, even though these boundaries are formed from alternating contrast polarities. We can also recognize the long horizontal boundaries of the checkerboard in the same way.

The punch line is: if the brain adds together signals from opposite contrast polarities in order to build boundaries around objects, then these boundaries cannot tell the difference between light and dark, and thus cannot carry a visible signal of lightness or darkness. In other words, *all boundaries are invisible* because the brain needs to build boundaries around objects that lie in front of textured backgrounds whose relative contrast with respect to the object can reverse as the boundary is traversed.

Apart from having serious problems with object recognition if boundaries did not pool opposite contrasts, there would also be serious perceptual consequences. For example, if a connected boundary did not form around objects, then its lightnesses and colors could spread, through a surface filling-in process, between objects in a scene and we could not properly see object surfaces.

This contrast-pooling property occurs at an early stage in visual cortical processing, namely, at the complex cells in layer 2/3 of cortical area V1. Perhaps the simplest example of this boundary-pooling property can be perceived by viewing Figure I.7-4. The image to the left of Figure I.7-4 is perhaps the most famous image due to Kanizsa. This image consists of four black pac-man, or pie-shaped, figures on white paper. The percept of the square that abuts the pac-men interiors is a visual illusion that is called the Kanizsa square. The enhanced brightness of the square is also an illusion.

Now consider the image to the right of Figure I.7-4. The percept that it elicits is called a "reverse-contrast Kanizsa square." Again a square boundary can be *recognized*. If the contrasts of the gray background relative to the white and black pac-men are chosen correctly, then we will *not see* a brighter or darker square within the square boundary. In this percept, the two vertical boundaries of the square are formed between pac-man edges of opposite contrast polarity: white to gray (light to dark) and black to gray (dark to light). To do this, the brain pools signals from opposite directions of contrast to form an object boundary, albeit the boundary of an illusory object. In addition, the two horizontal boundaries of the square are formed between like directions of contrast, and both sets of boundaries join together to form the square. Thus both like-contrast and opposite-contrast polarities are pooled together to form an object's boundary, thereby illustrating once again why all boundaries are invisible. In other words, boundaries are *insensitive to contrast polarity*, in the sense that they pool signals from cells that are sensitive to opposite contrast polarities.

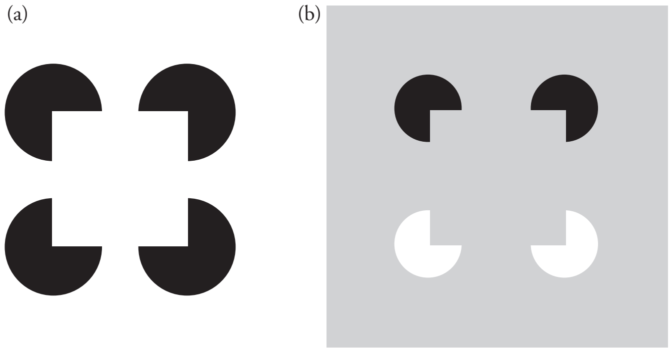


Figure I.7-4. A Kanizsa square (a) and reverse-contrast Kanizsa square (b) illustrate complementary properties of perceptual boundaries and surfaces. Whereas the Kanizsa square may appear brighter than its background, the reverse-contrast Kanizsa square may be recognized, but not seen, if the interior and exterior filled-in gray brightnesses are the same due to the balance between black-to-gray and white-to-gray pac man inducers of filling-in.

If, in fact, all boundaries are invisible, then how do we see the Kanizsa square? How, indeed, do we see anything? Something else must also be going on that lets us see the visible world of surface brightnesses and colors. What is this additional surface perception process, and how is it related to the process whereby the brain forms boundaries? We next build up an answer to these basic questions in several simple stages.

LIVING WITH YOUR BLIND SPOT

To get started, let us first acknowledge that perceptual boundaries and surfaces are not just high-level constructs that have to do with sophisticated processes like object recognition or hypothesis testing. Boundaries and surfaces play a role in processing signals from the very earliest stage of visual processing, the eye itself.

The top of Figure I.7-5 shows a cross-sectional view through an eye. Light streams in from the left, passes through the lens and vitreous humor, and finally hits the photosensitive retina. A ray of light is shown hitting the most sensitive regions of the retina, which is called the fovea. The fovea is the region of the retina with the highest acuity and the highest concentration of color-sensitive photoreceptors. The fovea processes a region of visual space that is approximately 1° wide. Retinal acuity gradually decreases with distance from the fovea. Our eyes move in our heads to direct our sensitive foveas to new points of interest.

Although the retina is a miracle of biological design, it is also a very *noisy* detector. Without the help of additional processing by the brain, retinal noise could easily defeat our ability to see. The bottom of Figure I.7-5 illustrates one source of noise. This figure looks down on the surface of the retina. To the left of the figure is the fovea. To the right, is the *blind spot*. The blind spot is a region of the retina that is totally blind; it has no light-sensitive photodetectors. The blind spot exists because of the way in which the photodetectors in other parts of the retina send light-activated signals to the brain along nerve pathways, called axons. These axons are brought together in the optic nerve, as shown at the top of Figure I.7-5. The blind spot lies right in front of the place where all the axons come together to be bundled into the optic nerve.

The blind spot is large: notice how big it is relative to the fovea. Despite this fact, many of us manage to go through life without ever realizing that we have a large blind hole in our retinas. We certainly do not see a roving hole in the world as our eyes scan a scene. Why not? In addition to the blind spot, Figure I.7-5 shows that a lot of veins lie between the source of light and the retina. These veins nourish the eye, but we do not see these veins either.

There is a second sense in which retinal processing is noisy. When light hits the retinal photodetectors, it causes an electrical response in them. This response is processed by several layers of retinal cells before the processed signals are collected in the optic nerve and sent to the brain. However, these retinal layers are upside-down: light passes through all of them before it ever hits the photodetectors and activates them. All these layers of cells can absorb and scatter the light before it ever reaches the photodetectors.

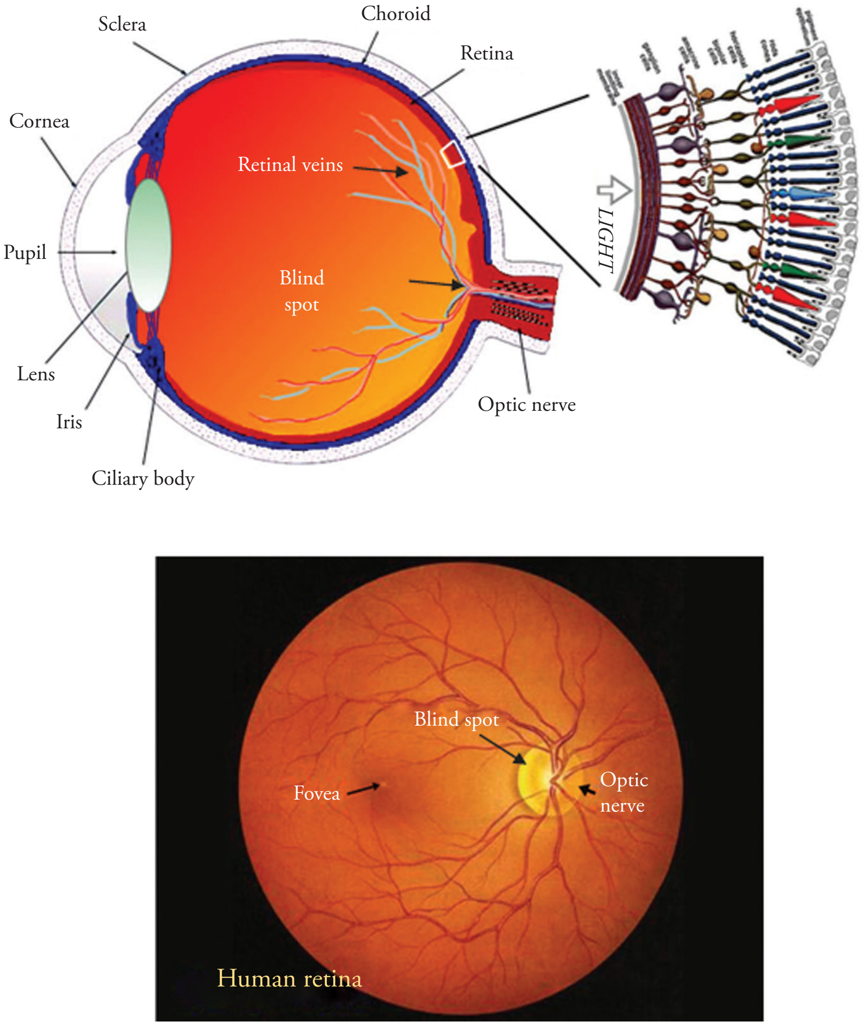


Figure I.7-5. Side and top views of the eye, showing the relationships between the fovea, blind spot, retinal veins, and optic nerve. (Images adapted with permission from Webvision-University of Utah.)

In the next few sections, we suggest how seemingly strange ideas such as "all boundaries are invisible" help to explain how the brain compensates for the noisy design of the retina. From a more general viewpoint, this discussion illustrates how deficiencies of image processing at one stage of brain organization—in this case, due to physical and developmental constraints on how the retina and optic nerve grow and how the retina is nourished by blood—are compensated by processing at higher stages of brain organization. We call this general principle *hierarchical resolution of uncertainty*. This principle helps us to understand how the noisy and rapidly metabolizing meat that makes up our eyes and brain cells give rise to the Platonic world of visual forms of which we are consciously aware as we go through our lives. It also clarifies a sense in which most of the percepts of which we are consciously aware are, mechanistically speaking, visual illusions since they reconstruct, within the brain, image properties that are occluded by the blind spot and retinal veins.

JIGGLING EYES AND OCCLUDED OBJECTS

How does the brain avoid seeing the retinal veins and the blind spot? One contributing factor is that, even when we think our eyes are fixated on a stationary object in the world, they are rapidly jiggling in their orbits, back and forth, by such a small amount that we are not consciously aware of these tiny movements. This jiggle acts like a "refresh" operation that maintains the sensitivity of visual nerve cells to the stationary object. It does so by creating small relative motions between the object and the retina. Many visual cells are sensitive to these motion transients and fire in response to them. Images that do not move relative to the jiggling retina are said to be *stabilized*. Percepts of stabilized images fade away. They cannot be seen after a little while because they are not refreshed. The retinal veins and blind spot are stabilized because they are attached to the retina. They therefore are not visible, at least not under ordinary viewing conditions.

The retinal jiggle helps to discount the veins and blind spot, but more is needed to see the world of continuous objects that we take for granted every day. To get a feel for what is missing, imagine that a simple image, like one of a thick, straight edge, is registered by the retina in a region that includes the blind spot and some retinal veins. Parts of the edge cannot effectively activate retinal photoreceptors because they are received where there are veins or the blind spot. Because the veins and blind spot are stabilized images on the retina, they fade. What remains is not a representation of the entire edge, however, but discontinuous fragments of the edge's unoccluded regions. How does the brain transform these discontinuous fragments into a percept of a continuous edge?

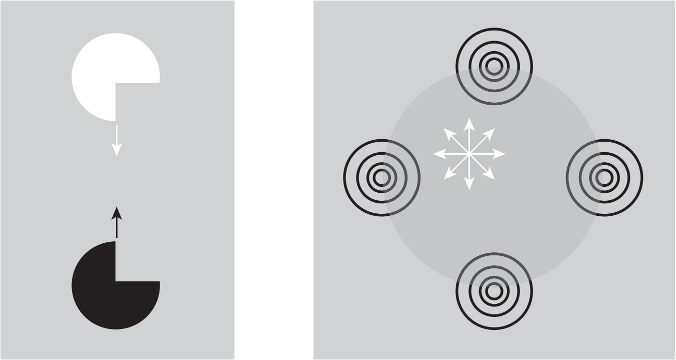
HOW ILLUSIONS MAY REPRESENT REALITY

We can now begin to understand how the brain can use the discontinuous fragments of an edge to complete a continuous edge percept. The occluded fragments of the edge are surrounded by unoccluded fragments that are approximately collinear, or aligned. The brain uses these collinear fragments to complete edge boundaries wherever they are missing. In particular, the brain has cells that can form boundaries *inwardly* between pairs, or greater numbers, of cells that are (nearly) collinear across space and whose orientational preferences are (nearly) collinear as well. In other words, a boundary can form in a region that receives no visual inputs of its own if it has collinear inducers on both sides of the region. Then, all the boundaries that were occluded by the blind spot or retinal veins can be completed. On the other hand, boundaries will not grow out of the ends of a single edge, because boundary formation takes place only inwardly between pairs or greater numbers of inducers. If these boundary-forming cells can also pool collinear signals from opposite contrast polarities, as in the case of the reverse-contrast Kanizsa square of Figure I.7-5b, then they can complete the edge even if the unoccluded edge fragments have opposite contrasts with respect to a textured background.

Using these ideas, we can begin to see how Kanizsa square percepts form in response to the images in Figure I.7-4. In particular, the brain does not known if there is a retinal vein or blind spot between the pac-men in Figures I.7-4a and I.7-4b. It will therefore complete boundaries between pairs of such collinear contrasts whenever they are long enough and not too far apart. Indeed, the existence of the Kanizsa square illusion is a small price to pay for having percepts of the world that are not occluded by the blind spot and retinal veins.

These examples illustrate why boundaries form *inwardly* between pairs, or greater numbers, of inducing contrasts. It is fortunate that boundaries cannot form *outwardly* from a single inducer. Imagine if they could: then any speck of contrast in an image or scene could force our brains to build an expanding web of boundaries that could fill our field of view. Such boundaries would do more to drive us mad than to accurately represent objects in the world. It is also clear from these examples that boundaries form in an *oriented* way between collinear, or nearly collinear, inducers, as in the case of the Kanizsa squares in Figure I.7-4a. Finally, we know from the example of the reverse-contrast Kanizsa square in Figure I.7-4b, and the Kanizsa disk-on-texture display in Figure I.7-3, that boundaries are *insensitive* to contrast polarity. These three properties of boundary completion are summarized in Figure I.7-6.

In contrast to the *inward* completion of boundaries, the filling-in process that completes percepts of visible surfaces propagates *outward* from individual sources of brightness or color. For example, the brightness and color of the unoccluded parts of an edge—after the stabilized blind spot and veins fade—flow outward into the occluded parts of the edge, thereby restoring brightness and color to the edge percept at positions that were occluded by the blind spot or retinal veins. This flow of brightness and color behaves much like a fluid that diffuses away from a source. It is called *filling-in*. Boundaries act like a dam that contains the flow of brightness and color and keeps it from flowing outside the contours of the edge. By acting like barriers, or obstructions, to the flow of brightness and color, the positions of boundaries can become visible when a different brightness or color occurs on opposite sides of the boundary. In the case of the Kanizsa square of Figure I.7-4a, for example, the four black pac-men induce local regions of enhanced brightness, called "brightness buttons," just inside the illusory square. These brightness buttons then fill-in within the square boundary to create a percept of enhanced brightness throughout the square.



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 I.7-6. Complementary properties of boundary completion and surface filling-in. (Reprinted with permission from Grossberg, 2014.)

EVERY EDGE IS AN ILLUSION

To briefly summarize: a combination of inward boundary completion and outward filling-in of surface brightness and color help to complete the representation of an edge that is occluded by retinal veins or the blind spot. A remarkable implication of this discussion is that *every edge is an illusion*.

In other words, our conscious percepts of those parts of the edge that are occluded by retinal veins or the blind spot are actively constructed by the brain, just as are the more obviously illusory percepts like the Kanizsa square. The unoccluded parts of the edge are more "real" than the occluded parts, because they generate their own visual signals at the retina. But we do not know which parts are which. Both types of parts look just as real to us. Moreover, as our eyes move along the edge, these real and illusory regions may be interchanged, with no one the wiser, as different parts of the edge are occluded. In summary, even the percept of such a simple image as a real edge will often include a series of illusory percepts that just happen to look real.

If even real percepts can be illusory, then what criteria do we use to decide in our everyday lives that a percept like the Kanizsa square is an illusion? The concepts of boundary completion and surface filling-in help to do this, especially if we keep in mind the fact that all boundaries are invisible. One possible pragmatic definition of a percept that we may tend to *call* an illusion is: *a visual illusion is a percept whose combination of boundaries and surfaces looks unfamiliar*.

For example, in the percept of the offset grating of Figure I.7-1a, there is a vertical boundary that has no visible brightness or color. We call this unfamiliar combination of boundary and surface color an illusion. On the other hand, the part of an edge that is perceived at the locations over the blind spot is believed to be a "real" percept, even though it is not due to direct visual input from the retina, because it combines boundary and surface properties in a familiar way. Indeed, many percepts that are completed over the blind spot and retinal veins seem to be real even though they are illusions in the narrow mechanistic sense of being constructed by the brain over retinal positions that transmit no contrastive inputs to the brain. The following section attempts to distinguish between the reconstructive processes that we perceive as illusions versus those that are reconstructed but interpreted to represent what is really in the world.

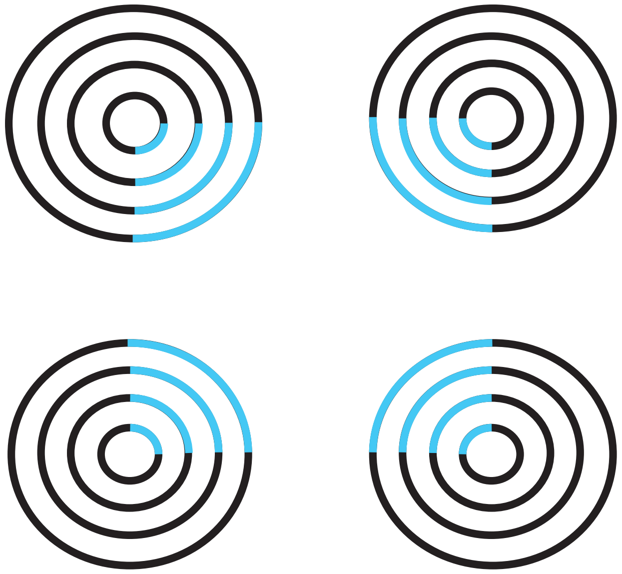


Figure I.7-7. An example of neon color spreading. Although the only blue in the figure is within the blue strips, end gaps in the boundaries that are induced by the figure where the black and blue strips intersect allow blue color to flow out of the blue strips and fill in the illusory square that is created when the end cuts at the black-to-blue edge interfaces cooperate by bipole grouping. (Reprinted with permission from Grossberg, 2008.)

EVERY LINE END IS AN ILLUSION

We noted previously that boundary completion is oriented. To initiate oriented boundary processing, the brain uses oriented detectors, called simple cells, in cortical area V1. However, such orientation-selective tuning of cells can prevent them from responding at the ends of lines. Without further processing, the open gaps in the boundaries at line ends would allow color to flow out of every line end. This additional processing can be accomplished in two stages: Spatial competition among cells at different positions that share a similar orientational preference, followed by a push-pull orientational competition among cells with different orientational preferences at the same position, with maximal competition occurring between cells sensitive to perpendicular orientations, can create line ends capable of containing the line's color within its borders. These constructed line ends are called *end cuts*.

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

NEON COLOR SPREADING

Hierarchical resolution of uncertainty can prevent color from flowing out of some line ends, but it does not prevent color from flowing out of all line ends. A famous example of color spreading from line ends is shown in Figure I.7-7. It is called *neon color spreading*. Neon color spreading illustrates key properties of color filling-in, in addition to how end cuts are formed. The image in Figure I.7-7 consists of circular annuli, part of which are black and part of which are blue. When we view this figure, we can see an illusory square filled with blue color, even though the only blue in the image is in the concentric blue circular strips.

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 I.7-7. Van Tuijl (1975) independently introduced images that gave rise to percepts that he called "neon-like color spreading." The following explains more fully how such illusory percepts arise from an interaction between boundary and surface system properties that are useful for survival (Grossberg & Mingolla, 1985a). For now, let us note that the black arcs and the blue arcs in Figure I.7-7 both create boundaries in our brains. At the positions where these boundaries join, the boundaries caused by the black arcs cause small breaks, called *end gaps*, to occur in the boundaries caused by the blue arcs. This happens because, in order for the percept to occur, the contrast of the black arcs with respect to the white background is chosen to be 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 spatial competition. When these boundaries are inhibited, the cells that are tuned to other orientations, notably the perpendicular orientation, are disinhibited as part of the end-cut process. These disinhibited cells can cooperate across space to complete a boundary with the shape of an illusory square. This long-range cooperative process is often called *bipole grouping* (Grossberg, 1984). The predicted properties of bipole grouping have been supported by psychophysical experiments (Field, Hayes, & Hess, 1993; Kellman & Shipley, 1991) and neurophysiological experiments (von der Heydt, Peterhans, & Baumgartner, 1984; Peterhans & von der Heydt, 1989) that discovered cells with these properties in cortical area V2, among other cortical areas.

Blue color can flow out of the end gaps in the broken boundaries. This filling-in of blue color across space continues until the color hits the square "illusory" boundary, which prevents its further spread. Thus boundaries act as barriers to the filling-in of brightness or color. In summary, the process of end cutting that completes line ends 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.

Neon color spreading illustrates how color can spread *outwardly* from the individual blue strips in all directions. Its spread is thus *unoriented* (e.g., Grossberg & Todorovic, 1988). The blue color continues to spread outwardly in all directions until it fills the illusory square boundary, beyond which it cannot spread any further. In addition to the two filling-in properties of unoriented and outward spread, we can also add the obvious property that filling-in is *sensitive* to contrast polarity, because we can consciously see its effects. These three properties of surface filling-in are summarized in Figure I.7-6.

COMPLEMENTARY PROPERTIES OF BOUNDARIES AND SURFACES

Figure I.7-6 illustrates the prediction, first made in Grossberg (1984), that the brain is organized into parallel processing streams that exhibit computationally *complementary* properties (Grossberg, 2000). Visual boundary completion and surface filling-in provide just one of many examples of brain processes with computationally complementary properties. Figure I.7-6 summarizes that boundaries are completed *inwardly* between pairs or greater numbers of inducers, in an *oriented* fashion, and are *insensitive* to contrast polarity, in the sense that they pool over opposite contrast polarities. In contrast, surface filling-in proceeds *outwardly* in an *unoriented* fashion, and is *sensitive* to contrast polarity, since surfaces can create visible percepts. These properties are manifestly complementary: inward versus outward, oriented versus unoriented, insensitive versus sensitive. They fit together like lock and key, or yin and yang.

Complementarity of boundary and surface processing is important for the success of each process. For example, filling-in needs to be unoriented so that it can cover an entire surface. On the other hand, the unoriented flow of brightness can be efficiently contained only by an oriented boundary. Likewise, a seeing process cannot efficiently build boundaries around objects in front of textured backgrounds. Both types of process are needed for either process to work well. Moreover, both types of process need to interact to overcome each other's complementary deficiencies.

The ability of boundaries to serve as barriers to filling-in provides one example of how interactions between the boundary and surface systems help them to achieve a better result than either one could attain by itself. One of the challenges facing the visual system is to combine these complementary properties into a higher synthesis, which we experience as unified visual percepts. In other words, the challenge is to reconcile the *complementarity* of boundary and surface properties with the *consistency* of conscious percepts. We call this resolution the property of *complementary consistency*.

COMPLEMENTARY PROCESSING STREAMS IN VISUAL CORTEX: BOUNDARIES AND SURFACES

How are these complementary processes represented in the brain? Much evidence suggests that they are carried out by parallel processing stream in the visual cortex. Figure I.7-8 illustrates how visual signals activate the light-sensitive retinas within our eyes. The retinas, in turn, send signals to the lateral geniculate nucleus (LGN). The LGN has a beautiful laminar organization that may, at first approximation, be broken up into a *parvocellular* part and a *magnocellular* part, so named after their parvocellular and magnocellular cell types. Output signals from the LGN branch out and activate several parallel subsystems of the visual cortex.

Three streams are easy to track in Figure I.7-8. Two of these streams proceed from the parvocellular LGN to regions of the first cortical stage, called area V1 in monkeys and area 17 in cats. One of these streams goes through

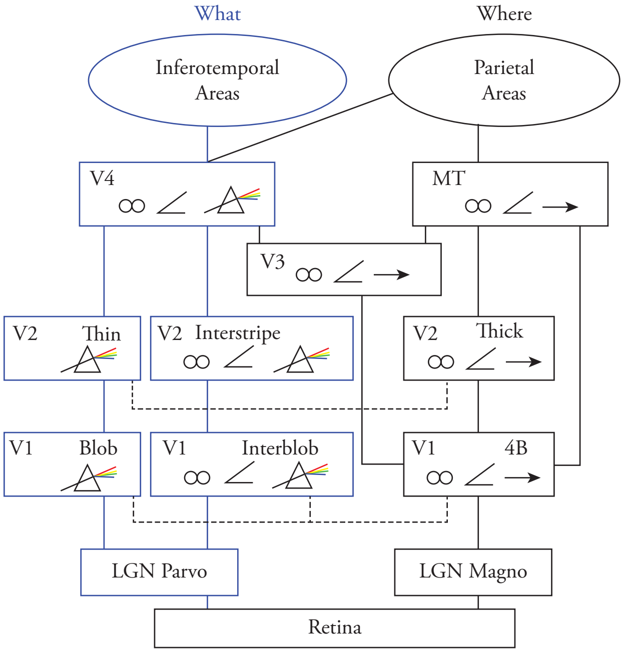


Figure I.7-8. Schematic diagram of anatomical connections and neuronal selectivities of early visual areas in the macaque monkey. LGN = lateral geniculate nucleus; V1 = striate visual cortex; V2, V3, V4, MT = prestriate cortical areas. The boundary stream goes through the blobs and thin stripes to cortical area V4 and inferotemporal areas. The surface stream goes through interblobs and interstripes to V4. The motion stream goes through V1 and MT to the parietal areas. (Reprinted with permission from DeYoe & Van Essen, 1988.)

structures that are called the *blobs* of V1. The blobs are small regions that are distributed in regular arrays throughout V1. They are highly active metabolically and therefore light up when probed by a chemical marker called cytochrome oxydase. The blobs project, in turn, to a region that is called the *thin stripes* of the prestriate cortex, in an area that is called V2 in monkeys and area 18 in cats. The thin stripes then project to prestriate area V4. Grossberg (1984) predicted that the LGN→blob→thin stripe→V4 processing stream generates visual *surface* representations and that the parallel LGN→interblob→interstripe→V4 processing stream generates visual *boundary* representations. Let us call these streams the blob and interblob streams, respectively.

Other investigators, notably Livingstone and Hubel (1984), have made related but distinct proposals. They suggested that the blob stream computes "color" and the interblob stream computes "orientation." These two proposals lead to different sets of predictions. In particular, a boundary system can complete boundaries—both "real" and "illusory"—over positions in space that receive no inputs, let alone oriented inputs; and a surface system can generate filled-in representations of figure–ground relationships that do not directly represent the local brightnesses and colors of a scene, as illustrated by the example of Kanizsa stratification in Figure I.7-2. In particular, FACADE theory predicts why completed boundaries and filled-in surfaces of the occluded parts of objects, including those in Figure I.7-2, are amodal, or invisible, even though they can have profound effects on object recognition (Grossberg, 1994, 1997). These completed representations of occluded regions receive no bottom-up oriented signals and, although they are predicted to control filling-in using double-opponent cells, generate no visible color signals. Many vision scientists now routinely use the boundary/surface distinction to interpret their experiments.

How do the complementary properties of these parallel streams arise? Grossberg (2000) predicted that they arise during brain development as part of a process of *symmetry-breaking*. This concept envisages that pairs of streams may arise from an earlier shared set of cells that bifurcate into complementary streams, while maintaining and elaborating their reciprocal connections, as developmental specialization occurs. Symmetry-breaking is a widespread phenomenon in many parts of physics and biology. In the case of the brain, current neural models suggest that, during brain development, there is a *cascade of symmetry-breaking operations on multiple spatial scales*, with some complementary processes arising as substreams of larger streams that also exhibit complementary relationships.

WATERCOLOR ILLUSION AND NEON COLOR SPREADING

Before going on to discuss other types of complementary processing during visual perception, let us consider some other examples of how complementary processing creates visual illusions of form and color. Here we focus on conceptual themes and examples rather than on mechanistic explanations of visual illusions. All of these illusions have, however, been given quantitative mechanistic neural explanations, supported by computer simulations. For readers who prefer more heuristic discussions. Grossberg & Pinna (2012) provides a heuristic summary of enough modeling concepts and mechanisms to explain a wide range of illusory percepts devised by Pinna (2009).

One of the most famous visual illusions that was studied by Pinna (1987; also see Pinna, Brelstaff, & Spillmann, 2001) is the *watercolor illusion*. The watercolor illusion is a long-range spread of color diffusing from a thin colored contour running parallel and contiguous to a darker chromatic contour. The illusion can also impart a strong figural effect across large regions. In Figure I.7-9, light blue undulating contours flanked by darker blue contours are perceived as irregular curved shapes evenly colored by a veil of light blue spreading from the inner light blue contours.

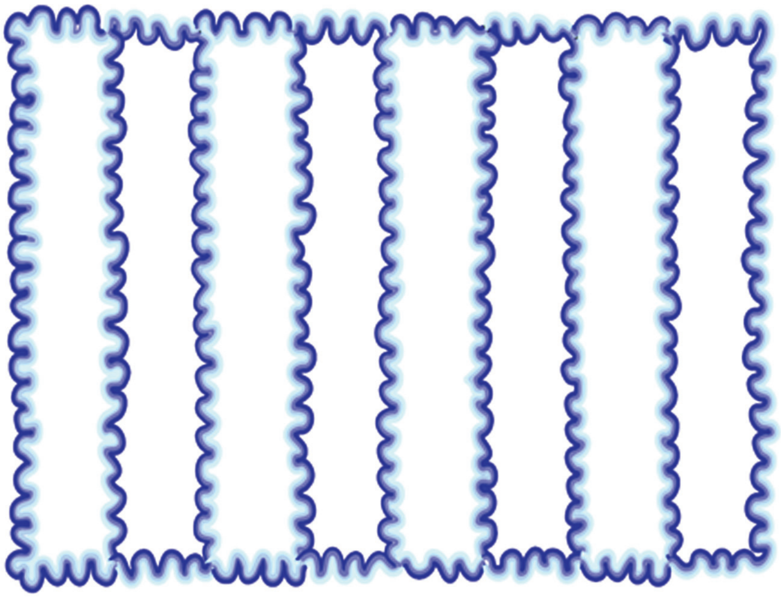


Figure I.7-9. An example of the watercolor illusion. When dark blue and light blue contours are juxtaposed, the stronger boundaries that are induced by the dark blue contours inhibit the weaker boundaries that are induced by the light blue contours, more than conversely. As a result, light blue color flows past the weaker boundaries and fills in the intervening spaces. How this creates a percept of blue figures on a white background is explained in Pinna and Grossberg (2005). (Reprinted with permission from Pinna and Grossberg, 2005.)

The regions into which the light blue spreads look like figures, whereas the remaining white regions look like background. Pinna and Grossberg (2005) explain this and other examples of the watercolor illusion. One relevant neural mechanism is the spatial competition that occurs between nearby boundaries, which also helps to explain how line ends are completed and how neon color spreads out of end gaps in boundaries (see previous discussions in this chapter). In a watercolor illusion image, just as in a neon color spreading image, one boundary is stronger than the other. The stronger boundary in this case is induced by the dark blue contours. The stronger boundary can inhibit the nearby weaker boundary of the light blue contours, more than conversely, via spatial competition. The color that the weaker boundary would otherwise contain can then partially spread, or fill in, onto the contiguous surface.

As noted earlier, the spatial competition that triggers the watercolor illusion is part of a process of hierarchical resolution of uncertainty whereby the boundary completion system can overcome uncertainties that are created by the orientationally tuned processing of edge, texture, and shading signal, thereby completing object boundaries that can usually, but not always, contain the filling-in of brightnesses and colors within the correct object regions.

TEXTURE GROUPING, BREAKING CAMOUFLAGE, AND VISUAL ART

We can now better see how visual illusions help to understand how the visual brain works. Because of their unexpected combinations of boundary and surface properties, they provide important additional information about perceptual processes that we might otherwise overlook. Second, they are crucial for our survival, as is illustrated by the fact that the percept of even of an object's edge may be, from a mechanistic perspective, "illusory." There are hundreds of examples of visual illusions that are important for our survival, including examples that depend on invisible boundaries.

Figure I.7-10 provides one such example, a famous picture of a Dalmatian dog in a snowy terrain. The white coat of the Dalmatian blends into the snowy terrain, so that the Dalmation first looks like a disconnected series of black blotches on white paper. We gradually recognize the Dalmatian when emergent boundaries form that connect the black blotches. These boundaries provide important information about the Dalmatian's shape. Although these boundaries are illusory and invisible, the brain cells that compute the boundaries can trigger signals that activate the object recognition centers in our brains. The brain uses these invisible boundaries to *recognize* the Dalmatian even if the boundaries are not themselves *seen*. The Dalmatian can hereby be recognized using its invisible boundaries in just the same way that the invisible vertical boundary of the offset grating in Figure I.7-1a can be recognized. The boundaries will only be seen if they fill-in the surface representation of the Dalmatian with a different brightness, color, or depth than its background, as in the case of the Kanizsa square (Fig. I.7-4).

Much neurophysiological evidence suggests that object recognition occurs in the *inferotemporal cortex* (e.g., Lueschow, Miller, & Desimone, 1994; Zoccolan, Kouh, Poggioa, & DiCarlo, 2007), whereas perceptual boundaries and surfaces are formed at earlier processing stages that are found in the interblob and blob processing streams of the *prestriate visual cortex*. We can recognize boundaries that we cannot see in part because recognition and seeing are carried out by different parts of the brain.

The Dalmatian example is important because it illustrates a problem that our ancestors may have confronted in the wild. Imagine that the image is not of a Dalmatian but rather of a deadly predator, like a snow leopard, that is stealthily tracking a man in the snow. All that he can see are the predator's black spots, because its white coat blends into the snow. By using his emergent invisible boundaries, he can hopefully recognize the predator in time to initiate evasive or other survival techniques. Emergent boundaries, whether invisible or not, help to

Figure I.7-10. Dalmation in snow. When this image is first seen, it may look like a textured pattern of meaningless black and white textures. After sustained inspection, amodal boundaries are completed and enable recognition of a Dalmation in snow.

break through camouflage in the many situations where local properties of an object—such as its curves and surface elements—would be insufficient for its recognition. In fact, one way to disable our ability to break through camouflage is to generate images whose orientations create emergent boundary groupings that compete with those of the target.

Figures I.7-11 through I.7-13 provide some other examples of how invisible boundary groupings can segregate textured regions from one another, or separate figures from their backgrounds. Figure I.7-11 is a famous image due to the distinguished American psychologist Jacob Beck. It is a texture composed of the letters L and T. Although the regions with vertical and diagonal T's are contiguous, the vertical T's group together with the vertical L's, rather than with the diagonal T's. One reason for this is that the vertical T's and L's share both vertical and horizontal lines, whereas the diagonal T's do not. Invisible vertical and horizontal emergent boundaries can connect, and thereby group together, the vertical T's and L's but not the diagonal T's with the vertical T's. The boundary grouping process is thus highly sensitive to the relative orientation of image contrasts. As in the Kanizsa square, inducer orientations that are *collinear*—that is, have the same orientation and are aligned across space—are strongly favored in the grouping process. Because the properties of grouping take precedence over the different identities of the T's and L's, this percept is consistent with the hypothesis that collinear grouping occurs before object recognition and cannot be overridden by top-down cognitive factors when the emergent groupings are unambiguous.

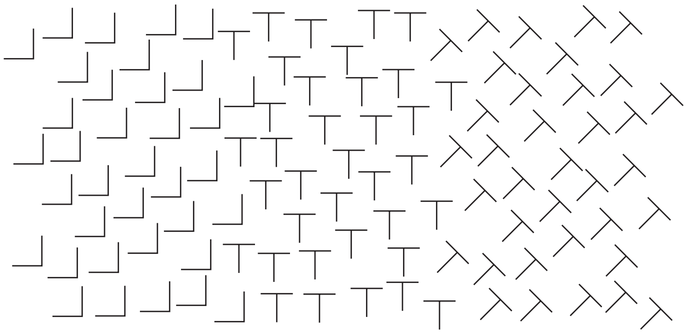


Figure I.7-11. A texture composed of T's and L's. The vertical T's in the middle group with the vertical L's on the left, rather than the oblique T's on the right, thereby illustrating that, when boundary grouping can occur unambiguously, it cannot be overridden by higher-level recognition processes. (Reprinted with permission from Beck, Prazdny, & Rosenfeld, 1983.)

Figure I.7-12 illustrates that emergent boundary groupings are sensitive to both the distance and the relative

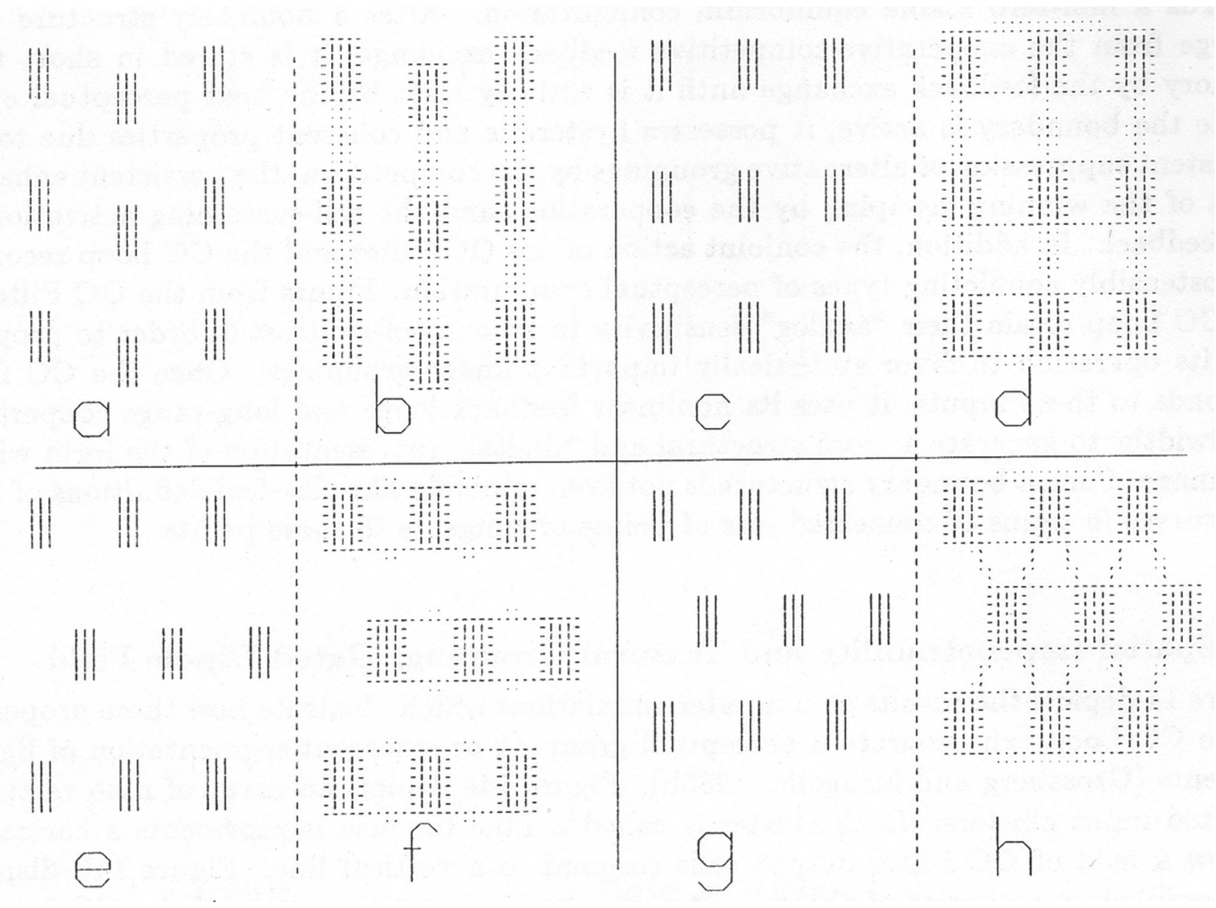


Figure I.7-12. Examples of how emergent perceptual groupings can form with orientations that are not in the inducing pictorial elements. Figures I-7-12a, I.7-12c, I.7-12e, and I.7-12g are inputs. Figures I.7-I.7-12b, I.7-12d, I.7-12f, and I.7-12h are model simulations of the emergent groupings. (Reprinted with permission from Grossberg & Mingolla, 1987.)

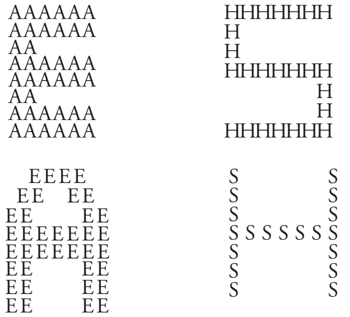


Figure I.7-13. Emergent boundary groupings can generate recognizable letters from local image inducers that do not have the orientations of the letters.

orientations between the inducing elements. It is easier for boundaries to form between elements that are closer together, other things being equal. The importance of such "Gestalt rules" for visual perception was advocated by German psychologists such as Wolfgang Köhler and Max Wertheimer during the first half of the 20th century. Although each of the four images in Figure I.7-12 is composed of three rows of six vertical line segments, their spatial arrangement can generate strikingly different percepts. The percept in Figure I.7-12a is of emergent vertical boundaries, the one in Figure I.7-12c is one of both emergent vertical and horizontal boundaries, the one in Figure I.7-12e is of emergent horizontal boundaries, and the one in Figure I.7-12g can generate both emergent horizontal and diagonal boundaries. Percepts of horizontal and diagonal boundaries emerge in response to Figures I.7-12c, I.7-12e, and I.7-12g images even though there are only vertical line segments in the inducing images. The global arrangement of the line elements determines which grouping will win. Figures I.7-12b, I.7-12d, I.7-12f, and I.7-12g show model simulations of these emergent boundary groupings. These boundary properties use the same combination of spatial competition, orientational competition, and bipole grouping that were used to explain neon color spreading.

One implication of this result can immediately be illustrated in Figure I.7-13, where the boundaries of the letters E, S, A, and H are generated as emergent properties of groupings of the letters A, H, E, and S, respectively. This sort of global "forest before trees" structuring of a percept was particularly emphasized by David Navon. In the vertical parts of the global S, collinear grouping of local orientations is used, whereas in the horizontal parts of the global S, groupings perpendicular to the local orientations form the emergent boundary. The top of the global letter A, in contrast, uses oblique groupings, even though only local horizontal and vertical orientations occur in its constituent letters E.

Many percepts that we observe in real life are illusions of this type. We are continually separating figures from their backgrounds based on boundary groupings that form between regions with different textures and colors. Many visual artists, notably the Impressionists and Fauvists, exploited this property of visual perception to create some of the most beautiful paintings ever made. They were masters at showing how to arrange small patches of color on a canvas in such a way that the groupings that were thereby induced in the brain can be recognized as objects in the world.

In particular, many artists instinctively exploited the fact that invisible boundaries indirectly assure their own visibility through their interactions with the surface stream. Within the surface stream, boundaries trigger the filling-in of lightness and color within their contours and restrict the spreading of lightness and color to the surface region that they bound. In addition, the same visual stimuli activate both the boundary and the surface streams. Thus a visual scene or painting activates both the boundary and the surface streams in parallel, the boundary stream generates an emergent perceptual grouping, and this grouping controls the filling-in of potentially visible surface lightnesses and colors within the surface stream. These facts about brain processing have influenced the creative struggles and decisions of many artists.

Every artist faces the choice of whether to draw edges around the surfaces that his or her painting will represent, or to allow the brain to create boundaries through the process of amodal perceptual grouping. Drawing in an edge with a particular color may influence the appearance of nearby colors either through color assimilation, which may darken the nearby colors, or by color contrast. Matisse (1947/1992), among others, struggled with this issue for many years. Later in his life, when his health led him to work with paper cut-outs, he 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 was already "drawing directly in color" in his paintings from the Fauve period, as illustrated in his painting from 1905 called *The Roofs of Collioure* (Fig. I.7-14). Matisse realized instinctively that, if he painted directly with appropriately shaped color patches, these patches 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 enable the viewer to understand the painting, but without the risk that edges with colors of their own could darken the perceived surface colors (Fig. I.7-15). These concepts clarify the sense in which a masterpiece like *The Roofs of Collioure* is a visual illusion. Grossberg (2008) analyzes various artistic theories and struggles of famous artists from the perspective of neural models of boundary and surface formation.

FROM 2D TO 3D: EFFECTS OF CONTEXT ON 3D BOUNDARY ORIENTATION AND SURFACE FILLING-IN

The previous summary includes discussions of how figure–ground illusions can arise in which partially occluded objects can be completed behind their occluders (e.g., Figs. I.7-2 and I.7-4). However, this barely grazes the huge field of 3D vision and figure–ground perception, which goes beyond the scope of this chapter, although a great deal of neural modeling work in the FACADE and 3D LAMINART theories has clarified how it works and explained many perceptual and neurobiological experiments along the way (e.g., Cao & Grossberg, 2005, 2012; Fang & Grossberg, 2009; Grossberg, 1987, 1995, 1997; Grossberg & Howe, 2003; Grossberg & McLoughlin, 1997; Grossberg

Figure I.7-14. *The Roofs of Collioure* was painted by Matisse in 1905.

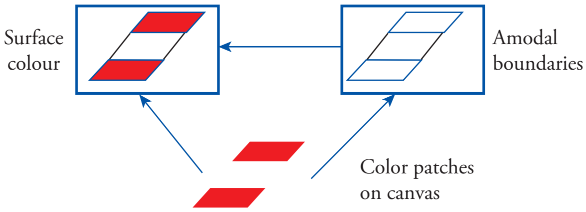


Figure I.7-15. *The Roofs of Collioure* in Figure I.7-14 illustrates how Matisse "paints in color" without using explicitly drawn edges that could darken the painting's perceived colors. These color patches generate amodal boundaries that can capture the surface colors into surface percepts that are meaningful to viewers. (Reprinted with permission from Grossberg, 2008.)

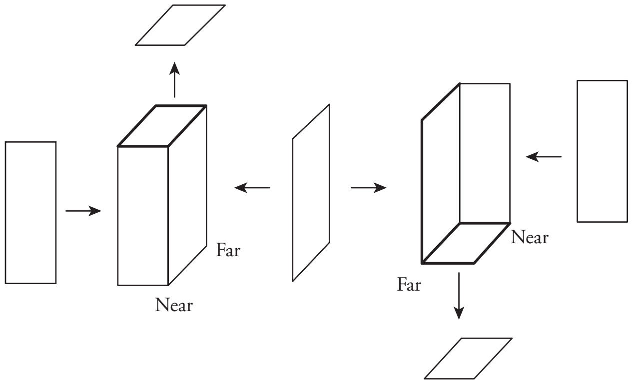
& Swaminathan, 2004; Grossberg & Yazdanbakhsh, 2005). Here, only one main point is noted using Figures I.7-16 and I.7-17 to illustrate how 3D illusions can occur.

The top image in Figure I.7-16 shows how a rectangular parallelopiped can be constructed in a 2D picture from sides that individually generate flat percepts of rectangles and parallelograms. Consider, in particular, the flat-looking parallelogram in the middle of the figure. If it is surrounded by sides as in the left-hand parallelopiped, then its lower side seems to project from near to far. However, if it is surrounded by sides as in the right-hand parallelopiped, then the same parallelogram's lower side appears to project from far to near. This simple demonstration shows that the percepts of the parallelopipeds are highly context-sensitive and depend on the way in which the sides are joined together at their shared edges.

When both types of parallelopipeds are joined together in the famous Necker cube at the bottom of the figure, then the 3D interpretations of both parallelopipeds can alternate through time. Focusing attention on one of the edges can, moreover, pull it forward and cause the side to which it belongs to look closer in space.

Figure I.7-17, which is due to Peter Tse, shows in a striking way that lines that are perceived as parallel when interpreted in 2D no longer look parallel when interpreted as part of a 3D percept, even though the entire figure is a 2D picture.

Grossberg and Swaminathan (2004) explain and quantitatively simulate such percepts, including the bistable 3D surface percepts of the Necker cube as they oscillate through time. In order to explain such percepts, the concept of bipole grouping was generalized from 2D to 3D, and the fact that it could be so generalized is another testimonial to its existence in the brain. In particular, Grossberg and Swaminathan (2004) show through computer simulations how bipole grouping cells can develop in response to visual inputs to represent collinear 2D and 3D groupings. The latter cells, called *disparity gradient cells*, can represent tilted edges such as those that occur in the percepts of the parallelopipeds. The same developmental laws gives rise to *angle cells* that are sensitive to the angular relationships that occur when two or more edges terminate at the same position in space. Interactions between angle cells, disparity gradient cells,



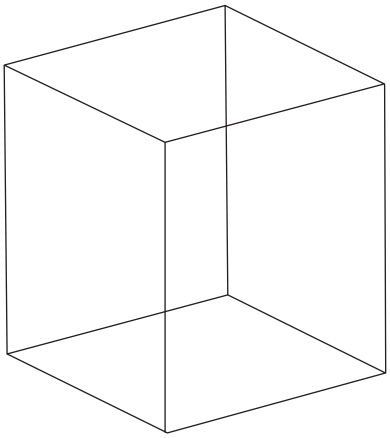


Figure I.7-16. The top figure shows how flat rectangles and parallelograms can generate 3D slanted percepts of parallelopipeds. (Reprinted with permission from Grossberg & Swaminathan, 2004.) The bottom figure of the Necker cube generates a bistable percept through time of both 3D interpretations of these parallelopipeds. (Reprinted with permission from 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 and Edinburgh Philosophical Magazine and Journal of Science.* **1** (5): 329–337.)

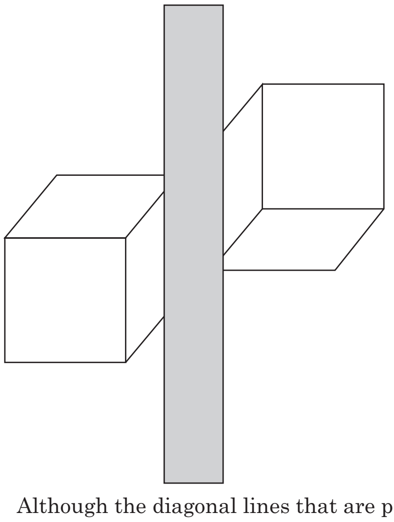


Figure I.7-17. Although the diagonal lines that are partially occluded by the vertical gray bar are colinear in 2D, they are not perceived to colinear as part of their respective parallelopipeds. (Reprinted with permission from Tse, 1999.)

and bipole cells give rise to the context-sensitive boundaries of the parallelopiped percepts. In addition, Grossberg and Swaminathan explain and simulate how these 3D boundaries capture surface filling-in that can occur over surfaces that are tilted in depth.

COMPLEMENTARY PROCESSING STREAMS IN VISUAL CORTEX: FORM, MOTION, AND FORMOTION

The organization of the visual cortex into parallel processing streams with complementary properties, and the way in which interactions across the streams generates consistent percepts, helps to explain many visual illusions. These complementary relationships may occur both within and across the "What" and "Where" visual cortical streams (Grossberg, 2000). For example, neural models explain perceptual and brain data about interactions of visual form and motion by explicating their complementary mechanisms (Baloch & Grossberg, 1997; Berzhanskaya, Grossberg, & Mingolla, 2007; Francis & Grossberg, 1996b), Object tracking and spatial navigation also use complementary mechanisms (Browning, Grossberg, & Mingolla, 2009a, 2009b; Grossberg, Mingolla, & Pack, 1999; Pack, Grossberg, & Mingolla, 2001).

The pale stripes of cortical area V2, also called interstripes, carry out 3D boundary grouping of object form, whereas the middle temporal, or MT, cortical area carries out motion processing. Grossberg (1991) predicted that the form and motion processing streams exhibit complementary properties due to the conflicting computational requirements of computing depth-selective object form versus those of computing directionally-selective object motion. According to this view, by itself, V2 can compute good estimates of object depth but only coarse estimates of object motion direction, whereas MT can compute good estimates of object motion direction but only coarse estimates of object depth. Form-to-motion interactions from V2 to MT, also called *formotion* interactions, were predicted to select motion direction signals in MT with the correct depths, and thus to enable good object tracking in depth to be carried out by subsequent processing stages of the Where cortical processing stream. This Grossberg (1991) prediction has been embodied in a model called the 3D FORMOTION model. The formotion prediction differs from that of Livingstone and Hubel (1988) and DeAngelis, Cumming, and Newsome (1998), who propose that precise depth estimates are created in MT. The Grossberg prediction has been supported by neurophysiological experiments of Ponce, Lomber, and Born (2008), who cooled V2, thereby eliminating the V2-to-MT influence, and recorded from MT to find the predicted good estimates of object motion but coarse estimates of object depth.

Formotion interactions have been used to explain many visual illusions. For example, in Grossberg (1991), they were used to explain the apparent motion of subjective surfaces (Fig. I.7-18) that was reported by Ramachandran, Rao, and Vidyasagar (1973) and Ramachandran (1985). To generate this percept, two visual frames alternate through time. In one frame, there are four pac-man figures that induce a Kanizsa square percept. To the right of the pac-men is a

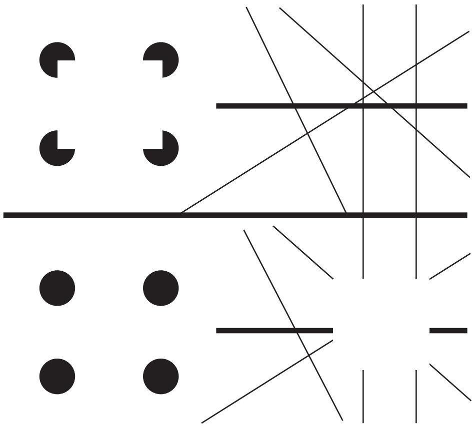


Figure I.7-18. Apparent motion of illusory contours. These images demonstrate that apparent motion of illusory contours arises through the interaction of the static illusory contours via formotion interactions that are predicted in Grossberg (1991) to occur from cortical area V2 to cortical area MT. Frame 1 (row 1) is temporally followed by Frame 2 (row 2) and conversely. (Reprinted with permission from Ramachandran, 1985.)

random array of intersecting oriented lines. In the second frame, the pac-men are replaced by four filled circles and the lines are deleted in the region of an imagined square, again producing an illusory square percept. By this construction, there are no features that can be matched across the two figures. Despite this fact, when the frames alternate through time, a percept of a square moving from left to right and back again is clearly perceived. Mechanisms of boundary completion and surface filling-in were used to explain the illusory square percepts. Mechanisms of long-range apparent motion (Grossberg & Rudd, 1992) were used to explain the motion percept in response to boundary signals in MT that were received from V2. Many other visual illusions have also been explained using these formotion mechanisms, including the line motion illusion, motion induction, and transformational apparent motion (Baloch & Grossberg, 1997).

COMPLEMENTARY PROCESSING STREAMS IN VISUAL CORTEX: MOTION PERCEPTION AND TRACKING

Area MT is broken into substreams (Born & Tootell, 1992): MT^- is involved in generating percepts of object motion and the tracking of objects moving relative to an observer, whereas MT^+ supports visually-based navigation as an observer moves with respect to the world. MT^- inputs to the ventral medial superior temporal (MSTv) cortex, whereas MT^+ inputs to the dorsal medial superior temporal (MSTd) cortex. These substreams obey computationally complementary subtractive versus additive laws, respectively, and interact to overcome their complementary deficiencies. How these complementary substreams interact to carry out object tracking and visual navigation is modeled by Browning et al. (2009a, 2009b), Elder, Grossberg, and Mingolla (2009), and Pack et al. (2001).

One example of the functional role of these interstream interactions occurs during the maintenance of smooth pursuit eye movements (SPEM). Such movements use a negative feedback system whereby the oculomotor system continuously attempts to match the velocity of the eye to that of the object being tracked. This process takes place in the MT^-–MSTv stream. What happens when this system stabilizes the target near the fovea? At these times, there is little or no motion of the target on the retina. Therefore, the pursuit system cannot rely wholly on retinal target velocity to maintain an accurate SPEM. A number of additional signals have been hypothesized to guide pursuit, including target position, target acceleration (Lisberger, Morris, & Tychsen, 1987) and a “memory” of target velocity (Young, Forster, & van Houtte, 1968), which is often described in terms of an oculomotor efference copy of eye velocity (von Holst, 1954). An efference copy duplicates the neural signal sent to the muscles that move the eye and carries information about the movement of the eye that is independent of the retinal image. An efference copy can hereby maintain a prediction of pursuit velocity from moment to moment. Thus the brain may combine retinal information about target motion or position with extraretinal information about the velocity of eye rotation.

An additional source of information relevant to pursuit maintenance is the existence of the visual background. A SPEM is typically made across a visual scene that contains stationary objects, and these objects sweep across the retinal image with a velocity opposite that of the target. This results in large-field coherent motion across the retina. Such large-field processing goes on in the MT^+–MSTd stream. It can trigger an involuntary eye rotation known as optokinetic nystagmus (OKN). OKN causes the eye to move in the same direction as the large stimulus, so that an OKN movement to track retinal motion of a visual background during pursuit would be in the opposite direction of the ongoing pursuit movement. As such, it is crucial that optokinetic signals be suppressed during execution of a SPEM.

The visual motion of the background can provide information about the velocity of ongoing SPEMs, even when the pursuit target is relatively stable on the retina. Such a signal could also be used to generate a prediction of target velocity. The visual motion of the background therefore has contradictory effects on the pursuit system, providing a potentially useful signal for pursuit maintenance and a potentially destructive OKN signal. Pack et al. (2001) have modeled how feedback from the MT^+–MSTd stream to the MT^-–MSTv stream can be used to maintain accurate tracking speed during periods when the moving target is successfully fixated.

The complementary organization of motion tracking and spatial navigation also leads to visual illusions. For example, during an accurate SPEM, target motion on the retina is very small, while objects in the background move across the retina. Psychophysical experiments indicate that human subjects are able to estimate the velocity of objects during a SPEM, but that this ability is somewhat limited. Specifically, observers underestimate the velocity of a moving target during a SPEM when no visible background is present (the Aubert-Fleischl phenomenon; Aubert, 1886) and perceive slight motion of a stationary visual background during a SPEM (the Filehne illusion; Filehne, 1922). Pack et al. (2001) simulate these illusions as consequences of the above interactions.

ALL OBJECT MOTION IS ILLUSORY: THE APERTURE PROBLEM

Essentially all percepts of moving objects are illusions, in the mechanistic sense that they all involve active reorganization and reconstruction of the visual signals received by the retina. For example, imagine a predator or prey darting intermittently behind protective cover in a forest. As the animal moves, patterns of light and shade play upon its body through the overhanging foliage. These moving regions mingle with the movements caused on the animal's body as its limbs and muscles deform the textured patterns on its coat. The luminance and color contours of these moving regions may move in a variety of directions that do not necessarily point in the direction of the animal's physical movement. Rather, a scintillating mosaic of moving contours may be generated that could easily prevent its detection, as in camouflage.

Detecting an animal under these rather typical conditions in the wild is obviously of great ecological importance to an observer, whether predator or prey. Doing so, however, requires that the visual system of the observer solve several basic perceptual problems, such as: How does the observer separate the forest cover from the moving animal, so that cover and animal can be independently recognized? How does the cover pop out in front of the animal so that the contours of the animal are not confused with those of the cover? How is the scintillating mosaic of moving contours reorganized into a coherent object percept with a unitary direction and speed of object motion? How are the intermittent appearances of the animal integrated into a continuous motion percept that is amodally completed behind the cover, thereby enabling the animal to more easily be tracked? How does this amodally completed motion percept adapt itself to the animal's variable speed of locomotion? How does attention track the animal's moving trajectory?

Grossberg (1998) reviews and unifies neural models that simulate how all of these processes work to generate motion percepts (e.g., Berzhanskaya et al., 2007; Chey, Grossberg, & Mingolla, 1997; Grossberg, 1994; Grossberg, Mingolla, & Viswanathan, 2001; Grossberg & Rudd, 1992; Kelly & Grossberg, 2000). This chapter briefly discusses two of these types of percepts because they are both visual illusions, at least from a mechanistic perspective, albeit illusions with high survival value: the percept of a continuous trajectory of the animal's movement behind the occluding cover, and the percept of the direction and speed of the animal's motion.

The process whereby intermittent views, or flashes, of the animal moving behind the forest cover are continuously interpolated as it moves at variable speeds is a type of long-range apparent motion. This process may use form-motion interactions to group the individual target "flashes" that are intermittently seen by the observer before these groupings are interpolated by the motion system. This sort of percept is illustrated in Figure I.7-18. Rather than being a laboratory curiosity, such long-range apparent motion is used to help track a target that is intermittently occluded.

The percept of object direction and speed can use a relatively small number of unambiguous and correct motion signals, called *feature tracking* signals, to "capture" compatible motion signals across the entire object and to suppress a much larger number of incompatible motion signals. In particular, the form boundaries of the moving animal's outer contours may generate feature tracking signals, which propagate to the ambiguous scintillating mosaic within these contours to capture compatible motion signals and suppress incompatible ones. This is a visual illusion that transforms the actual motion inputs received from the animal into a different motion percept. Doing so requires a solution of one of the most important problems facing the motion system, namely, the *aperture problem*.

Wallach (1935; see also Wuerger, Shapley, & Rubin, 1996) further analyzed the observation of Stumpf in 1911 that the motion of a featureless line seen within a circular aperture is perceptually ambiguous: Given any real direction of motion of the line, its perceived direction is perpendicular to its orientation. This phenomenon was called the aperture problem by Marr and Ullman (1981). The aperture problem is faced by any localized neural motion sensor, such as a neuron in the early visual pathway, which responds to a moving local contour through an aperture-like receptive field. Only when the contour within an aperture contains features, such as line terminators, object corners, or high contrast blobs or dots, can a local motion detector accurately measure the direction and velocity of motion (Shimojo, Silverman, & Nakayama, 1989). For example, when the aperture is rectangular, as during the barberpole illusion (Wallach, 1935), moving lines may appear to move in the direction parallel to the longer edges of the rectangle within which they move, even if their actual motion direction is not parallel to these edges. The brain must solve the aperture problem, despite the fact that every neuron's receptive field defines an "aperture," in order to detect the correct directions of moving objects in the world. The examples of circular and rectangular apertures, or occluders, provide important cues about how the brain often succeeds in doing this in the real world.

When an object moves behind multiple occluders, aperture ambiguities can again lead to nonveridal percepts of its real motion. Despite the fact that the object may be segmented into many visible parts by the occluders, the visual system can often integrate these parts into a percept of coherent object motion that crosses the occluders. Studying conditions under which the visual system can and cannot accomplish this provides important cues to the processes that are used by the visual system to create useful and predictive object motion percepts during normal viewing conditions.

The aperture problem shows that, under many viewing conditions, the visual system perceives an illusory direction of motion that is perpendicular to a contour's orientation. The brain tries to solve the aperture problem by transforming one illusion into another. The latter illusion, however, helps to predict the direction and speed in which objects really move.

As noted earlier, the visual system uses the relatively few unambiguous motion signals arising from image features, called feature tracking signals, to select the ambiguous motion signals that are consistent with them, while suppressing the more numerous ambiguous signals that are inconsistent with them. For example, during the barberpole illusion, feature tracking signals from the moving line ends along the longer edges of the bounding rectangle of the barberpole compute an unambiguous motion direction. These feature tracking signals can become strong in the brain's motion processing stream within cortical areas V1, MT, and MST because directionally-selective cells can accumulate consistent evidence for that motion direction through time. In contrast, within the interior of a moving line, multiple directions become partially activated, and these compete using a self-normalizing competition, thereby weakening each other.

Later feedback processes between MT and MST enable these sparse, but relatively strong, feature tracking signals to gradually propagate through time into the interior of the rectangle. This *motion capture* process selects the feature tracking motion direction from the ambiguous directions along the lines within the rectangle. Motion capture also suppresses the other ambiguous motion signals (Ben-Av & Shiffrar, 1995; Bowns, 1996, 2001; Castet, Lorenceau, Shiffrar, & Bonnet, 1993; Chey et al., 1997; Chey, Grossberg, & Mingolla, 1998; Grossberg et al., 2001; Lorenceau & Gorea, 1989). When a scene does not contain any unambiguous motion signals, the ambiguous motion signals cooperate to compute a consistent object motion direction and speed (Grossberg et al., 2001; Lorenceau & Shiffrar, 1992). Chey et al. (1997) simulated the time course of such motion capture, and Pack and Born (2001) described neurophysiological data from cortical area MT of monkeys that fit this predicted time course.

MOTION CAPTURE AND THE TEMPORAL DYNAMICS OF DECISION-MAKING

The way in which the 3D FORMOTION model solves the aperture problem through motion capture helps to explain many visual illusions of motion. In particular, it sheds new light on how the brain makes movement decisions by selecting the direction of saccadic eye movements to match the direction of probabilistically defined motion stimuli (Grossberg & Pilly, 2008). This type of movement decision is based on a visual illusion whereby the brain converts an image that consists mostly of randomly moving dots, with a subset of dots briefly moving in a fixed direction, into an estimate of a coherent direction of dot motion. This conversion seems, once again, to be due to the brain's mechanisms for solving the aperture problem.

It is well known that speed and accuracy of perceptual decisions covary with certainty in the input, and correlate with the rate of evidence accumulation in parietal and frontal cortical neurons. The 3D FORMOTION model creates an emerging estimate of the direction of motion as motion capture sets in. This estimate needs to be converted into an estimate of directional eye movement. The Motion Decision (MODE) model (Fig. I.7-19) adds to 3D FORMOTION processing a subsequent processing stage for making the coordinate change from visual motion direction to eye movement direction. This subsequent stage is interpreted to occur in the lateral interparietal region (LIP) of the parietal cortex, where a number of neurophysiologists have recorded cells whose dynamics covary with the behaviorally observed properties of saccadic eye movements in response to the moving dots (e.g., Roitman & Shadlen, 2002; Shadlen & Newsome, 2001).

This LIP circuit is modeled using a kind of decision circuit that has become classical in the neural modeling literature, namely, a network of cells that obey the membrane equation, or shunting, dynamics that are known to occur in neurons, when they interact via a recurrent on-center off-surround network. Such a network is often called a recurrent competitive field (RCF), and its variants have been used by a number of authors to model the dynamics of perceptual or motor decisions in both deterministic models (e.g., Brown, Bullock, & Grossberg, 2004; Chey et al., 1997; Francis, Grossberg, & Mingolla, 1994; Francis & Grossberg, 1996a, 1996b) and stochastic models (Boardman, Grossberg, Myers, & Cohen, 1999; Cisek, 2006; Grossberg, Roberts, Aguilar, & Bullock, 1997; Usher & McClelland, 2001). Grossberg (1973) was the first to mathematically define and analyze these networks, proving along the way how they can make winner-take-all and distributed decisions. See Grossberg (1980) for a heuristic review of these early results. In order to explain the motion decision-making data, the RCF network needs to have its outputs gated by the basal ganglia. How the basal ganglia can gate the release of saccadic eye movement outputs was modeled in detail by Brown, Bullock, and Grossberg (1999, 2004). Cisek, Puskas, and El-Murr (2009) have also modeled how the buildup of evidence for movement decisions depends on basal ganglia gating.

Newsome, Shadlen, and colleagues studied neural correlates of perceptual decision-making in macaque monkeys who were trained to discriminate motion direction. Random dot motion displays covered a 5° diameter aperture centered at the fixation point. Motion coherence was controlled by varying the fraction of dots moving nonrandomly in a particular direction from one frame to the next in each of three interleaved sequences. Varying motion coherence provided a quantitative way to control the ambiguity of directional information that the monkey used to make a saccadic eye movement to a peripheral choice target in the perceived motion direction, and thus the task difficulty. Apart from being few in number, these correct directional signals also have a short life span because a new set of signal dots are chosen in every frame.

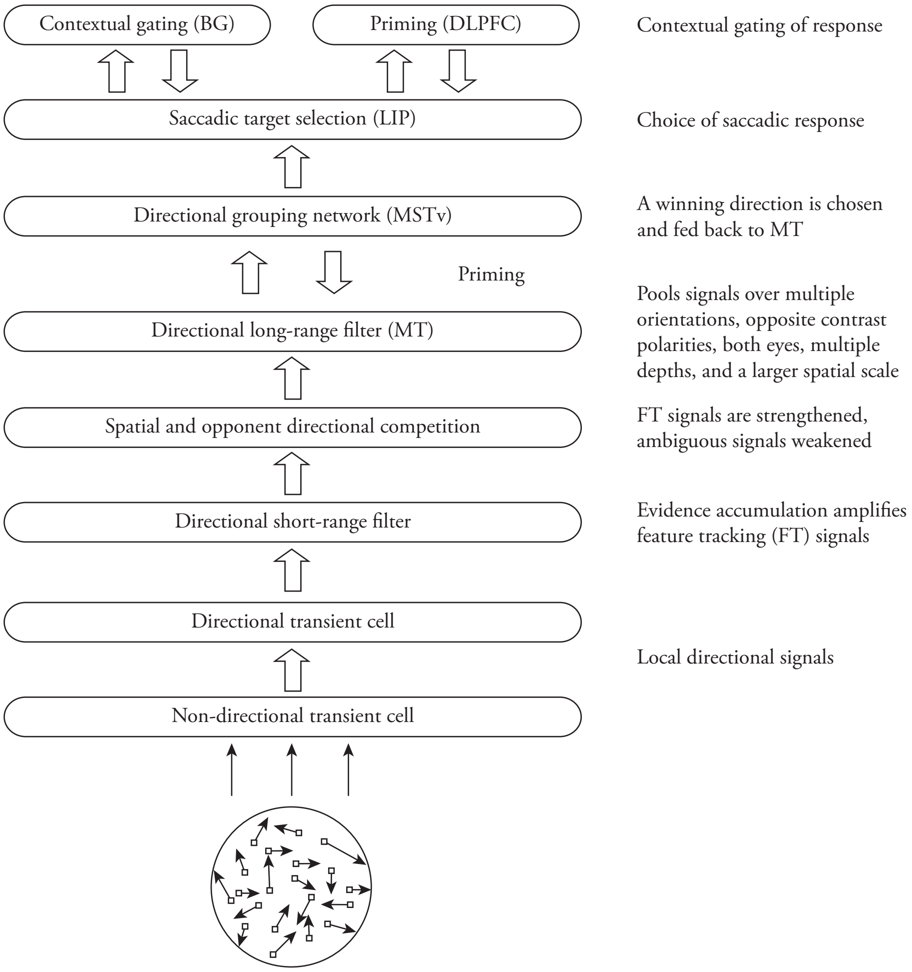


Figure I.7-19. MODE model processing stages retina/LGN–V1–MT–MST–LIP–BG. (Reprinted with permission from Grossberg & Pilly, 2008.)

Neurophysiological recordings were done in LIP while the monkeys performed these tasks. Figures I.7-20 though I.7-22 summarize some of the behavioral and LIP neurophysiological data, along with the MODE model simulations of them. These simulations are included to illustrate the power of current neural models for quantitatively explaining and predicting perceptual and brain data, including data about visual illusions. Note that, despite the discrete and probabilistic nature of the input stimuli, both the behavioral and neurophysiological data change continuously in time, reflecting the growing coherence of the neuronal responses due to motion capture interactions.

CONTEXT CAN CREATE AMBIGUITY AND THE INFORMATION NEEDED TO OVERCOME IT

Why does a solution of the aperture problem help to solve a movement decision problem in response to moving dots, given that individual dots have unambiguous motion directions and thus can generate feature tracking signals? Why is motion capture needed when many dots move in the same image? A moment's thought clarifies that a similar question needs to be asked when there are multiple moving elements in any scene. Our answer to this question illustrates how context can create computational ambiguities, as well as provide information that brain interactions can use to overcome them, and thereby generate "illusory" perceptual representations that are more informative than local cues could ever be.

The directional filters that accumulate directional evidence to generate feature tracking signals in response to a single dot can be fooled when there are multiple dots in each frame and some dots move incoherently. These directional filters can generate local directional signals between *any* two dots that occur with an appropriate spatiotemporal displacement. Typically some directions will be amplified more than others, but lower motion coherence, higher dot density, and more interleaving of stimulus frames increase the probability that incorrect directional signals will be generated, thereby reducing the impact of correct local groupings in determining a clear motion directional percept. In addition, apart from being few in number, these correct directional signals also have a short life span because a new set of signal dots are chosen every frame.

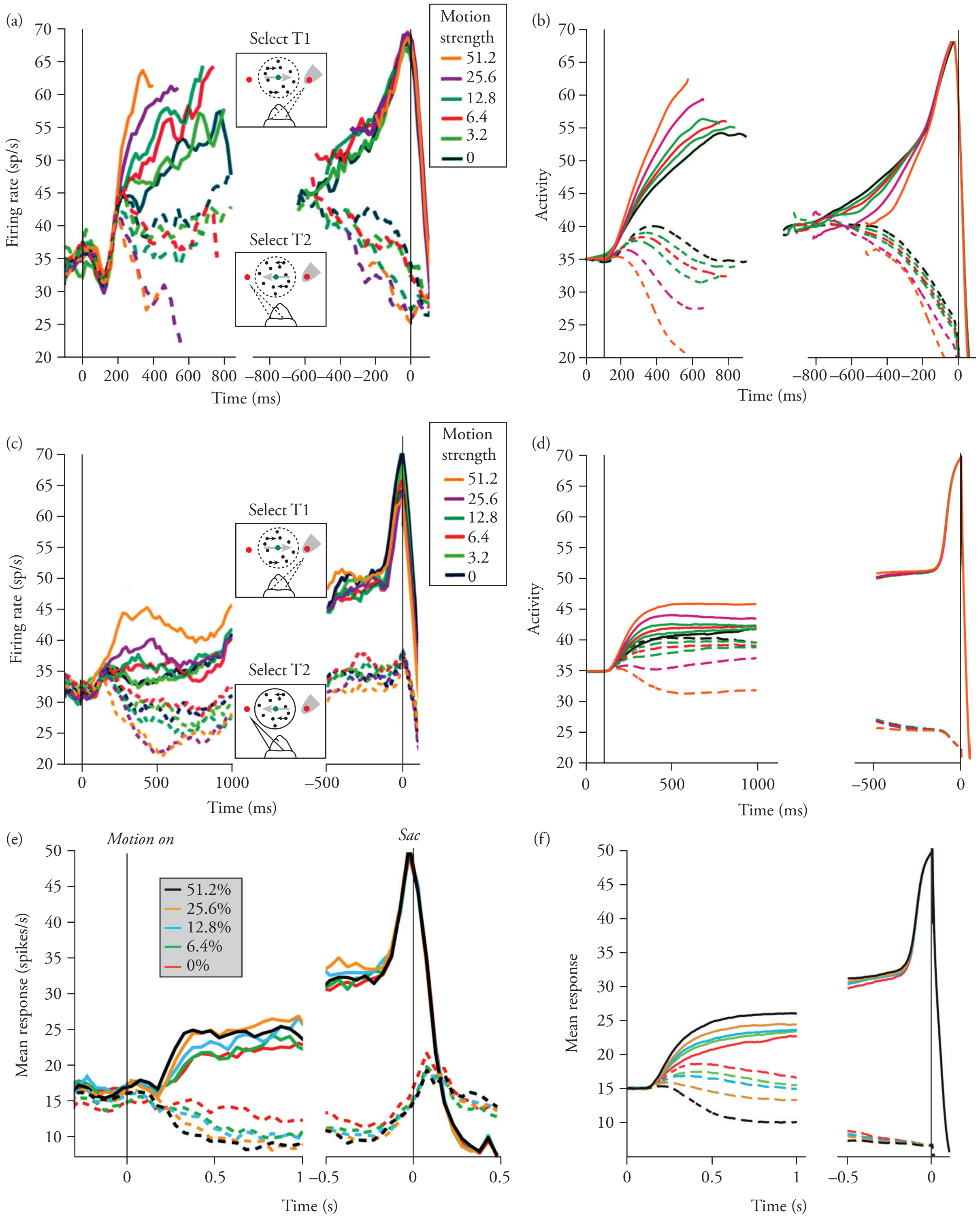


Figure I.7-20. Temporal dynamics of LIP neuronal responses during fixed duration (FD) and reaction time (RT) tasks. (a) Average responses of a population of LIP neurons among correct trials during the RT task. (b) Model simulations replicate LIP cell recordings during the RT task. (c) Average responses of a population of LIP neurons among correct trials during the 2002 FD task. (d) Model simulations mimic LIP cell recordings during the FD task. (e) Average responses of a population of LIP neurons among correct trials during another FD task. (Reprinted with permission from Grossberg & Pilly, 2008.)

Therefore, the motion stream must somehow enable a relatively sparse set of short-lived and correct feature tracking signals to gradually discount the more numerous incorrect local directional groupings. The motion capture process that is predicted to occur as a result of motion grouping feedback signals between MT and MST can do this, with the results shown in Figures I.7-18 through I.7-20.

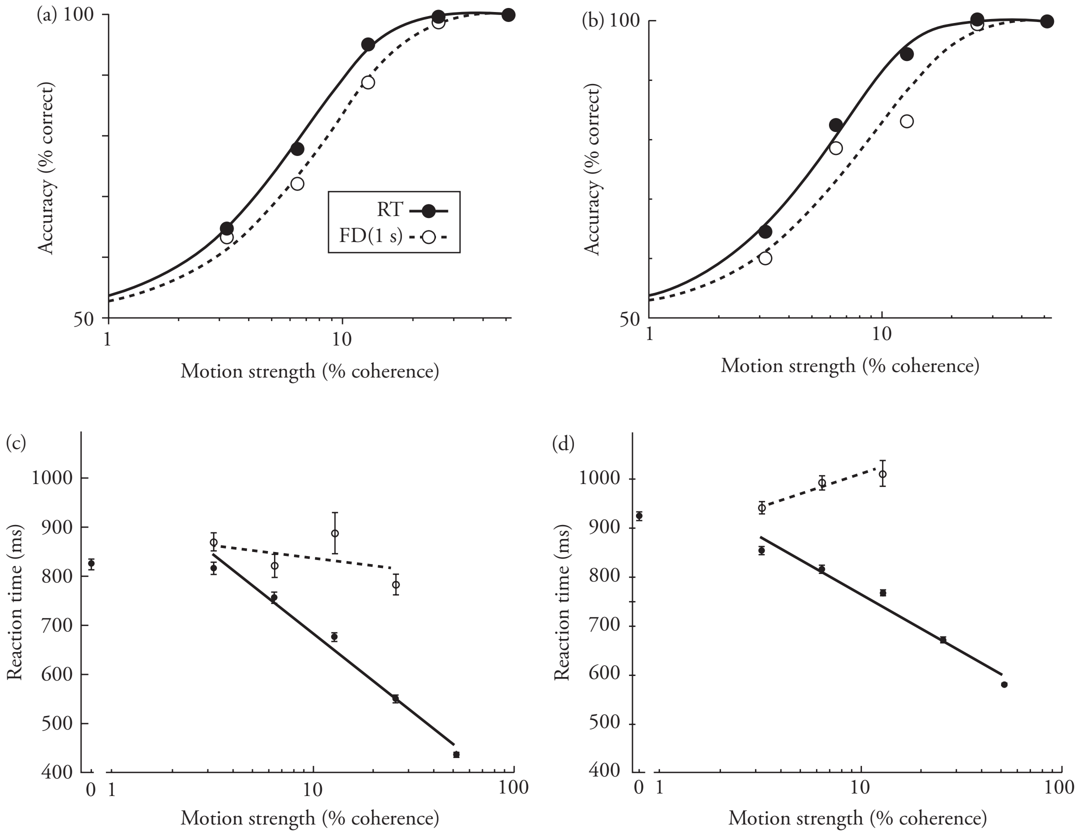


Figure I.7-21. Psychometric and chronometric data during the FD and RT tasks. (a) Accuracy data (percentage correct) as a function of motion coherence. (b) Model simulations emulate these data. (c) Reaction time data (ms) as a function of motion coherence (percent certainty). (d) Model simulations emulate the data on both correct and error trials. In (a–d), the abscissa is in the $\log_{10}$ scale. (Reprinted with permission from Grossberg & Pilly, 2008.)

HOW DOES THE BRAIN COPE WITH A PROBABILISTIC WORLD?

Earlier we mentioned the contributions of Kanizsa to better understanding how primarily bottom-up visual processes contribute to visual perception, whereas those of Helmholtz were particularly relevant to top-down processes. Helmholtz's ideas about top-down expectations, or unconscious inferences, were precursors to the use of probabilistic Bayesian concepts by some recent vision theorists.

It is clear why probabilistic concepts are appealing in trying to understand mind and brain, if only because there is a lot of uncertainty in the world, and probability concepts try to conceptualize that uncertainty. But does the brain embody *classical* probability concepts, or does it deal with uncertainty in a new way? In this regard, it should be noted that the combination of shunting dynamics and competitive interactions with a RCF lead to automatic gain control and self-normalizing properties (Grossberg, 1973, 1980), so that the total activity of a processing channel in such a model is often approximately conserved. This total activity plays the role of a *real-time probability distribution*, that is, a deterministically computed distribution of normalized activity through time. Self-normalization enables such a network to maintain its sensitivity in response to distributed inputs whose total number and size can vary wildly through time. Grossberg (1973) called this property a solution of the *noise-saturation dilemma*. This property is useful when estimating the directional coherence of inputs whose total number and distribution can vary randomly through time, as occurs in the motion decision-making experiments. The robustness of this property when processing noisy data is illustrated by the fact that MODE model properties remain qualitatively unchanged when cellular noise is added to the model MT and MST processing stages. Indeed, all the processing stages of the MODE model—indeed all the models from the Grossberg school of modeling—use shunting neurons with self-normalizing properties in on-center off-surround networks. Thus self-normalizing neural dynamics capture key properties of classical probability theory, but they go beyond classical probability theory by being able to compute coherent, context-sensitive, time-varying measures that can exploit global properties of a scene.

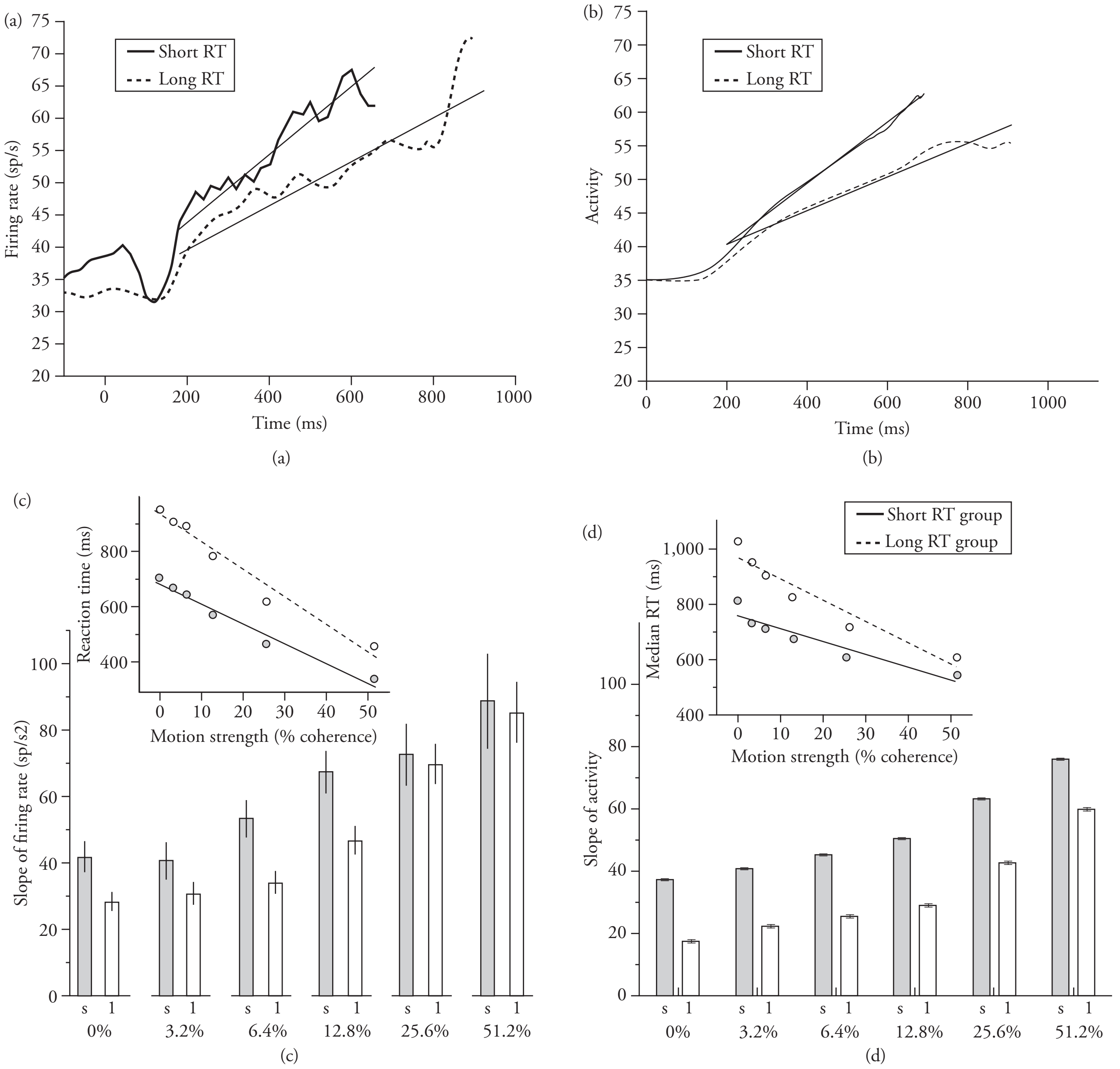


Figure I.7-22. Relationship between LIP response and reaction time. (a) Average LIP responses. (b) Model simulations reflect this relationship. (c) The histogram shows for each coherence level and group the slope of the linear fit to the average LIP response. (d) Model simulations reproduce these data trends. In panels (c) and (d), the error bars represent 95% confidence intervals. (Reprinted with permission from Grossberg & Pilly, 2008.)

When these self-normalizing "probabilistic" properties interact with the coherence-producing, aperture-problem-reducing, motion capture properties of the MT–MST grouping network, MODE model LIP cell activities covary quantitatively with the activation rate and amplitude of LIP cells that control a monkey's observable behavior in these experiments. The MODE model hereby clarifies how brain circuits that solve the aperture problem, notably the circuits that realize motion capture, control properties of probabilistic decision-making in real time. These directional decision-making data, and the motion capture process that enables it, can hereby be understood as a resolution of ambiguity that selects the best consensus movement that is compatible with motion data.

DOES BAYESIAN INFERENCE SHAPE BRAIN DESIGN?

The question of whether the mathematical laws of Bayesian inference shape brain design is of particular relevance to Newsome and Shadlen's data (e.g., Roitman & Shadlen, 2002; Shadlen & Newsome, 2001), since these investigators have proposed that perception and decision-making can be described as Bayesian inference, which estimates the optimal interpretation of the stimulus given priors and likelihoods. Such an idea also seems appealing to people who believe that visual illusions are the brain's way of generating the optimal interpretation of the stimulus. However, Bayesian concepts do not provide a way to discover the brain mechanisms that make decisions, or indeed any behavioral and brain mechanisms. The Bayes rule, and related statistical concepts, are so general that they may be applied to problems in any science.

In particular, the Bayes rule follows from writing the probability of any two events I and S, namely *p* (I,S), in two different ways in terms of conditional probabilities, dividing both sides of this identity by a term on one side of the equation and then optimizing the result. In applications to perception, I is a stimulus that falls on the retina and the Bayesian model asks what percept S is most likely to have caused it. The basic idea is to find the S that maximizes the posterior probability *p* (S/I) via the Bayes rule.

This generality is part of its broad appeal but is also its weakness in not providing specific design principles or mechanisms that govern any particular science. Those who believe it can must ask themselves if the Bayes rule is sufficient, for example, to derive the laws of general relativity or high-energy physics. If not, then why can it do so for the brain?

The neural models that are discussed in this chapter have explained data that Bayesian models have heretofore failed to explain, do so in terms of basic brain design principles and mechanisms without an appeal to Bayesian inference, and, unlike other available models of these data, generate perceptual representations in response to the experimental visual stimuli. In particular, two neural design principles go into the MODE model, one that solves the aperture problem, thereby simulating neurophysiological data about interactions of cortical areas V1, MT, and MST, and the other that solves the noise-saturation dilemma and how this solution enables time-varying choices. Joining together these two designs enables simulation of the time course of LIP neurophysiological data, as well as of behavioral accuracy and reaction time properties, during both correct and error trials at different levels of input ambiguity in both fixed duration and reaction time tasks.

The succession of model processing stages V1–MT–MST embodies a hierarchical resolution of the directional uncertainties caused by the aperture problem. Model MST computes the global direction of random dot motion stimuli as the final stage of motion capture, while model LIP responds to MST outputs to compute the directional perceptual decision that leads to a saccadic eye movement. The intuitive idea is that the MT–MST feedback loop needs more time to capture the incoherent motion signals when there are more of them. In addition, this loop cannot achieve as high a level of asymptotic response magnitude, due to self-normalizing competition, when more incoherent motion signals compete with the emerging winning direction. In other words, the effectiveness of the motion capture process depends on input coherence. LIP then converts the inputs from MST into an eye movement directional command and thereby enables the (model) monkey to report its decision via a saccade. This self-organizing system thus trades accuracy against speed, with the brain running as fast as it can in response to the amount of uncertainty in the data that it is processing.

These various results cumulatively illustrate how cortical dynamics go beyond Bayesian concepts. They also clarify why probability theory ideas are initially so appealing in trying to explain visual percepts in general, and visual illusions in particular. Indeed, the brain is a self-organizing system that is designed to adapt autonomously to changing environments that include unexpected and rare events that may have no priors. These are not the types of processes for which Bayesian ideas are most appropriate. The brain needs to be studied and understood on its own terms, and we are well along the way to doing so, with visual illusions as one of many guides.

ART MODELS OF BRAIN LEARNING, RECOGNITION, AND CONSCIOUS PERCEPTION

If the Bayes rule is not the best way to represent top-down effects on perception, then what is? As the FACADE model was being developed, a parallel stream of modeling was aimed at understanding the neural mechanisms that support conscious experiences, notably how conscious visual percepts may arise. Such explanations are developed in Adaptive Resonance Theory (ART), which was introduced in Grossberg (1976a, 1976b). ART predicts that "all conscious states are resonant states." The 3D LAMINART theory describes a synthesis of ideas from the FACADE model about how 3D vision works in laminar visual cortical circuits that also embody ART mechanisms.

ART is a cognitive and neural theory of how the brain autonomously learns to attend, categorize, recognize, and predict objects and events in a changing world. To a remarkable degree, humans and other primates can rapidly learn new facts without being forced to just as rapidly forget what they already know. As a result, we can confidently go out into the world without fearing that, in learning to recognize a new friend's face, we will suddenly forget the faces of our family and friends. Otherwise expressed, ART avoids the problem of *catastrophic forgetting*.

ART accomplishes this by proposing a solution to the *stability-plasticity dilemma*, or how brains can learn quickly without also catastrophically forgetting already learned memories just as quickly. In so doing, ART clarifies key brain processes from which conscious experiences emerge. It describes a functional link between processes of consciousness, learning, expectation, attention, resonance, and synchrony (the CLEARS processes). ART predicted that all brain representations that solve the stability–plasticity dilemma use variations of CLEARS mechanisms (Grossberg, 1978, 1980, 2007, 2013). Through these CLEARS connections, ART clarifies why many animals are intentional beings who pay attention to salient objects, why "all conscious states are resonant states," and how brains can learn both *many-to-one maps* (representations whereby many object views, positions, and sizes all activate the same invariant object category) and *one-to-many maps* (representations that enable us to expertly know many things about individual objects and events).

ATTENTION, MATCHING, BIASED COMPETITION, RESONANCE, AND LEARNING

The views of Helmholtz enter the picture because ART accomplishes these properties by proposing how top-down expectations focus attention on salient combinations of cues, called "critical feature patterns", and characterizes how attention may operate via a form of self-normalizing "biased competition" (Desimone, 1998). In addition to anticipating the concept of biased competition by two decades, the exact form used for such attentive matching in explaining perceptual data (e.g., Bhatt, Carpenter, & Grossberg, 2007; Gove, Grossberg, & Mingolla, 1995) has been recently supported by the same form factor in the "normalization model of attention" (Reynolds & Heeger, 2009).

ART has developed Helmboltz's intuitions about unconscious inferences into rigorous science. Indeed, an active top-down expectation in ART plays the role of an unconscious inference. Top-down attentive matching in ART was predicted to be carried out by a top-down, modulatory on-center, off-surround network (Carpenter & Grossberg, 1987, 1991; Grossberg, 1980, 1999b) whose realization by identified cells within the laminar circuits of visual cortex has been described (Fig. I.7-23; Grossberg, 1999a; Grossberg & Versace, 2008; Raizada & Grossberg, 2001). The modulatory on-center in this top-down matching mechanism explains why unambiguous bottom-up data cannot usually be overridden by an active top-down expect at Matching Rule are combined with FACADE and 3D LAMINART mechanisms, many of the stresses between the views of Helmboltz and Kanizsa are resolved.

ART explains how such top-down attentive matching may help to solve the stability–plasticity dilemma by regulating cycles of *resonance* and *reset*, that is, of attentive matching and hypothesis testing (Fig. I.7-24). In particular, when a good enough match occurs between bottom-up

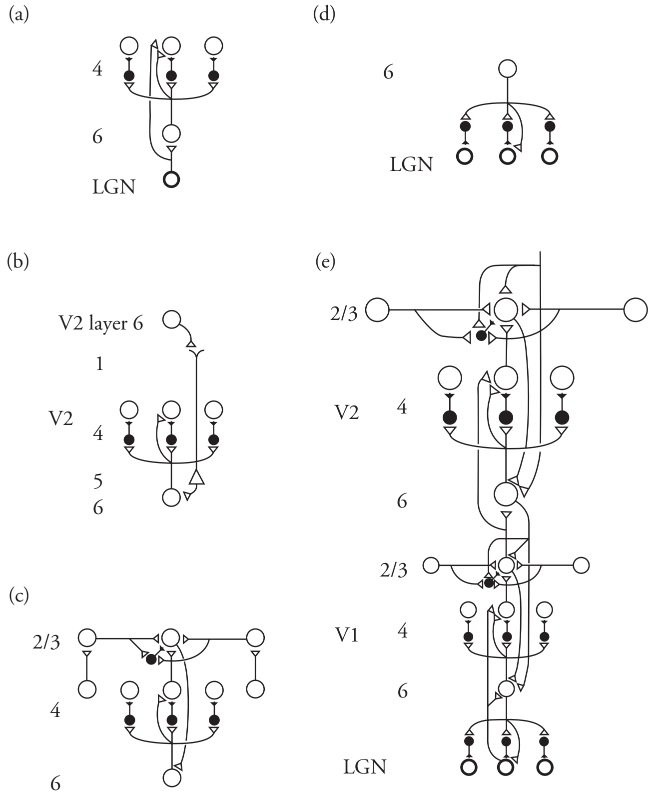


Figure I.7-23. Model LAMINART circuitry: How known cortical connections integrate bipole grouping with the ART Matching Rule for top-down attention. Inhibitory interneurons are shown filled-in black. (a) The LGN provides bottom-up activation to layer 4 via two routes. First, it makes a strong connection directly into layer 4. Second, LGN axons send collaterals into layer 6 and thereby also activate layer 4 via the 6 → 4 on-center off-surround path. This is an example of spatial competition. The combined effect of the bottom-up LGN pathways is to stimulate layer 4 via an on-center off-surround, which provides divisive contrast normalization. (b) "Folded feedback" carries attentional signals from higher cortex into layer 4 of V1, via the modulatory 6 → 4 path. Corticocortical feedback axons tend preferentially to originate in layer 6 of the higher area and to terminate in layer 1 of the lower cortex, where they can excite the apical dendrites of layer 5 pyramidal cells whose axons send collaterals into layer 6. The triangle in the figure represents such a layer 5 pyramidal cell. Several other routes through which feedback can pass into V1 layer 6 exist. Having arrived in layer 6, the feedback is then "folded" back up into the feedforward stream by passing through the 6 → 4 on-center off-surround path. This circuit realizes the top-down, modulatory on-center, off-surround circuit of the ART Matching Rule. (c) Connecting the 6 → 4 on-center off-surround to the layer 2/3 bipole grouping circuit: like-oriented layer 4 simple cells with opposite contrast polarities compete (not shown) before generating half-wave rectified outputs that converge onto layer 2/3 complex cells in the column above them. The long-range interactions within layer 2/3 realize the bipole grouping property. Just like attentional signals from higher cortex, as shown in (b), groupings that form among bipole cells in layer 2/3 also send activation into the folded feedback path, to enhance their own positions in layer 4 beneath them via the 6 → 4 on-center and to suppress input to other groupings via the 6 → 4 off-surround. There exist direct layer 2/3 → 6 connections in macaque V1, as well as indirect routes via layer 5. (d) Top-down corticogeniculate feedback from V1 layer 6 to LGN also has an on-center off-surround anatomy, similar to the 6 → 4 path, and realizes the ART Matching Rule from V1 to LGN. The on-center feedback selectively enhances LGN cells that are consistent with the activation that they cause, and the off-surround contributes to length-sensitive (endstopped) responses that facilitate grouping perpendicular to line ends. (e) The entire V1/V2 circuit: V2 repeats the laminar pattern of V1 circuitry but at a larger spatial scale. In particular, the horizontal layer 2/3 connections have a longer range in V2, allowing above-threshold perceptual groupings between more widely spaced inducing stimuli to form. V1 layer 2/3 projects up to V2 layers 6 and 4, just as LGN projects to layers 6 an 4 of V1. Higher cortical areas send feedback into V2, which ultimately reaches layer 6, just as V2 feedback acts on layer 6 of V1. Feedback paths from higher cortical areas straight into V1 (not shown) can complement and enhance feedback from V2 into V1. Top-down attention can also modulate layer 2/3 pyramidal cells directly by activating both the pyramidal cells and inhibitory interneurons in that layer. The inhibition tends to balance the excitation, leading to a modulatory effect. These top-down attentional pathways tend to synapse in layer 1. Their synapses on apical dendrites in layer 1 are not shown, for simplicity. (Reprinted with permission from Raizada & Grossberg, 2001.)

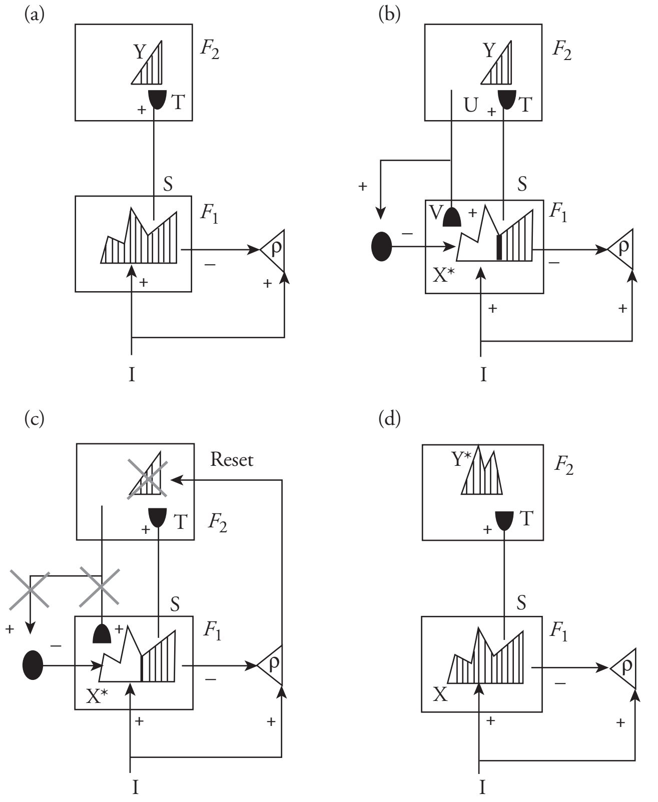


Figure I.7-24. How ART searches for a recognition category using cycles of resonance and reset. (a) Input pattern *I* is instated across feature detectors at level F_1 as an activity pattern *X*, while it nonspecifically activates the orienting system *A* with gain ρ. *X* inhibits *A* and generates output pattern *S*. *S* is multiplied by learned adaptive weights to form the input pattern *T*. *T* activates category cells *Y* at level F_2. (b) *Y* generates the top-down signals *U*, which are multiplied by adaptive weights and added at F_1 cells to form a *prototype V* that encodes the learned expectation of active F_2 categories. If *V* mismatches *I* at F_1, then a new STM activity pattern *X** (the hatched pattern) is selected at F_1. *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." Arousal resets F_2 by inhibiting *Y*. (d) After *Y* is inhibited, *X* is reinstated and *Y* stays inhibited as *X* activates a different activity pattern *Y**. A search for better F_2 category continues until a better matching or novel category is selected. When the search ends, an attentive resonance triggers learning of the attended data. (Adapted with permission from Carpenter & Grossberg, 1993.)

inputs and a top-down expectation, then a synchronous resonant state emerges that embodies an attentional focus and is capable of driving fast learning of bottom-up recognition categories and top-down expectations—hence the name *adaptive* resonance. If the match is not good enough, then the currently active recognition category is reset by a complementary orienting system, and interactions between the attentional and orienting systems drive a search for a new, or better-matching, category, as schematized in Figure I.7-24.

All of the main ART predictions have received increasing support from psychological and neurobiological data since ART was introduced in 1976 (see Grossberg, 1999b, 2003; Grossberg & Versace, 2008; and Raizada & Grossberg, 2003, for reviews). Synchronous resonances are, in particular, expected to occur between multiple cortical and subcortical areas. Recent data support this prediction (e.g., see Buschman & Miller, 2007; Engel, Fries, & Singer, 2001; Grossberg, 2009b; and Pollen, 1999, for reviews).

CONSCIOUS VISUAL PERCEPTION IS A SURFACE-SHROUD RESONANCE

ART predicts that "all conscious states are resonant states," and FACADE theory predicts that all consciously visible percepts are surface percepts (Fig. I.7-6). This raises the question: What sort of resonances support conscious visible surface percepts, whether real or illusory? An answer to this question arose indirectly through a study of how the brain learns to recognize objects.

The ARTSCAN model clarifies how we can learn to recognize objects when we see them from multiple perspectives (e.g., Cao et al., 2011; Fazl et al., 2009). ARTSCAN shows how this may be accomplished by coordinating spatial attention and eye movements in the Where cortical processing stream (Goodale & Milner, 1992) with object attention and learning in the What cortical stream (Mishkin, Ungerleider, & Macko, 1983). A *surface-shroud resonance*, or synchronous feedback interaction, between cortical area V4 and parietal cortex can maintain spatial attention on an object's surface while the eyes explore it and drive the learning of an invariant object category. I have predicted that such a surface-shroud resonance supports conscious perception of the object's surface. After such a resonance locks attention upon a surface, it can also propagate top-down to lower visual areas, such as V1 and V2, and bottom-up to regions like prefrontal cortex, to synchronize them all to support the conscious percept.

HABITUATIVE GATES IN PERCEPTUAL AND RECOGNITION PROCESSES: AFTERIMAGES

Some visual illusions, such as afterimages, occur after the inducing stimuli shut off. Why do afterimages occur? In all the FACADE, 3D LAMINART, and ART models of perception and recognition, there exist activity-dependent habituative gating processes at the synapses of various connections. These gating processes may be realized either at presynaptic transmitters or postsynaptic receptors. The laws for habituative gates were introduced in Grossberg (1968, 1969). Neurobiological data and modeling in support of them from visual cortex were reported by Abbott, Varela, Sen, and Nelson (1997) in visual cortex and by Tsodyks and Markram (1997) in somatosensory cortex, using the names synaptic depression and dynamic synapses, respectively. These gating processes play several related functional roles. During the development of cortical maps, they prevent perseverative activation of the same cells and enable new inputs to learn how to activate new cells across the developing map (e.g., Grossberg & Seitz, 2003; Olson & Grossberg, 1998). During percepts of changing visual inputs, they limit persistent activation of the same cells and enable percepts not to be pathologically smeared across a scene in response to moving objects (e.g., Francis & Grossberg, 1996a; Francis et al., 1994). During the learning of recognition categories, they enable the brain to reset categories whose top-down expectations mismatch bottom-up input patterns and to search for better-matching categories (e.g., Carpenter & Grossberg, 1987, 1991; Grossberg, 1980).

These adaptive properties also enable various illusions to occur. For example, the same activity-dependent habituative properties that limit persistence and help to drive the search for more predictive recognition categories can cause negative aftereffects and residual traces to be seen when images are suddenly shut off after being inspected for a while (Francis & Grossberg, 1996a). Habituative gating also plays a key role in generating bistable percepts, such as the Necker cube (Fig. I.7-16), since activity-dependent habituation of the pathways that support one percept can enable a competing percept to become dominant for a while (e.g., Grossberg & Swaminathan, 2004; Grossberg & Yazdanbakhsh, 2005; Grossberg et al., 2008).

In summary, both bottom-up and top-top mechanisms, and their intimate interaction, are needed to perceive and learn to recognize the world, and adaptive mechanisms of both types can cause the visual illusions that we recognize as such, or know, as well as those that we don't.

ACKNOWLEDGMENTS

This research was supported in part by the SyNAPSE program of DARPA (HR0011-09-C-0001).

REFERENCES

Abbott, L. F., Varela, K. Sen, K., & Nelson, S. B. (1997). Synaptic depression and cortical gain control. *Science*, *275*, 220–223.

Aubert, H. (1886). Die bewegungsempfinung. *Pfluggers Archive*, *39*, 347–370.

Baloch, A. A., & Grossberg, S. (1997). A neural model of high-level motion processing: Line motion and formotion dynamics. *Vision Research*, *37*, 3037–3059.

Beck, J., Prazdny, K., & Rosenfeld, A. (1983). A theory of textural segmentation. In J. Beck, B. Hope, & A. Rosenfeld (Eds.), *Human and machine vision* (pp. 1–38) New York: Academic Press.

Ben-Av, M. B., & Shiffrar, M. (1995). Disambiguating velocity estimates across image space. *Vision Research*, *35*, 2889–2895.

Berzhanskaya, J., Grossberg, S., & Mingolla, E. (2007). Laminar cortical dynamics of visual form and motion interactions during coherent object motion perception. *Spatial Vision*, *20*, 337–395.

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

Boardman, I., Grossberg, S., Myers, C., & Cohen, M. (1999). Neural dynamics of perceptual order and context effects for variable-rate speech syllables. *Perception & Psychophysics*, *6*, 1477–1500.

Born, R. T., & Tootell, R. B. H. (1992). Segregation of global and local motion processing in primate middle temporal visual area. *Nature*, *357*, 497–499.

Bowns, L. (2001). IOC, vector sum, and squaring: Three different motion effects or one? *Vision Research*, *41*, 965–972.

Brown, J., Bullock, D., & Grossberg, S. (1999). How the basal ganglia use parallel excitatory and inhibitory learning pathways to selectively respond to unexpected rewarding cues. *Journal of Neuroscience*, *19*, 10502–10511.

Brown, J. W., Bullock, D., & Grossberg, S. (2004). How laminar frontal cortex and basal ganglia circuits interact to control planned and reactive saccades. *Neural Networks*, *17*, 471–510.

Browning, A., Grossberg, S., & Mingolla, M. (2009a). A neural model of how the brain computes heading from optic flow in realistic scenes. *Cognitive Psychology*, *59*, 320–356.

Browning, A., Grossberg, S., & Mingolla, M. (2009b). Cortical dynamics of navigation and steering in natural scenes: Motion-based object segmentation, heading, and obstacle avoidance. *Neural Networks*, *22*, 1383–1398.

Buschman, T. J., & 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., & Grossberg, S. (2005). A laminar cortical model of stereopsis and 3D surface perception: Closure and da Vinci stereopsis. *Spatial Vision*, *18*, 515–578.

Cao, Y., & 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 Networks*, *26*, 75–98.

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

Carpenter, G. A., & Grossberg, S. (1987). A massively parallel architecture for a self-organizing neural pattern recognition machine. *Computer Vision, Graphics, and Image Processing*, *37*, 54–115.

Carpenter, G. A., & Grossberg, S. (1993). Normal and amnesic learning, recognition, and memory by a neural model of cortico-hippocampal interactions. *Trends in Neurosciences*, *16*, 131–137.

Carpenter, G. A., & Grossberg, S. (Eds.). (1991). *Pattern recognition by self-organizing neural networks*. Cambridge, MA: MIT Press.

Castet, E., Lorenceau, J., Shiffrar, M., & Bonnet, C. (1993). Perceived speed of moving lines depends on orientation, length, speed and luminance. *Vision Research*, *33*, 1921–1936.

Chey, J., Grossberg, S., & Mingolla, E. (1997). Neural dynamics of motion grouping: From a perture ambiguity to object speed and direction. *Journal of the Optical Society of America A*, *14*, 2570–2594.

Chey, J., Grossberg, S., & Mingolla, E. (1998). Neural dynamics of motion processing and speed discrimination. *Vision Research*, *38*, 2769–2786.

Cisek, P. (2006). Integrated neural processes for defining potential actions and deciding between them: A computational model. *Journal of Neuroscience*, *26*, 9761–9770.

Cisek, P., Puskas, G. A., & El-Murr, S. (2009). Decisions in changing conditions: The urgency-gating model. *The Journal of Neuroscience*, *16*, 11560–11571.

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

DeAngelis, G. C., Cumming, B. G., & Newsome, W. T. (1998). Cortical area MT and the perception of stereoscopic depth. *Nature*, *394*, 677–680.

Desimone, R. (1998). Visual attention mediated by biased competition in extrastriate visual cortex. *Philosophical Transactions of the Royal Society of London*, *353*, 1245–1255.

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

Elder, D., Grossberg, S., & Mingolla, E. (2009). A neural model of visually-guided steering, obstacle avoidance, and route selection. *Journal of Experimental Psychology: Human Perception & Performance*, *35*, 1501–1531.

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

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

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

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

Filehne, W. (1922). Uber das optische Wahrnehmen von Bewegunggen. *Zeitschrift fur Sinnephysiologie*, *53*, 134–145.

Francis, G., & Grossberg, S. (1996a). Cortical dynamics of boundary segmentation and reset: Persistence, afterimages, and residual traces. *Perception*, *35*, 543–567.

Francis, G., & Grossberg, S. (1996b). Cortical dynamics of form and motion integration: Persistence, apparent motion, and illusory contours. *Vision Research*, *36*, 149–173.

Francis, G., Grossberg, S., & Mingolla, E. (1994). Cortical dynamics of feature binding and reset: Control of visual persistence. *Vision Research*, *34*, 1089–1104.

Goodale, M. A., & Milner, D. (1992). Separate visual pathways for perception and action. *Trends in Neurosciences*, *15*, 10–25.

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

Grossberg, S. (1968). Some physiological and biochemical consequences of psychological postulates. *Proceedings of the National Academy of Sciences USA*, *60*, 758–765.

Grossberg, S. (1969). On the production and release of chemical transmitters and related topics in cellular control. *Journal of Theoretical Biology*, *22*, 325–364.

Grossberg, S. (1973). Contour enhancement, short term memory, and constancies in reverberating neural networks. *Studies in Applied Mathematics*, *52*, 217–257.

Grossberg, S. (1976a). Adaptive pattern classification and universal recoding: I. Parallel development and coding of neural feature detectors. *Biological Cybernetics*, *23*, 121–134.

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

Grossberg, S. (1978). A theory of human memory: Self-organization and performance of sensory-motor codes, maps, and plans. In R. Rosen & F. Snell (Eds.), *Progress in theoretical biology* (Vol. 5, pp. 233–374). New York: Academic Press.

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

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

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

Grossberg, S. (1991). Why do parallel cortical systems exist for the perception of static form and moving form? *Perception & Psychophysics*, *49*, 117–141.

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

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

Grossberg, S. (1998). How is a moving target continuously tracked behind occluding cover? In T. Watanabe (Ed.), *High level motion processing: Computational, neurobiological, and psychophysical perspectives* (pp. 3–52). Cambridge, MA: MIT Press.

Grossberg, S. (1999a). How does the cerebral cortex work? Learning, attention, and grouping by the laminar circuits of visual cortex. *Spatial Vision*, *12*, 163–187.

Grossberg, S. (1999b). The link between brain learning, attention, and consciousness. *Consciousness and Cognition*, *8*, 1–44.

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

Grossberg, S. (2003). How does the cerebral cortex work? Development, learning, attention, and 3D vision by laminar circuits of visual cortex. *Behavioral and Cognitive Neuroscience Reviews*, *2*, 47–76.

Grossberg, S. (2007). Consciousness CLEARS the mind. *Neural Networks*, *20*, 1040–1053.

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

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

Grossberg, S., (2014). How visual illusions illuminate complementary brain processes: Illusory depth from brightness and apparent motion of illusory contours. *Human Neuroscience,* doi:10.3389/fnhum.2014.00854/full.

Grossberg, S., & Howe, P. D. L. (2003). A laminar cortical model of stereopsis and three- dimensional surface perception. *Vision Research*, *43*, 801–829.

Grossberg, S., Markowitz, J., & 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 Networks*, *24*, 1036–1049.

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

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

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

Grossberg, S., & Mingolla, E. (1987). Neural dynamics of surface perception: Boundary webs, illuminants, and shape-from-shading. *Computer Vision, Graphics and Image Processing*, *37*, 116–165.

Grossberg, S., Mingolla, E., & Pack, C. (1999). A neural model of motion processing and visual navigation by cortical area MST. *Cerebral Cortex*, *9*, 878–895.

Grossberg, S., Mingolla, E., & Viswanathan, L. (2001). Neural dynamics of motion integration and segmentation within and across apertures. *Vision Research*, *41*, 2521–2553.

Grossberg, S., & Pessoa, L. (1998). Texture segregation, surface representation, and figure-ground separation. *Vision Research*, *38*, 2657–2684.

Grossberg, S., & Pilly, P. (2008). Temporal dynamics of decision-making during motion perception in the visual cortex. *Vision Research*, *48*, 1345–1373.

Grossberg, S., & Pinna, B. (2012). Neural dynamics of Gestalt principles of perceptual organization: From grouping to shape and meaning. *Gestalt Theory*, *34*, 399–482.

Grossberg, S., Roberts, K., Aguilar, M., & Bullock, D. (1997). A neural model of multimodal adaptive saccadic eye movement control by superior colliculus. *Journal of Neuroscience*, *17*, 9706–9725.

Grossberg, S., & Rudd, M. E. (1992). Cortical dynamics of visual motion perception: Short-range and long-range apparent motion. *Psychological Review*, *99*, 78–121.

Grossberg, S., & Seitz, A. (2003). Laminar development of receptive fields, maps, and columns in visual cortex: The coordinating role of the subplate. *Cerebral Cortex*, *13*, 852–863.

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

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

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

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

Grossberg, S., Yazdanbakhsh, A., Cao, Y., & Swaminathan, G. (2008). How does binocular rivalry emerge from cortical mechanisms of 3-D vision? *Vision Research*, *48*, 2232–2250.

Kellman, P. J., & Shipley, T. F. (1991). A theory of visual interpolation in object perception. *Cognitive Psychology*, *23*, 141–221.

Kelly, F., & Grossberg, S. (2000). Neural dynamics of 3-D surface perception: Figure–ground separation and lightness perception. *Perception & Psychophysics*, *62*, 1596–1618.

Lisberger, S. G., Morris, E. J., & Tychsen, L. (1987). Visual motion processing and sensory-motor integration for smooth pursuit eye movements. *Annual Review of Neuroscience*, *10*, 97–129.

Livingstone, M. S., & Hubel, S. H. (1984). Anatomy and physiology of a color system in the primate visual cortex. *Journal of Neuroscience*, *4*, 309–356.

Livingstone, M. S., & Hubel, S. H. (1988). Segregation of form, color, movement, and depth: Anatomy, physiology, and perception. *Science*, *240*, 740–749.

Lorenceau, J., & Gorea, A. (1989). "Blobs" are critical in perceiving the direction of moving plaids. *Perception*, *18*, 539.

Lorenceau, J., & Shiffrar, M. (1992). The influence of terminators on motion integration across space. *Vision Research*, *32*, 263–273.

Lueschow, A., Miller, E. K., & Desmone, R. (1994). Inferior temporal mechanisms for invariant object recognition. *Cerebral Cortex*, *4*, 523–531.

Marr, D. & Ullman, S. (1981). Directional selectivity and its use in early visual processing. *Proceedings of the Royal Society of London, B*, *211*, 151–180.

Matisse, H. (1992). *Jazz*. Scranton, PA: George Braziller. (Original work published 1947)

Mishkin, M., Ungerleider, L. G., & Macko, K. A. (1983). Object vision and spatial vision: Two cortical pathways. *Trends in Neurosciences*, *6*, 414–417.

Olson, S., & Grossberg, S. (1998). A neural network model for the development of simple and complex cell receptive fields within cortical maps of orientation and ocular dominance. *Neural Networks*, *11*, 189–208.

Pack, C. C., & Born, R. T. (2001). Temporal dynamics of a neural solution to the aperture problem in visual area MT of macaque brain. *Nature*, *409*, 1040–1042.

Pack, C. C., Grossberg, S., & Mingolla, E. (2001). A neural model of smooth pursuit control and motion perception by cortical area MST. *Journal of Cognitive Neuroscience*, *13*, 102–120.

Peterhans, E., & von der Heydt, R. (1989). Mechanisms of contour perception in monkey visual cortex: II. Contours bridging gaps. *Journal of Neuroscience*, *9*, 1749–1763.

Pinna, B. (1987). Un effetto di colorazione. In V. Majer, M. Maeran, & M. Santinello (Eds.), *Il laboratorio e la città: XXI Congresso degli Psicologi Italiani* (p. 158). Milano: Societá Italiana di Psiocologia, 1987.

Pinna, B. (2009). New Gestalt principles of perceptual organization: An extension from grouping to shape and meaning. *Gestalt Theory*, *32*, 11–78.

Pinna, B., Brelstaff, G., & Spillmann, L. (2001). Surface color from boundaries: A new "watercolor" illusion. *Vision Research*, *41*, 2669–2676.

Pinna, B., & Grossberg, S. (2005). The watercolor illusion and neon color spreading: A unified analysis of new cases and neural mechanisms. *Journal of the Optical Society of America A*, *22*, 2207–2221.

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

Ponce, C. R., Lomber, S. G., & Born, R. T. (2008). Integrating motion and depth via parallel pathways. *Nature Neuroscience*, *11*, 216–223.

Raizada, R. D. S., & Grossberg, S. (2001). Context-sensitive bindings by the laminar circuits of V1 and V2: A unified model of perceptual grouping, attention, and orientation contrast. *Visual Cognition*, *8*, 431–466.

Raizada, R. D. S., & Grossberg, S. (2003). Towards a theory of the laminar architecture of cerebral cortex: Computational clues from the visual system. *Cerebral Cortex*, *13*, 100–113.

Ramachandran, V. S. (1985). Apparent motion of subjective surfaces. *Perception*, *14*, 127–134.

Ramachandran, V. S., Rao, V. M., & Vidyasagar, T. R. (1973). Apparent motion with subjective contours. *Vision Research*, *13*, 1398–1401.

Reynolds, J. H., & Heeger, D. J. (2009). The normalization model of attention. *Neuron*, *61*, 168–185.

Roitman, J. D., & Shadlen, M. N. (2002). Response of neurons in the lateral intraparietal area during a combined visual discrimination reaction time task. *Journal of Neuroscience*, *22*, 9475–9489.

Shadlen, M. N., & Newsome, W. T. (2001). Neural basis of a perceptual decision in the parietal cortex (area LIP) of the rhesus monkey. *Journal of Neurophysiology*, *86*, 1916–1936.

Shimojo, S., Silverman, G. H., & Nakayama, K. (1989). Occlusion and the solution to the aperture problem for motion. *Vision Research*, *29*, 619–26.

Tse, P. U. (1999). Volume completion. *Cognitive Psychology*, *39*, 37–68.

Tsodyks, M. V., & Markram, H. (1997). The neural code between neocortical pyramidal neurons depends on neurotransmitter release probability. *Proceedings of the National Academy of Sciences USA, 94*, 719–723.

Usher, M., & McCllelland, J. L. (2001). The time course of perceptual choice: the leaky competing accumulator model. *Psychological Review*, *108*, 550–592.

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

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

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

von Holst, E. (1954). Relations between the central nervous system and the peripheral organs. *British Journal of Animal Behaviour*, *2*, 89–94.

Wallach, H. (1935). On the visually perceived direction of motion. *Psychologische Forschung*, *20*, 325–380.

Wuerger, S., Shapley, R., & Rubin, N. (1996). "On the visually perceived direction of motion" by Hans Wallach: 60 years later. *Perception*, *25*, 1317–1367.

Young, L. R., Forster, J. D., & van Houtte, N. (1968). A revised stochastic sampled model for eye tracking movements. Paper presented at the 4th Annual NASA-University Conference on Manual Control, University of Michigan, Ann Arbor.

Zoccolan, D., Kouh, M., Poggioa, T., & DiCarlo, J. J. (2007). Trade-off between object selectivity and tolerance in monkey inferotemporal cortex. *Journal of Neuroscience*, *27*, 12292–12307.