

CONTEXTUAL INFLUENCES ON VISUAL PROCESSING

Thomas D. Albright and Gene R. Stoner

Howard Hughes Medical Institute, Systems Neurobiology Laboratories, The Salk Institute for Biological Studies, La Jolla, California 92037; email: tom@salk.edu; gene@salk.edu

Key Words occlusion, depth-ordering, figure-ground interpretation, filling-in, visual cortex

■ **Abstract** The visual image formed on the retina represents an amalgam of visual scene properties, including the reflectances of surfaces, their relative positions, and the type of illumination. The challenge facing the visual system is to extract the “meaning” of the image by decomposing it into its environmental causes. For each local region of the image, that extraction of meaning is only possible if information from other regions is taken into account. Of particular importance is a set of image cues revealing surface occlusion and/or lighting conditions. These information-rich cues direct the perceptual interpretation of other more ambiguous image regions. This context-dependent transformation from image to perception has profound—but frequently under-appreciated—implications for neurophysiological studies of visual processing: To demonstrate that neuronal responses are correlated with perception of visual scene properties, rather than visual image features, neuronal sensitivity must be assessed in varied contexts that differentially influence perceptual interpretation. We review a number of recent studies that have used this context-based approach to explore the neuronal bases of visual scene perception.

INTRODUCTION

The *Oxford English Dictionary* defines “context” as “the parts which immediately precede or follow any particular passage or ‘text’ and determine its meaning.” More generally, context is the “whole situation, background, or environment relevant to a particular event, etc.,” which reveals its meaning. These definitions, of course, beg the question of “meaning,” to which there are no simple or all-encompassing answers. For our purposes, the definition of meaning is implicit in the central problems of vision, which are identifying the environmental causes of the patterns of light falling on the retinae, and the behavioral significance of those causes for the observer. Accordingly, the meaning of a contour of light on the retina includes, among other things, the particular object in the scene that reflected that pattern and the information that the object conveys about the world. Context, in turn, includes sensory cues that enable the image feature to be assigned to an object,

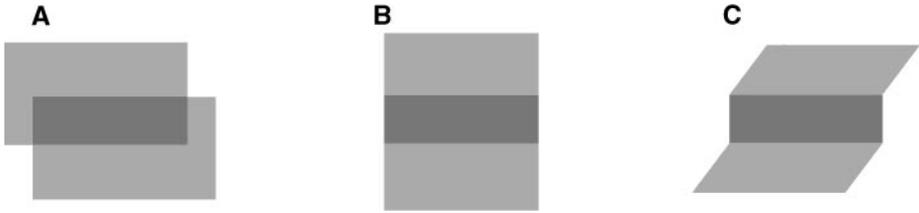


Figure 1 Illustration of the influence of local sensory context on visual perception. Each image contains a horizontal dark-gray rectangle. Although the rectangles are physically identical, the surrounding features differ in the three images. As a result, the rectangle is perceptually attributed to different environmental causes in the three instances, i.e., it conveys different meanings. (A) The rectangle appears to result from overlap of two surfaces, one of which is transparent (e.g., a piece of tinted glass). (B) The rectangle appears to result from a variation in surface reflectance (e.g., a stripe painted across a flat canvas). (C) The rectangle appears to result from variation in the angle of the surface with respect to the illumination source. These markedly different perceptual interpretations argue for the existence of different neuronal representations.

the observer's history with the object, his motivational state, and the reward value of the object to the observer. As a practical matter, extraction of meaning can be identified with perception, and context is the sensory/behavioral/cognitive milieu that influences the way each sensory feature is perceived.

Contextual influences on perception are necessarily manifested as interactions, rather than passive associations, between sensory, behavioral, or cognitive elements. For example, the percepts elicited by the images in Figures 1A–C are not merely of three different sets of features. On the contrary, the meaning of the dark gray bar is changed dramatically through its interactions with surrounding features. The perceptual whole is more than the sum of the sensory parts.

With these definitions as a foundation, we begin our review with a brief historical account of views on the role of context in vision. This history originates with philosophical doctrine and empirical psychology and is offered with the belief that past controversies, misunderstandings, and insights on this topic can guide and inform modern-day visual neuroscience. Indeed, the growing bond between psychology and neuroscience has led to remarkable new discoveries regarding the role of context in the neuronal bases of perception, as well as paradigmatic changes in the way neuronal representations are viewed and approached experimentally. These changes, and the data that support them, are the primary focus of our review.

HISTORICAL CONTEXT

The Reductionist Psychology of Perception

Discussions leading to our present-day understanding of the role of context in perception can be traced to the concept of “associationism” in the seventeenth-century

philosophy of John Locke (1690). The crux of Locke's view was that perception results from a passive association of "ideas," of which immediate sensation is the primary source. Associationism remained a dominant theme in philosophy well into the nineteenth century. Its early proponents adopted the metaphor of "mechanical compounding" to characterize their conclusion that sensations are linked noninteractively to render the objects of perceptual experience.

In the mid-nineteenth century, associationism was embraced by many of the great figures in the emerging field of experimental psychology. Wilhelm Wundt (1902), in particular, pressed further the deconstruction of psychological reality by advancing "elementism." Whereas the associationists believed percepts to be the product of linked sensations, elementism held the inverse to be true: Any percept may be reduced to independent internal states (sensations) elicited by individual elements of the proximal stimulus (such as luminance or chrominance). Wundt's arguments were highly influential and fitted squarely with the emerging physiology of his time, specifically Johannes Müller's (1838) law of specific nerve energies and the growing belief that different sensory pathways convey specific and distinct—perhaps elemental—types of information.

Despite the prevalence of the Wundtian view, there were dissenting voices. William James (1890) was an associationist to be sure, but he promoted the position that perception is an emergent product of interactions between associated sensations. The Austrian physicist Ernst Mach held similar beliefs. In *Contributions to the Analysis of Sensations* (1886), Mach discussed context explicitly and provided striking demonstrations of its role in visual perception, which anticipated the Gestalt movement in experimental psychology.

Gestalt Psychology

Flawed by its failure to account for emergent properties of perception, elementism fell from favor and Gestalt psychology rose in its place. Max Wertheimer [1961 (1912)] founded the Gestalt movement and rejected elementism following his studies of "phi motion," now commonly known as "apparent motion." This phenomenon is the illusory percept of smooth or continuous motion that results when a stimulus (a spot of light, for example) is displaced by discrete intervals over space and time. (This is, of course, the basis for motion picture photography.) Elementism argues that this motion percept should be reducible to the elemental sensations, which consist of static points of light occurring at different points in space and time. This reduction is impossible, Wertheimer observed; perceived motion under these conditions is thus an emergent property that comes about through the relations between discrete sensations—again, the whole is more than the sum of the parts.

The Gestaltists offered a set of laws to characterize the sensory interactions underlying "perceptual organization." Despite the empirical validity of these laws and the many compelling demonstrations of contextual influences on visual processing, the Gestalt school did not exert a pervasive influence on the developing field of neurobiology in the latter half of the twentieth century. There are several

reasons for this, including the heavy emphasis placed by the Gestaltists on phenomenology and their attendant failure to build firm neurobiological and mechanistic foundations for their beliefs. As a consequence, the Gestalt doctrine was rapidly eclipsed by a new reductionism, the neurophysiology of vision.

Neurophysiology of Vision: A Neurobiological Elementism?

With the development of techniques for recording neuronal action potentials and the subsequent application of these techniques to the study of the mammalian visual system, the field of visual science underwent a paradigm shift. Fundamental to the conduct and interpretation of neurophysiological experiments has been the “classical” concept of the receptive field (RF) as a discrete, spatially restricted, representational unit (Hartline & Graham 1932). Implicit in this concept is an assumption of functional independence: Once the information represented by individual RFs is known, it should be possible to deduce how neurons collectively represent the objects of perceptual experience. This assumption of independence, in turn, has fostered a localizationist view, in which different sensory elements are thought to be represented in different brain regions that are mechanistically and functionally independent of one another (e.g., Livingstone & Hubel 1987).

Conjunctively—and in spite of the many merits of the neurophysiological approach—these premises are a throwback to the nineteenth-century views of the associationists and the elementists. An association of sensory attributes represented by functionally independent neurons is reminiscent of “mechanical compounding,” and the notion that different brain regions represent different attributes smells of Wundtian elementism. Neurophysiological approaches shackled by these premises cannot reveal in toto the neural substrates of perceptual experience—for the same reasons that psychological elementism cannot explain the contextual phenomena highlighted by the Gestaltists.

As a simple illustration, consider an orientation-selective neuron in primary visual cortex (Hubel & Wiesel 1968). The RF of the neuron illustrated in Figure 2 was characterized without contextual manipulations, and the data clearly reveal how the neuron represents the proximal (retinal) stimulus elements of orientation and direction of motion. From such data, however, it is frankly impossible to know whether this neuron would respond differentially to locally identical image features viewed in different contexts, such as the dark gray bars in Figures 1A–C. In other words, the full meaning of the pattern of responses in Figure 2 is unclear, as that pattern is not sufficient to reveal what the neuron conveys about the visual scene.

The Emergence of a Neurobiology of Perception

The ready ability to record neuronal activity in alert nonhuman primates has allowed one of the most notable achievements of modern neuroscience: the establishment of direct links between the behavior of single neurons and that of the whole organism (e.g., Mountcastle et al. 1972, Newsome et al. 1989; see also Barlow 1972).

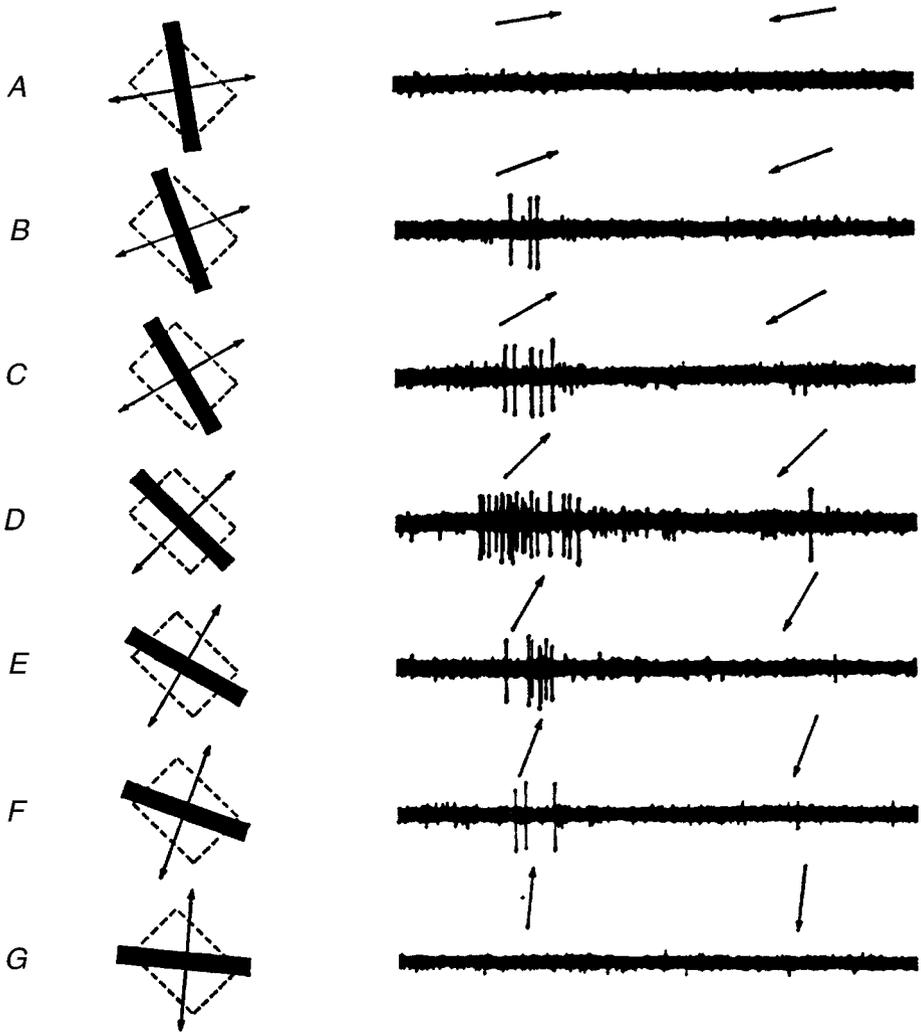


Figure 2 Neuronal directional selectivity, as first observed by Hubel & Wiesel (1968) in primary visual cortex (area V1) of rhesus monkey. The neuronal receptive field is indicated by broken rectangles in the left column. The visual stimulus was moved in each of 14 directions (rows A–G, opposing directions indicated by arrows) through receptive field. Recorded traces of cellular activity are shown at the right, in which the horizontal axis represents time (2 s/trace) and each vertical line represents an action potential. This neuron responds most strongly to motion up and to the right (row D). (From Hubel & Wiesel 1968. Permission from The Physiological Society.)

As impressive as such links are, their relevance to the extraction of “meaning” from visual images depends upon the meaningfulness of the stimuli being perceived. Unless neurophysiologists use stimuli embodying the “semantics” of natural images, they will advance—with or without the inclusion of behavioral links—only a few fledgling steps toward an understanding of the neuronal bases of perception.

To illustrate this assertion, imagine attempting to understand the function of language areas of the human brain armed only with nonlanguage stimuli such as frequency-modulated audible noise. Imagine further that, using these stimuli, you make the remarkable discovery that the responses of neurons within Wernicke’s area (for example) are correlated with behavioral judgments (of, for instance, whether the frequency modulation was high-to-low or low-to-high). Although this finding would offer an interesting (and perhaps satisfyingly parametric) link between single neurons and perceptual decisions, it seems clear that stimuli constructed of words and sentences would yield results more likely to illuminate language processing. Just as we can only progress so far using nonsense sounds to explore language function, so are we constrained by visual stimuli lacking the rich cue interdependencies that permit the perceptual interpretation of natural scenes.

In recognition of this limitation, some have advocated an approach in which neuronal activity is recorded while animals are viewing natural scenes (Gallant et al. 1998, Stanley et al. 1999, Vinje & Gallant 2000). Clearly the sensory features of such scenes are replete with contextual cues, but disentangling the roles of specific cues from among the many that exert parallel influences on neurons and perception is a formidable challenge.

Much promise lies in a simpler and more controlled method inspired by the well-documented effects of local context on visual perception—such as the influence of one image region on the surface color appearance of an adjacent region (e.g., Shevell 1978, von Helmholtz 2000 [1860/1924], Wachtler et al. 2001) or the influence of depth cues on surface brightness (e.g., Figure 1) (e.g., Gilchrist 1977, Adelson 1993, Purves et al. 1999) and perceived direction of motion (Shimojo et al. 1989, Duncan et al. 2000). By systematically varying one contextual cue at a time, neuronal-perceptual links can be established that reveal how the visual system uses these cues to recover the properties of more complex natural scenes.

Using this controlled contextual approach, several recent studies have obtained evidence for parallel influences on neuronal and perceptual sensitivity. Before considering these studies in detail, we note a development that has profoundly impacted the evolution of a context-based approach: the revamping of the concept of the RF.

Influences from Beyond the Classical Receptive Field

Interest in neuronal effects of context became focused with the discovery of modulatory influences from beyond the classical RF. There were several early indications of such effects (e.g., Barlow 1953, McIlwain 1964, Hubel & Wiesel 1965, Nelson & Frost 1978), suggesting that the classical RF concept was flawed and perhaps limited efforts to understand neural substrates of perception. By the mid-1980s, however, the accumulation of evidence for response modulation via the

“nonclassical RF,” or RF “surround,” reached a critical level, and the topic found a broad audience (Allman et al. 1985).

Perhaps the most well-documented nonclassical RF effect is the modulation, by motion in the surround, of neuronal responses in the middle temporal visual area (area MT) of primate cerebral cortex (Figure 3). MT neurons are well known to exhibit a high degree of selectivity for the direction of motion of a stimulus within the classical RF (e.g., Albright 1984). The presence of motion outside the classical RF is, by definition, incapable of eliciting a response. Nonclassical RF motion leads, however, to marked modulation of the response to a stimulus within the classical RF (Figure 4) (Allman et al. 1985, Tanaka et al. 1986, Xiao et al. 1997). Similar surround effects have been reported for other forms of feature contrast, including line orientation (e.g., Gilbert & Wiesel 1990; Knierim & van Essen 1992; Kapadia et al. 2000; Li et al. 2000, 2001) and binocular disparity (Bradley & Andersen 1998).

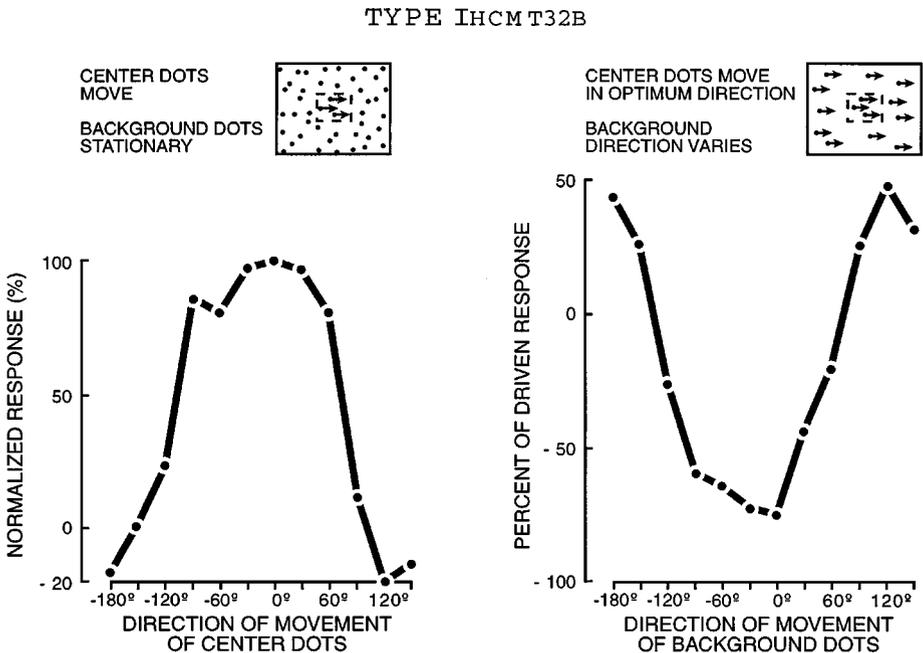


Figure 4 Responses to stimulus motion in the classical receptive field (RF) of a middle temporal (MT) neuron are modulated by motion in the RF “surround.” (*Left*) Direction tuning for a moving random dot pattern confined to the classical RF. (*Right*) Modulation of the response to classical RF motion in the preferred direction as a function of surround motion. When center and surround motion were in the same direction (0°), the response to classical RF motion was suppressed. When center and surround moved in opposite directions, the response to classical RF motion was facilitated. (From Allman et al. 1985. Permission from the Annual Review of Neuroscience, Volume 8, © 1985 by Annual Reviews.)

Although these findings suggested a neuronal substrate through which perceptual context effects might be exerted, there exists an important difference between the definition of context as it has been applied to neuronal RF surround effects and the perceptual definition we have adopted. The neuronal definition is based strictly upon an interesting spatial interaction in the RF, i.e., contextual information from the surround modulates the neuronal response to the classical RF stimulus. By contrast, the perceptual definition of context is based on the ability of contextual information to disambiguate the environmental origins of a sensory stimulus. As we show below, there are phenomena that seem to satisfy both definitions, suggesting that surround modulation may indeed underlie specific types of perceptual disambiguation. However, surround modulation is surely not the only means by which perceptual context effects are implemented neuronally. For example, because surround modulation takes on a very specific spatial form (by definition), it is difficult to imagine how it might account for perceptual effects that occur when contextual information is spatially comingled with stimulus features that are disambiguated by that context [e.g., the ability of a feature's color to disambiguate, both perceptually and neuronally, the motion of that feature (Croner & Albright 1999a; for review see Croner & Albright 1999b)].

ON THE VARIETIES OF CONTEXTUAL INFLUENCES

Broadly defined—as sources of information used to identify the meaning of a sensory stimulus—the scope of contextual influences on visual processing is vast and may include stimuli present at other points in space or time, as well as effects of attention, memory, and self-movement. For practical reasons, and in an effort to present our arguments in a coherent framework, we have chosen to concentrate on spatial contextual interactions in which information in one region of the visual image influences the interpretation of another region. Except where noted otherwise, the perceptual and neuronal phenomena we have reviewed are based on observations made in primates (humans and monkeys).

There is also a common functional theme fundamental to the context-based processes we consider: recovery of the multiple visual scene attributes associated with each single point in the visual image. The optical projection of a three-dimensional world onto a two-dimensional retina reduces the information at each spatial location to a single image value, which characterizes the light emanating from that location. Thus, although each value is a direct product of specific surface and lighting conditions in the visual scene, it is only possible to identify those conditions by taking into account the context in which the image value occurs. That context, of course, consists of other values in the image. But how is this context used to infer visual scene content? What are its critical attributes, and how do those attributes reveal surface and lighting relationships?

One way to approach these questions is to rephrase the problem as that of recovering “missing” information. Indeed, there exists a variety of circumstances, mostly intrinsic to the visual scene, but some intrinsic to the visual system itself, that cause attributes of the visual scene to be literally, fully or partially, missing from view. As an extreme example—but one of great functional significance—consider what happens when one opaque surface partially blocks the observer’s view of another surface. In such cases a single image value arising from the occluder reflects only the properties of that surface. Contextual cues from surrounding spatial regions, however, lead to perceptual decomposition or “scission” of the image value into visible and implied surfaces and to perceptual interpolation (“filling-in” or “completion”) of the “missing” attributes of the occluded surface.

A similar recovery of missing information occurs when the foreground surface is not opaque but transparent. Under these conditions each image value from the foreground surface is a physical composite of the separate reflectance and transmission properties of the overlapping surfaces, i.e., the actual properties of those surfaces are missing in the mix. Once again, surrounding contextual cues promote a scission of each image value into perceptually distinct properties of the transparent foreground and background surfaces (see Figure 1A).

A precisely analogous recovery process is engaged more generally—and ubiquitously—owing to the fact that each image value is a product of surface reflectance and illumination. Here also contextual cues promote scission: Just as context enables filling-in of surface properties missing behind an opaque or transparent occluder, context enables the viewer to see “behind” the illuminant to recover the surface reflectance properties of objects (see Figure 1C).

There are several computational steps that are broadly associated with these (and other related) context-mediated recovery processes. These include the interpretation of surface “depth-ordering,” “boundary assignment,” and filling-in. How these processes operate, what specific contextual cues are utilized, and what neuronal substrates and mechanisms are employed are the primary topics addressed in the research summarized below.

Although the visual scene information supplied by every visual image value is incomplete, this incompleteness, and hence the need for contextual resolution, is nowhere greater than when information about a part of the visual scene is absent altogether. We turn first to a discussion of how these informational gaps are filled-in using contextual cues.

FILLING-IN I: USING CONTEXT TO PREDICT MISSING INFORMATION

Filling-In the Blind Spot

Large enough to conceal 76 full moons (Campbell & Andrews 1992), the “blind spot” went unnoticed until the seventeenth century (Mariotte 1668). The blind spot is a roughly elliptical ($\sim 5^\circ$ wide by $\sim 8^\circ$ high) hole in the photoreceptor mosaic

through which information exits (via the optic nerve) and nourishment enters (via blood vessels). The trunks of blood vessels arborize atop each retina, casting shadows and creating “angioscotomas” (Le Gros 1967). The blind spot and angioscotoma form a large, continuous, irregularly shaped interruption in the central visual field of each eye (sparing the fovea), which, in collaboration with other minor scotomas and those caused by neuronal pathologies, conspire to prevent us from seeing all that lies within our visual field. That the blind spot and scotomas go unnoticed is testament to the importance of filling-in of these visual holes (Ramachandran & Gregory 1991).

A question of great importance to our understanding of neuronal representation is whether filling-in is an active process in which spatial context explicitly provides the missing information, or simply ignorance of the fact that anything is absent (e.g., von Helmholtz 2000 [1860/1924], Dennet 1991, Churchland & Ramachandran 1993, Ramachandran 1994, De Weerd et al. 1995). About the blind spot, Helmholtz concluded (contradicting many of his contemporaries), “Nothing bright or coloured or dark is to be seen in the gap in the visual field. What we see there is literally nothing—a nothing that is not even a hole . . .” (von Helmholtz 2000 [1860/1924], 3:208). Equally emphatically, he asserts “. . . and *especially* [our italics] . . . no sensations whatever are transferred from the surrounding neighborhood to the gap.”

Helmholtz’s denial of a role for context in filling-in has not prevailed. On the contrary, in recent years the field has witnessed several intriguing perceptual demonstrations of phenomenal filling-in of the blind spot with information from adjacent regions of the visual field (e.g., Ramachandran & Gregory 1991, Ramachandran 1992), as well as quantitative psychophysical support for an active process (Paradiso & Hahn 1996). For example, He & Davis (2001) showed that filled-in content from the blind spot locus of one eye could suppress real image content from the corresponding visual field locus of the other eye. These investigators reasoned that such binocular rivalry favors an active filling-in mechanism because ignored (i.e., nonactively processed) information should not be able to compete with actively processed information. The most direct evidence for active context-based filling-in comes, however, from neurophysiological studies of the blind spot representation.

The blind spots of the two retinæ do not correspond to the same parts of visual space, and hence the blind spot “representation” in primary visual cortex (V1) (see Figure 3) is a large [about 5×10 mm in humans (Tong & Engel 2001)] monocular island within a larger sea of neurons fed by both eyes. Fiorani et al. (1992) found that a bar swept across the blind spot of anesthetized Cebus monkeys (monocular viewing) elicited a response from neurons in these blind spot zones. Neurons with this “completion property” required stimulation of opposite sides of the blind spot to be activated. Similarly, Komatsu et al. (2000) found that uniform rectangles, with edges well outside the blind spot, elicited responses in the blind spot representation in alert rhesus monkeys. Together these studies indicate that perceptual filling-in of the blind spot results from an active neuronal process fed by informational content present in surrounding regions of the visual

field. Furthermore, filling-in occurs for both oriented features (such as surface boundaries) and homogenous image regions (such as surface interiors). From a functional perspective, this type of process is scarcely surprising, as local spatial context offers the best available clues to the missing content from the visual scene, and thus provides a meaningful resolution to sensory gaps.

Filling-In of Acquired or Induced Visual Gaps

Monocular retinal lesions, like the blind spot, rob information from a localized region of one eye. Whereas the visual system has had plenty of time to evolve adaptations to the visual gap created by the optic disk, the same is not true of pathological scotomas. Evidence nonetheless suggests that these visual holes are filled-in in humans (Gerrits & Timmerman 1969) and in monkeys with experimentally induced retinal lesions (Murakami et al. 1997).

Gerrits et al. (1966) asked whether filling-in extended to “artificial scotomas,” created by temporarily removing localized information from the retina. These artificial scotomas were formed from a homogenous black patch that was positioned over a restricted region of the retina and surrounded by light. When eye position was fixed relative to the patch, the patch faded over the course of a few seconds and was replaced by a percept of the surrounding light. In subsequent experiments in which the patches were surrounded by dynamic colored texture, Ramachandran & Gregory (1991) found that the color and dynamic properties of the contextual texture filled-in, but color did so faster, implicating the involvement of different mechanisms and perhaps different cortical areas (De Weerd et al. 1998).

To explore the neuronal bases of these filling-in effects, De Weerd et al. (1995) positioned similarly constructed artificial scotomas over the classical receptive fields of neurons in cortical areas V1, V2, and V3 (see Figure 3) of alert rhesus monkeys. Following stimulus onset, neuronal activity in areas V2 and V3 (but not V1) gradually increased as if responding to the missing texture. The time course of this “climbing activity” correlated well with perceptual filling-in, suggesting that these neuronal events in V2 and V3 may underlie the perceptual effect.

In a study bearing on mechanism, Pettet & Gilbert (1992) found that, when presented with artificial scotomas, V1 RFs increased in size up to fivefold. Neurons, deprived of information within their classical RF, extend their sensitivity to surrounding regions of visual space. This enlarged field could serve to fill-in the scotoma with information from the surrounding spatial field, just as sensitivity to regions surrounding the blind spot permits neurons in the blind spot representation to fill-in that visual hole. Similar changes in RF size have been seen minutes after focal retinal lesions (Chino et al. 1992, Gilbert & Wiesel 1992). Although longer-term plasticity had been documented in adult animals after restricted deafferentation (e.g., Kaas 1991, Garraghty & Kaas 1992), the results of Pettet & Gilbert’s study suggest that perceptual filling-in reflects a form of neuronal plasticity in which visual input alters RF structure and cortical topography from moment to moment. This remarkable suggestion challenges the long-standing view that a neuron’s visual RF is a stable entity within the healthy adult.

It should be pointed out that the time scale for perceptual filling-in of artificial scotomas (De Weerd et al. 1998) is not as fast as the seemingly instantaneous perceptual filling-in of the blind spot, nor is it as fast as measured for other types of filling-in (Paradiso & Hahn 1996, Paradiso & Nakayama 1991). This suggests that different neuronal mechanisms are involved in these different types of filling-in (De Weerd et al. 1998). Furthermore, the neuronal effects observed by Pettet & Gilbert (1992) occurred over a period of several minutes, not seconds, and hence the exact relationship between their observed RF changes and perceptual filling-in remains to be determined. Taken together, the different time courses observed for different types of neuronal and perceptual filling-in suggest the involvement of multiple mechanisms.

FILLING-IN II: SURFACE OCCLUSION

Intrinsic, pathological, and artificial scotomas of the sort discussed above are all characterized by loss of visual information from a fixed retinal locus—i.e., the gap moves with the eye or is present only when the eye is in a fixed position. There are, however, other circumstances of transient visual information “loss” that are caused by surface occlusion in the visual scene. These circumstances are, of course, endemic to normal visual experience, and the resolution of information loss is made possible by context in a manner that is functionally—and perhaps mechanistically—similar to the resolution of loss from scotomas.

Although we provide evidence for common principles, occlusion-based informational gaps differ importantly from those caused by defects intrinsic to the visual system. Occluders elicit a sense of depth-ordered surfaces in the visual scene. Gaps caused by the optic disk and retinal vasculature do not. Accordingly, an observer cannot usually perceive his/her own optic disk or retinal vasculature but generally will see occluding surfaces.

It is useful to distinguish two context-based processes in the resolution of information gaps caused by surface occlusion. The first involves establishing which image regions constitute “figure” and “ground” (occluder and occluded) using a set of characteristic image cues. The second process, which is intimately tied to the first, involves filling-in or perceptual completion of occluded content from the visual scene and depends upon local contextual cues that offer evidence of that missing content.

Surface Depth-Ordering, Figure-Ground Interpretation, and Border Ownership

Figure-ground interpretation is coupled to the establishment of the three-dimensional spatial arrangement of surfaces; image regions in the foreground depth plane are generally perceived as figure. There are, of course, well-known sources of depth information (as a scalar quantity) present in the visual image, such as binocular disparity and motion parallax, which serve this function. In addition, there are

image consequences that are specific to surface occlusion, which supply reliable contextual cues for surface depth-ordering (an ordinal quantity). One of the most common of such cues is due to the horizontal displacement of the two eyes, which enables one eye to see part(s) of the visual scene that a foreground surface occludes from the other eye. Human observers are particularly sensitive to this monocular occlusion cue, which is known as Da Vinci stereopsis (Anderson 1994, Nakayama & Shimojo 1990).

Occlusion produces several other types of depth-ordering cues that emerge from the distinctive characteristics of overlapping surfaces. One important cue consists of spatially defined “X-” and “T-junctions,” which are stereotyped contour arrangements in visual images that generally occur where edges of two surfaces overlap (Kanizsa 1979, Beck et al. 1984). A second cue consists of temporally defined “accretion-deletion,” which results from the progressive uncovering and/or covering of a background surface by a foreground surface (Kaplan 1969, Stoner & Albright 1995). A third cue consists of the coincident alignment of unconnected line endings or edges, which is indicative of occlusion by a common surface (Kanizsa 1979). The presence of any of these contextual cues generally biases perceived depth-ordering and figure-ground interpretation in human observers.

The Gestalt psychologists emphasized a number of additional image cues that lead to figure-ground interpretation but do not offer explicit evidence of occlusion. For example, regions within a closed boundary usually appear as figure (Koffka 1935). Furthermore, if the boundary resembles the profile of a known object, the image region corresponding to the implied object will tend to be seen as figure (Rubin 1921, Peterson 2002).

In some of the earliest neurophysiological studies of this topic, Lamme (1995) and Zipser et al. (1996) searched for correlates of figure-ground interpretation in area V1. Stimuli consisting of texture-defined rectangles (perceptual figures defined by boundary closure) were positioned such that individual texture elements drawn from figure or ground were within the classical RF. V1 responses to figure elements were reported to be significantly larger than responses to ground elements. This result implies that figure versus ground is encoded by differential response rate, and that larger spatial context elicits this differential. Using stimuli similar to those of Lamme (1995), however, a more recent study (Rossi et al. 2001) failed to find any response differential in V1 for figure versus ground elements in the classical RF (see also Cumming & Parker 1998). On the other hand, using stimuli consisting of two overlapping rectangles, which possess robust occlusion-based depth-ordering cues (T-junctions), Chang et al. (1999, 2001) detected small populations of both V1 and V2 neurons that responded differentially to the interiors of figure versus ground rectangles. Although differential responses to figure versus ground interiors could constitute the neuronal substrate for the heightened salience of figure versus ground (e.g., Kanizsa 1979), further evidence is needed to resolve the discrepancies between these different studies.

The attribution of figures to one side or the other of an image boundary is commonly referred to as assigning “border ownership” (Koffka 1935, Rubin 1921,

Zhou et al. 2000). Border ownership identifies which side of the boundary is seen as the occluder and which side as the occluded background. Several recent studies have found evidence of neuronal correlates of this perceptual process at early stages within the visual pathway. Zhou et al. (2000) stimulated V1, V2, and V4 (see Figure 3) neurons using stimuli that consisted of single rectangles (figure defined by closure) and overlapping pairs of rectangles (figure defined by occlusion-based depth-ordering cues, i.e., T-junctions). These investigators placed the border between figure and ground within the classical RF and varied border ownership, such that the figure lay on one side of the RF or the other. Slightly more than half of the recorded neurons in areas V2 and V4 exhibited significantly different responses for contrast edges as a function of border ownership (Figure 5). A smaller fraction of V1 neurons showed similar effects. The spatial integration of these contextual cues was found to extend to at least 20°. Moreover, they emerged almost immediately following stimulus onset, suggesting a rapid conveyance of contextual information from the nonclassical RF. Zhou et al. also found that different cues for border ownership (closure, T-junctions, relative size) yielded similar effects, which indicates that these neurons were encoding border ownership *per se*, rather than a particular cue for border ownership.

Chang et al. (2001) used a different strategy to differentiate selectivity for border ownership from sensitivity to incidental aspects of stimulus geometry. Specifically, stimuli with occlusion cues (T-junctions) yielding ambiguous figure-ground interpretation were compared with those with unambiguous interpretation. The unambiguous stimuli were overlapping rectangles similar to those of Zhou et al. This enlarged stimulus set allowed identification of neurons responding merely to the presence of T-junctions, rather than to depth-ordering *per se*. With the incorporation of these experimental controls, Chang et al. (2001) confirmed that responses correlated with border ownership existed within V2, although these effects were modest in magnitude and prevalence and were rarely seen in V1.

Two additional studies sought neuronal correlates of border ownership using stimuli with occlusion cues local to the stimulating border and thus present in the classical RF. Baumann et al. (1997) stimulated neurons in areas V2 and V3/V3A with spatially defined occlusion cues and observed responses that were consistent with border ownership. Stoner et al. (1998) stimulated neurons in area MT using temporally defined occlusion cues (i.e., accretion-deletion) and found evidence for figure-ground selectivity. Much more work is needed to establish the precise role(s) of these different cortical areas in figure-ground interpretation. The existing neurophysiological evidence suggests, nonetheless, that neuronal representations of figure-ground relationships are pervasive and are manifested early in the visual pathway.

Completion of Occluded Information

As we have seen, figure-ground interpretation follows the occlusion of one surface by another. Figure-ground interpretation itself leads to another important process: the filling-in of occluded features. There are two types of filling-in associated

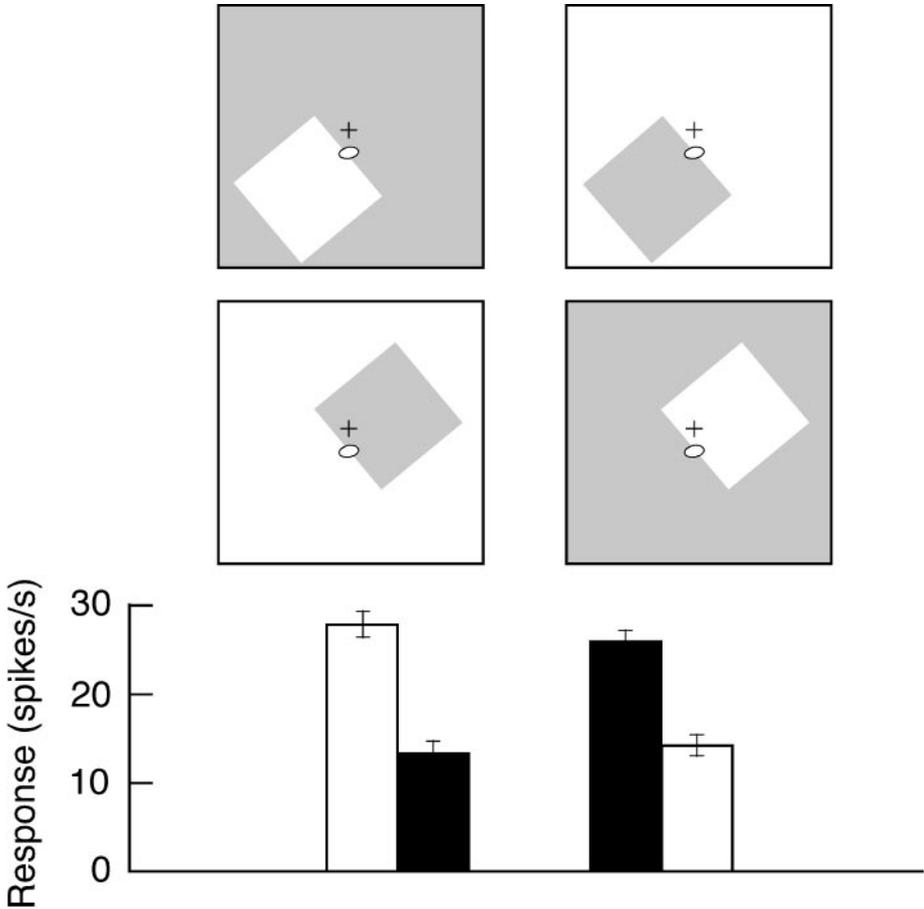


Figure 5 Border ownership responses from a V2 neuron. Responses to white squares on gray background are indicated by open bars. Responses to gray squares on white background are indicated by filled bars. Stimuli vary in luminance contrast polarity (i.e., white vs. gray or gray vs. white) across columns and implied border ownership across rows. This cell preferred the border ownership implied by the top row of stimuli in which figures were on the bottom left side of the receptive field (RF). Ellipses indicate classical RF (note that the preferred orientation is that of the short axis of the ellipse). This cell differentiated displays that were identical in an $8 \times 16^\circ$ region around the RF. (From Zhou et al. 2000. Copyright 2000 by the Society for Neuroscience.)

with occlusion, termed “modal” and “amodal” completion (Michotte et al. 1964). Modal completion refers to illusory image features that appear to have resulted from direct stimulation of the visual modality (Kanizsa 1979). The classic Kanizsa triangle (Figure 6) elicits illusory contours, which comprise a case of modal completion driven by contextual cues for surface occlusion. As illustrated in the

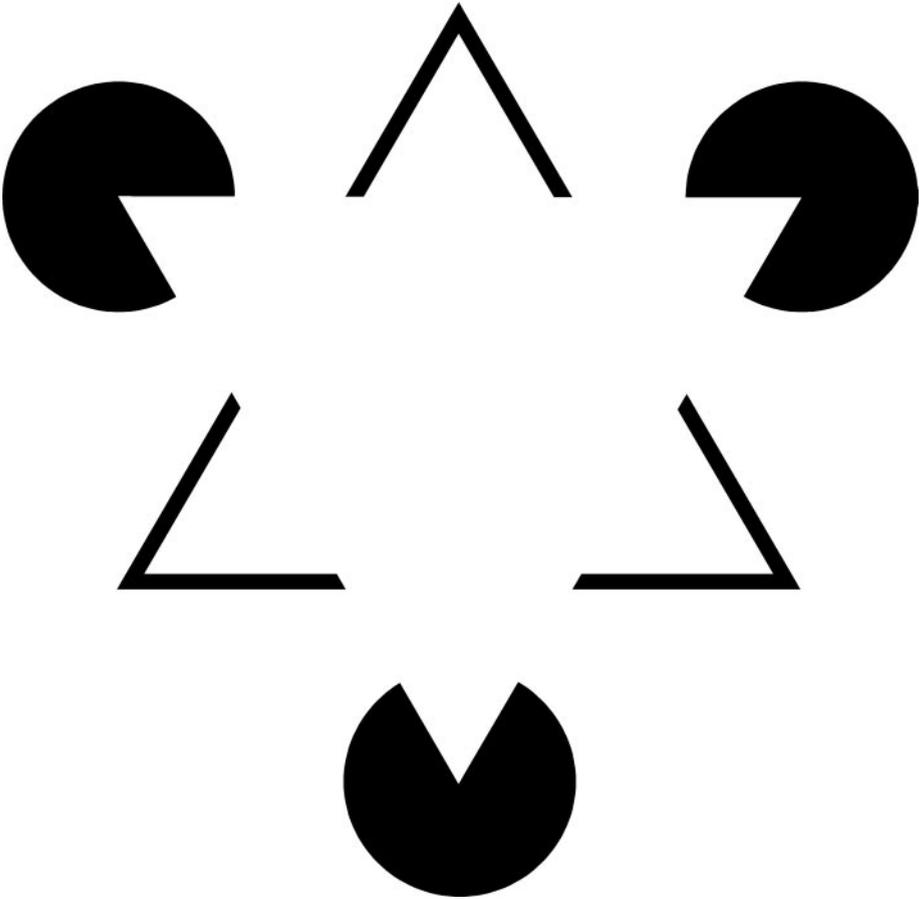


Figure 6 The Kanizsa triangle illustrates two forms of filling-in—modal and amodal completion—elicited by contextual cues for surface occlusion. Modal completion: Coincident alignment of contrast edges and line endings allows for perceptual interpolation of a foreground occluding triangle. Although the inferred foreground occluder and background are physically indistinct, contextual cues enable this “missing” boundary information to be filled-in. The resulting illusory contours have a sensory-like quality or “modal presence.” Amodal completion: The existence of a triangular foreground occluder is implied, in part, by the V-shaped segments missing from each of three black features. The occluder implies, in turn, that the black features are actually partially occluded full circles. Similarly, the three black-outlined Vs along the edges of the foreground occluding surface appear to compose a second, partially occluded, triangle. Although observers do not perceive the filled-in occluded features of these objects with the verisimilitude of their unoccluded portions, they are nonetheless fully and directly aware of the presence of occluded features. Moreover, the filled-in occluded features possess very specific forms, which are largely dictated by contextual cues, and observers can readily direct actions toward filled-in features, suggesting that they are explicitly represented in the brain. (From Kanizsa 1979.)

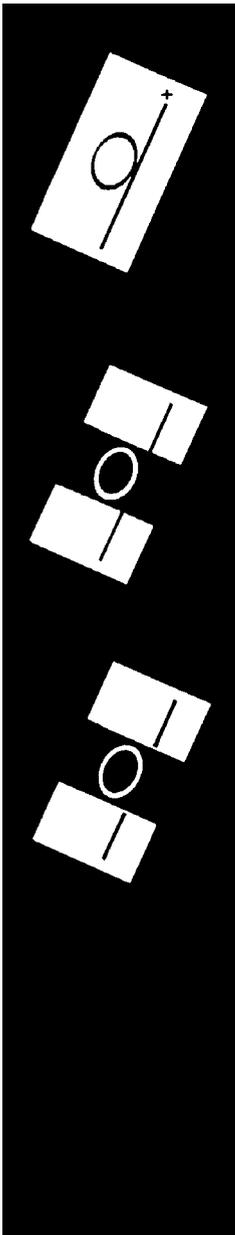
Kanizsa triangle, modal completion usually requires that foreground and occluded background have the same apparent intensity—a relatively rare occurrence in natural scenes. [The illusory “phi motion” discovered by Wertheimer [1961 (1912)] and perceptual filling-in of the blind spot/scotomas (all reviewed above) are also striking examples of modal completion, though unrelated to surface occlusion.]

In seminal experiments von der Heydt et al. (1984) explored neuronal correlates of modal completion. Stimuli consisted of illusory contours, which were induced by contextual cues for the presence of a foreground occluding surface (analogous to the case illustrated in Figure 6). The illusory contours were moved through the classical RFs of V1 and V2 neurons by moving the inducing context through the RF surround. Remarkably, when contextual cues were configured to elicit a percept of illusory contours, they also elicited neuronal responses in V2 (but rarely V1) that were comparable to responses elicited by real contours (Figure 7)—despite the fact that the illusory contours presented no physical contrast within the classical RF! Redies et al. (1986) obtained similar results from cat cortical areas 17 and 18.

Although these data support a model in which modal completion evolves in V2 from real contour signals present in V1, the V1-V2 distinction made by von der Heydt and colleagues has been challenged. Grosf et al. (1993) and Lee & Nguyen (2001) found that a significant fraction of V1 neurons responded to illusory contours (see also Sheth et al. 1996, Ramsden et al. 2001). In addition, the neuronal blind spot filling-in effects reported in V1 by Komatsu et al. (2000) (see above) suggest the existence of a mechanism for contour completion in V1. Lee & Nguyen reported, however, that V1 responses to illusory contours lag behind responses to real contours by about 55 ms, whereas this difference is only ~30 ms in area V2. They suggest that this lag could reflect feedback latencies from area V2 or from horizontal connections intrinsic to V1. As we shall see, the debate about the contribution of V1 to modal contour completion is further complicated by recent evidence that V1 neurons represent amodal contours.

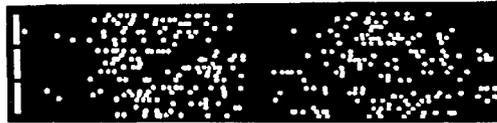
Amodal completion, which is the second type of filling-in, refers to perceptual interpolation of features seen to lie behind a foreground occluding surface [termed amodal, according to Kanizsa (1979), because the completed information “is not verified by any sensory modality”]. Matching the ubiquity of occlusion in the visual environment, amodal completion is much more common, and hence arguably of greater functional significance than the modal variety.

The specific features that are amodally completed are determined by contextual cues surrounding the occluding surface. Given that we do not actually “see” amodal features (unlike modally completed features), it might be thought that they are not explicitly represented. Counter to this supposition, however, are the observations that amodal features have a specific form, surface character, and extent, and are not easily modifiable by cognitive intent (e.g., try imagining the occluded triangle contours in Figure 6 as anything other than straight lines). Moreover, an observer can easily direct a motor action to amodal features (as one might reach behind a book to grab the edge of a desk it partially occludes). Quantitative psychophysical evidence supports these anecdotal phenomena (e.g., Nakayama & Shimojo 1990)—all of which argues that amodal features are explicitly represented by the brain.



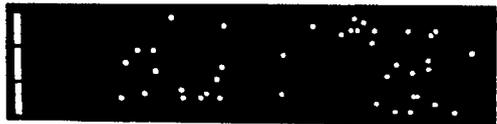
23.1

A



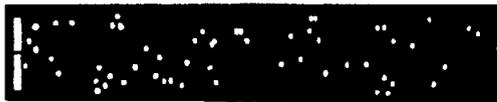
13.9

B



1.9

C



3.9

D

Even if explicitly represented, the phenomenally “nonvisual” character of amodal features would seem to place them in the domain of higher-order visual areas. This appears not to be the case. Sugita (1999) reported neuronal correlates of amodal completion within area V1. When presented with two collinear line segments placed on opposite sides of the classical RF, over 10% of V1 neurons responded if the gap between the segments was placed stereoscopically near (hence consistent with occlusion) (Figure 8). By contrast, these same neurons failed to respond if the gap was positioned stereoscopically behind the line segments, such that amodal completion was prohibited. Sugita found that response latencies to amodal contours were similar to those elicited by real contours presented within the classical RF. On this basis, Sugita concluded that these responses are likely to be mediated by horizontal connections within V1 or perhaps by short-latency feedback connections from V2.

The sensitivity of V1 to amodal features implies neuronal “X-ray vision”; with the help of contextual cues, area V1 constructs a representation of what lies hidden behind occluders. The apparent qualitative similarity between neuronal representations of amodal, modal, and real contours seems to conflict, however, with their differing perceptual qualities. On the other hand, as we have noted, amodally completed features are tangible enough to expect some processing overlap with real features. The degrees of processing differences and overlap are sure to emerge as research begins to focus on the cellular mechanisms by which context gives rise to these forms of filling-in.

INTERRUPTED CONTOURS: FACILITATION AND GROUPING

As we have seen, contours can be interrupted in the visual image by a variety of means, including surface occlusion and shadows. Depending upon the visual configuration, contours may be perceptually completed across these gaps (modally or amodally). Two related behavioral paradigms have also been used to explore

← **Figure 7** Responses of a V2 neuron to modal (illusory) contours. Stimuli are shown at left. Ellipses indicate the classical receptive field (RF). Crosses indicate the position of the monkey’s fixation point. Each row of dots in the right column corresponds to a forward and backward sweep across the neuronal RF. Each dot represents an action potential. (A) Responses to a real contour. (B) Responses to a modally completed (illusory) contour were qualitatively similar, even though no physical contrast was present in its classical RF. (C) Placing thin intersection lines between black contours and intervening black patch prevented modal completion. Consistent with perception, no response was observed. (D) Spontaneous activity. Numbers below each panel of spike trains indicate mean spike counts per stimulus cycle. (From Peterhans & von der Heydt 1989. Copyright 1989 by the Society for Neuroscience.)

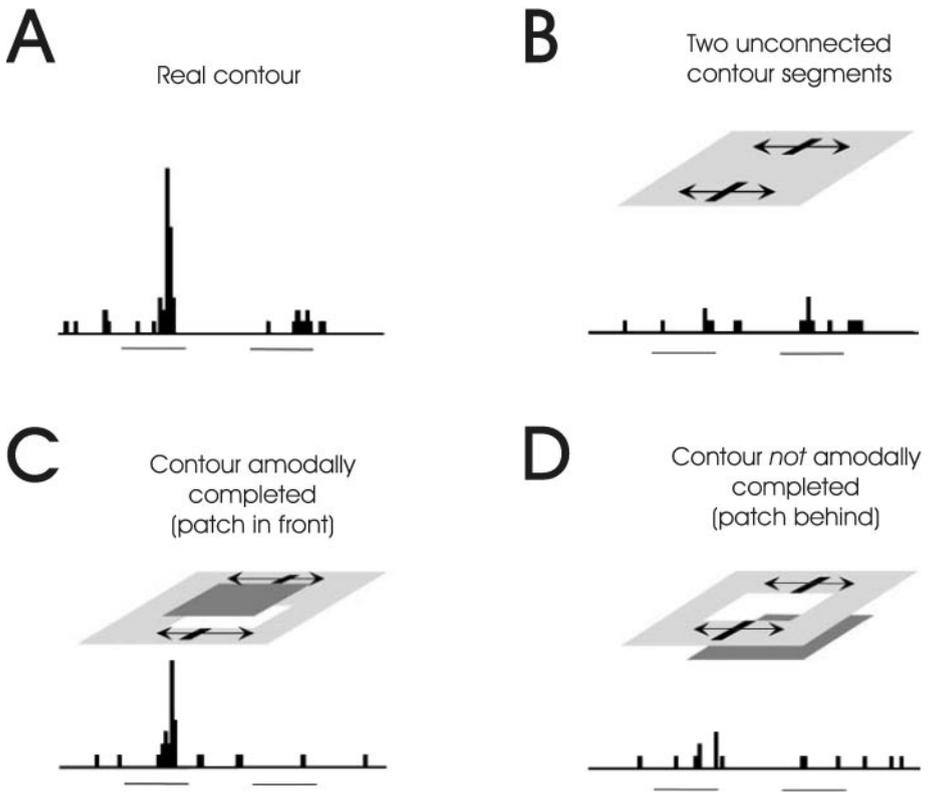


Figure 8 Responses of a V1 neuron to amodal contours. (A) Responses to a real contour. (B) Responses to two unconnected line segments (no stimulus in classical receptive field). (C) Responses to two collinear segments with a patch placed stereoscopically in front. (D) Responses to two collinear segments with a patch placed stereoscopically behind. Each stimulus type was visible during periods indicated by bars under each histogram and was swept in each of two directions (left bar, preferred direction; right bar, not preferred direction). This neuron exhibited directionally selective responses to both real (A) and amodally completed (C) contours. Little response was observed for unconnected segments in which amodal completion was not engaged by the presence of a foreground occluding surface (B and D). (From Sugita 1999.)

contextual interactions between contour fragments. The first has been used to show that collinear contour fragments tend to be perceptually “grouped” (Field et al. 1993). The second paradigm of this genre has been used to demonstrate that the processing of oriented contour fragments can be facilitated by flanking collinear contours (Polat & Sagi 1993, 1994; Kapadia et al. 1995).

Neurophysiological experiments that aim to identify neuronal correlates of grouping and contour facilitation require assessment of the degree to which the response to an oriented RF stimulus is modulated by contours in the RF surround. Many studies have documented surround effects in area V1 using oriented stimuli

(e.g., Hubel & Wiesel 1965; Bishop et al. 1973; Maffei & Fiorentini 1976; Nelson & Frost 1985; Gulyas et al. 1987; Gilbert & Wiesel 1990; Knierim & van Essen 1992; Li & Li 1994; Kapadia et al. 1995, 2000; Sillito et al. 1995; Polat et al. 1998; Li et al. 2000, 2001). A subset of these have manipulated contour collinearity and have yielded evidence of response facilitation when discontinuous contours are placed collinearly along the long axis of a RF (Maffei & Fiorentini 1976, Nelson & Frost 1985, Kapadia et al. 1995). Most notably, Kapadia et al. (1995, 1999) found that the responses of V1 and V2 neurons to a suprathreshold line within the classical RF can be greatly increased by adding a flanking collinear line in the RF surround.

Contour response facilitation could represent a neuronal correlate of perceptual continuity in the face of contour fragmentation caused by occlusion. In an elegant examination of this hypothesis, Bakin et al. (2000) obtained evidence suggesting that the neuronal collinear facilitation effect is associated with amodal completion (see Figure 9). To control the degree of amodal completion, these investigators placed additional lines orthogonally between the RF stimulus and flanking stimuli, and manipulated the relative depth of these lines using binocular disparity cues. The effects of these manipulations on neuronal response facilitation appeared strikingly consistent with amodal completion: Facilitation was blocked by placing orthogonal bars in a depth plane behind the collinear fragments, in a manner suggesting that the fragments lie on separate foreground surfaces (no amodal completion). Facilitation recovered when the orthogonal bars were placed in the near depth plane, in a manner that allows the fragments to lie on the same background surface (amodal completion).

This differential effect of orthogonal bar depth was seen by Bakin et al. (2000) in half of the V2 neurons so tested but was rarely seen in area V1. Although interesting, these results are somewhat at odds with those of Sugita (1999). Unlike Bakin et al., Sugita found neuronal correlates of amodal completion in V1. This discrepancy may simply be due to different sampling of the visual field (Sugita's sample was limited to the central 2°, whereas that of Bakin et al. extended 3–6°). Somewhat more puzzling, perhaps, is the fact that Bakin et al. observed neuronal contour facilitation under one condition in which neuronal (and indeed perceptual) amodal completion does not appear to occur (Sugita 1999): two interrupted fragments with no intervening surface (compare the stimulus configuration of Figure 8B with that of 9C). This observation casts doubt on the conclusion that contour facilitation depends upon amodal completion. This intriguing incongruity might best be resolved by examining neuronal correlates of amodal completion and flank facilitation in tandem while simultaneously monitoring perceptual interpretation.

These neuronal findings of Gilbert and colleagues (Bakin et al. 2000; Kapadia et al. 1995, 1999) also appear to have a connection to the behavioral phenomena of contrast facilitation (Dresp 1993; Polat & Sagi 1993, 1994; Zenger & Sagi 1996) and contour grouping (Field et al. 1993). Plausible candidates for mediation of these varied contour phenomena are long-range horizontal connections within area V1 (Rockland & Lund 1982, Gilbert & Wiesel 1983, Martin & Whitteridge 1984). These connections are anisotropic and preferentially connect neurons that have collinear RFs (Bosking et al. 1997, Schmidt et al. 1997).

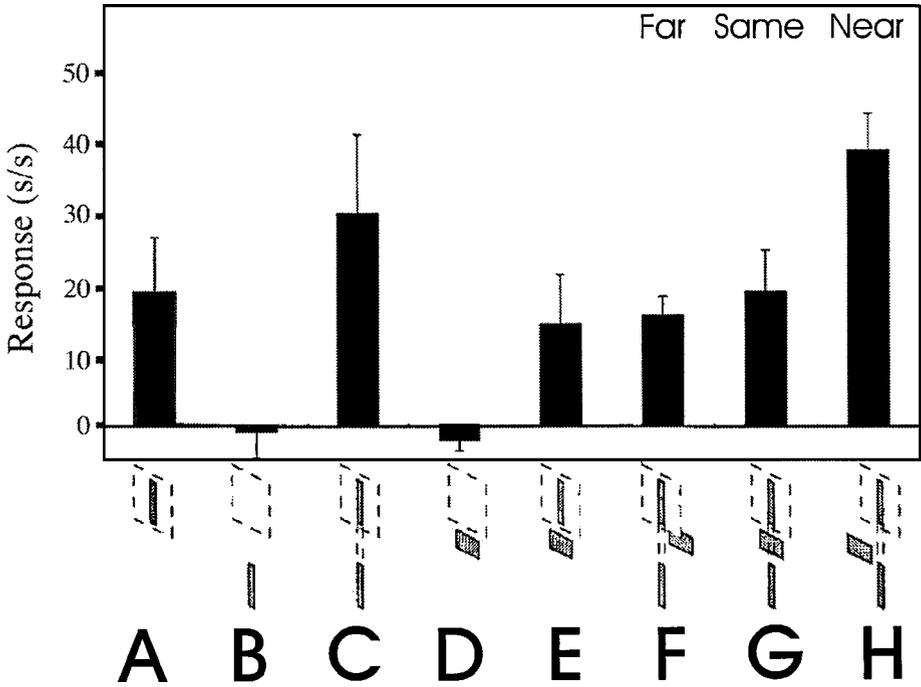


Figure 9 Collinear contextual response facilitation in area V2. Collinear contours in non-classical receptive field (RF) facilitate responses to RF contour, but not if contextual conditions are inconsistent with amodal completion (but see text). Flanking contour facilitates responses (compare *A* with *C*). An orthogonal bar placed in either the same plane (*G*) as the flank and target stimulus or in the far depth plane (*F*; 0.16° uncrossed disparity) blocked the flank-induced facilitation of the response to the target stimulus. Conversely, placing the orthogonal bar in the near depth plane (*H*; 0.16° crossed disparity) did not block this facilitation. (From Bakin et al. 2000. Copyright 2000 by the Society for Neuroscience.)

We have discussed evidence of neuronal filling-in/completion of oriented features and surface properties. We next discuss how completion of moving surfaces might manifest itself as a contextual effect arising from the nonclassical RF of motion-detecting neurons.

OCCCLUSION AND VISUAL MOTION PROCESSING

Although occlusion presents a ubiquitous challenge in the recovery of visual scene properties, it poses a special challenge to the recovery of visual motion. Where occluded, the edge of a moving surface appears terminated in the visual image, yielding a feature that traces an unambiguous trajectory across the image. Because such features and their trajectories are defined, however, not by the moving object, but by the object's coincidental intersection with a foreground occluder (Figure 10),

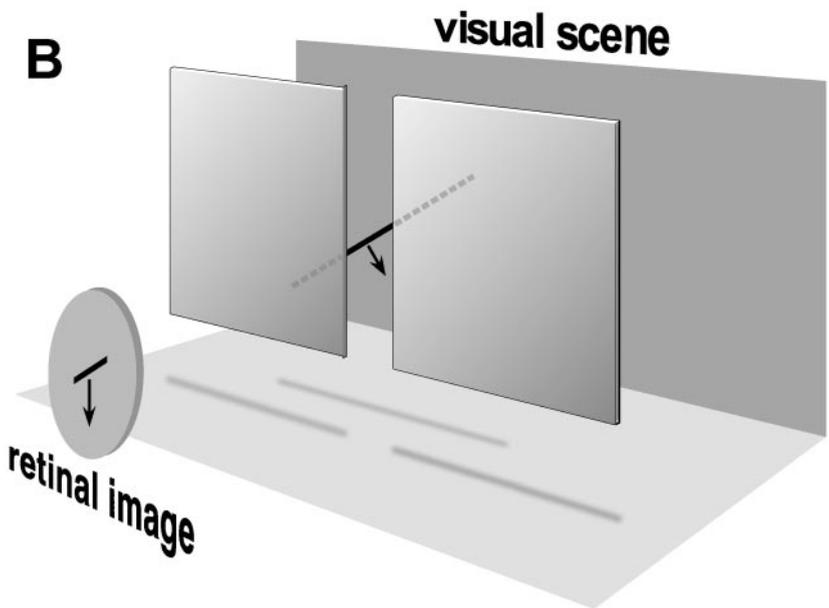
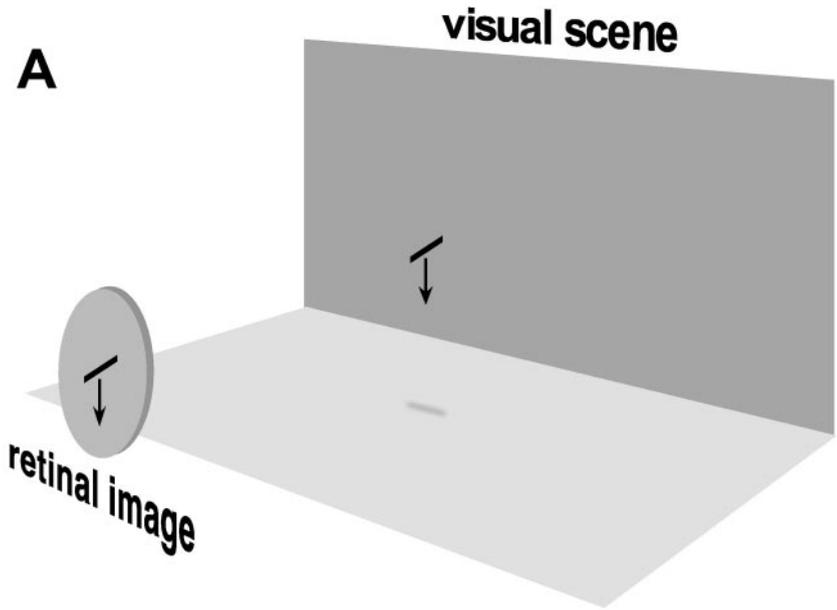
they provide spurious information about the motion of the object in the scene. For that reason, the visual motion system must distinguish the false motion information provided by features “extrinsic” to a moving object from the valid information provided by features “intrinsic” to that object (Nakayama et al. 1989). Shimojo et al. (1989) demonstrated that human subjects utilize depth-ordering cues to distinguish intrinsic and extrinsic features, which in turn enables recovery of the motions of partially occluded objects.

Duncan et al. (2000) devised a variant of the classic barber-pole illusion (Wallach 1935) to explore the neuronal substrates of contextual depth-motion interactions. Their “barber-diamond” stimuli consisted of a moving grating framed by a static diamond-shaped aperture (Figure 11). Two of four textured panels that defined the aperture were placed in front of the grating via stereoscopic depth cues, and the other two were placed behind. These depth manipulations simulated partial occlusion of the grating. The features formed by termination (“terminators”) of the grating stripes at the far panels were seen as intrinsic, and those formed at the near panels were seen as extrinsic to the grating. When the grating was moved horizontally, human observers reported that the grating moved diagonally upwards or downwards, in the direction of the intrinsic terminators. The depth cues at the margins of the moving stimulus thus provided a context for perceptual disambiguation of motion within the interior. [Similar perceptual motion effects can be obtained using simulated shadows, rather than binocular disparity, to elicit depth-ordering (G.R. Stoner, unpublished observations).]

Duncan et al. (2000) compared these perceptual effects to the responses of directionally selective neurons in cortical area MT. Barber-diamond stimuli were positioned such that the aperture edges (and hence depth-ordering cues) were outside of the classical RF. Many MT neurons exhibited directionally selective responses consistent with perceived direction of motion, i.e., depth-ordering cues that were restricted to the RF surround predictably altered responses to moving features within the classical RF (Figure 12).

These findings challenge previous characterizations of nonclassical receptive field (RF) surround effects within area MT, which have stressed relatively simple feature contrast effects. As noted above, the responses of many MT neurons to motion in the classical RF are modulated by motion in the surround (Figure 4) (Allman et al. 1985). A similar modulation has been reported for binocular disparity (Bradley & Andersen 1998). These two types of “intra-modal” (motion-motion and depth-depth) interactions may be helpful in signaling cue-specific image discontinuities (Bradley & Andersen 1998) or extracting depth variation based on either motion parallax or binocular disparity (Buracas & Albright 1996, Liu & Kersten 1998, Xiao et al. 1997). Bradley & Andersen (1998) found few interactions between depth and motion, however, which reinforced a view of surrounds as exerting intra-modal and mostly antagonistic modulation of responses to stimulation of the classical RF.

In the Bradley & Andersen study, depth and motion were independent stimulus attributes, neither bearing on the interpretation of the other. By contrast, in the barber-diamond stimuli employed by Duncan et al. (2000) (as in natural scenes)



depth cues provide a context that triggers recovery of occluded scene elements, which in turn modulates the representation of motion. Thus, as was the case for many of the contextual effects reviewed herein, it was the incorporation of occlusion cues that uncovered the sophisticated RF surround interactions observed by Duncan et al.

Although the precise mechanisms responsible for the depth-motion interactions of Duncan et al. are unknown, there are some plausible alternatives. One possibility is that the occlusion cues elicit a neuronal representation of the amodally completed grating behind the near panels, as suggested by the findings of Sugita (1999) (reviewed above). These completed contours might then lead to the recruitment of additional motion detectors, which may in turn shift the balance in favor of motion in the direction of intrinsic terminators. A second possible mechanism involves selective pooling of motion signals determined to be intrinsic to the same surface (Stoner & Albright 1993). For example, pooling of the signals arising from the horizontally moving gratings with the diagonally oriented far edge (intrinsic to the grating) would yield the observed diagonal motion of the barber-diamond. A third mechanism involves selective suppression of motion signals arising from extrinsic features (Liden & Pack 1999), which would tilt the balance in favor of intrinsic terminator motions. These three potential mechanisms are, of course, not mutually exclusive.

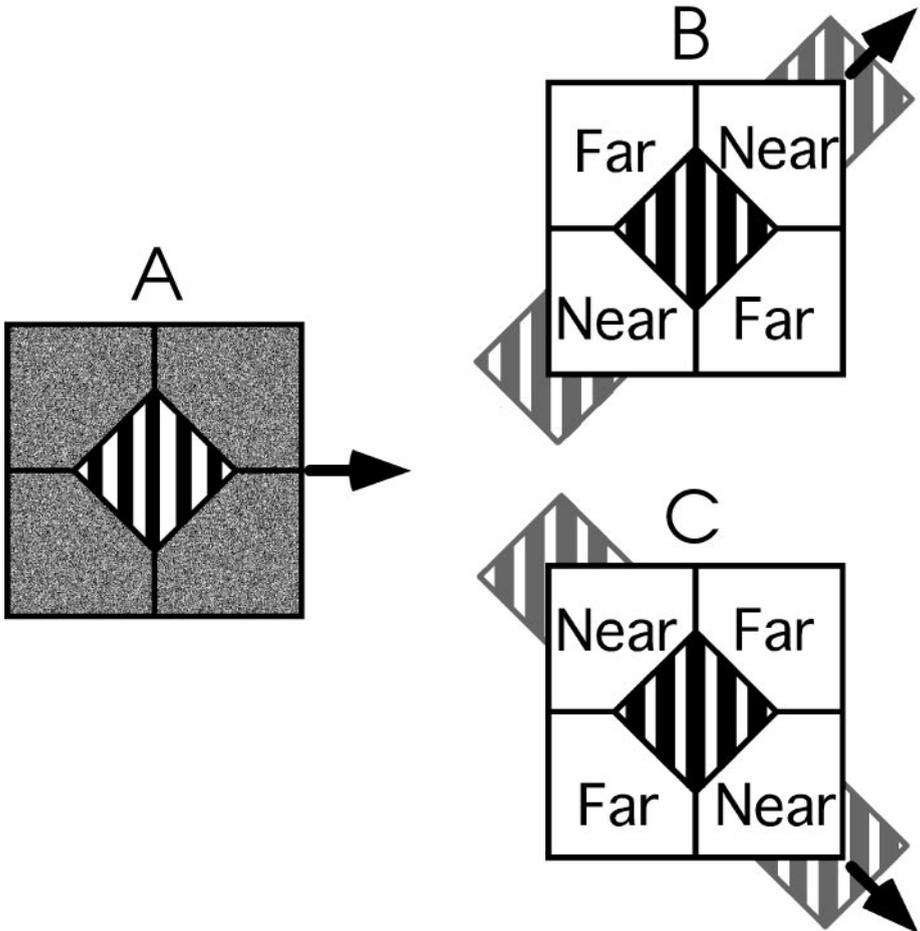
FILLING-IN REVISITED: NORMAL SURFACE PERCEPTION

At any one time, only a small proportion of cells (in visual cortex) are likely to be influenced (activated or suppressed), since contours of inappropriate orientation and diffuse light (within the receptive field) have little or no effect on a cell (Quotation from Hubel & Wiesel 1968).

← **Figure 10** Illustration of the contextual dependence of motion perception. Two simple visual scenes are shown, along with image motions they give rise to. (a) A single moving object oriented obliquely and moving downward. Image motion reflects object motion. (b) A single obliquely oriented bar moving to lower right (orthogonal to orientation). The scene also contains occluders, which block the observer's view of the ends of the bar. Resulting visual image motion in (b) is identical to that caused by the scene in (a), although the visual scene motions are clearly different. Observers perceive the different object motions correctly, and they do so because of different spatial contexts, which elicit different interpretations of the terminations of the moving line in the visual image. Terminators in (a) are interpreted as ends of an elongated object, i.e., they are intrinsic features of that object and their downward motions genuinely reflect object motion. Terminators in (b), by contrast, are extrinsic to the moving line (an accident of occlusion), and their downward motions are deemed spurious. Thus, terminator interpretation determines how the image motions are perceived.

Hubel & Wiesel's observation implies that visual neurons only respond to image contrast and that every homogenous region in the visual image is, in effect, a blind spot. Surfaces with uniform reflectance are common and, hence, so are uniform image regions. This poses a profound computational problem and a mystery: How is it that we are not blind to much of the world?

Part of the answer might be thought to lie in the functions of "luxotonic" neurons, which do respond to diffuse sustained light. However, the importance of these neurons is called into question by their relative rarity in the species in which they have been detected (DeYoe & Bartlett 1980, Kayama et al. 1979) and their absence in other species (Kahrilas et al. 1980). Moreover, as others have noted (Walls 1954, Gerrits & Vendrik 1970), the computational problem posed



by uniform regions in the visual image has an obvious parallel with that posed by the blind spot and scotomas, and the same mechanistic considerations and potential neuronal solutions apply. In particular, one can ask whether filling-in of homogenous image regions is an active or passive process.

As we have seen, perceptual completion of image gaps caused by the blind spot and scotomas, as well as by surface occlusion, involves an active neuronal process in which the missing features are explicitly reconstructed. It is possible, nonetheless, that our ability to perceive areas of constant coloration is accomplished implicitly (i.e., passively). For example, imagine a gray surface atop a white background. Neurons encoding one spatial location detect one surface border (light-dark); neurons encoding another location detect the complementary discontinuity (dark-light). According to an implicit encoding scheme, border-detecting neurons signal that the surface is darker than the background, while uniformity of

←

Figure 11 Schematic depiction of “barber-diamond” stimuli used to study the influence of context on perceived motion and its neuronal correlates (a demonstration can be seen at <http://www.cnl.salk.edu/~gene/>). (a) Stimuli consisted of a moving square-wave grating framed by a static, diamond-shaped aperture. The grating itself was placed in the plane of ocular fixation while four textured panels that defined the aperture were independently positioned in depth using binocular disparity cues. The grating was moved either leftward or rightward on each trial. Each of the four stimulus conditions used was a unique conjunction of direction of grating motion (right vs. left) and depth-ordering configuration. The two conditions illustrated (b and c) were created using rightward-moving gratings; two additional conditions (not shown) were created using leftward-moving gratings. “Near” and “Far” identify the depth-ordering of the textured panels relative to the plane of the grating. (b) Upper-right and lower-left panels were placed in the near depth plane, while the upper-left and lower-right panels were in the far depth plane. Line terminators formed at the boundary of near surfaces and grating are classified as extrinsic features resulting from occlusion, and the grating is amodally completed behind the near surface. (Note that gray stripes are not part of the stimulus and are used solely to illustrate amodal completion.) Conversely, line terminators formed at the boundary of the far surfaces and the grating are classified as intrinsic, i.e., they appear to result from the physical termination of the surface upon which the stripes are “painted.” As a result of this depth-ordering manipulation and ensuing feature interpretation, observers typically perceive the moving grating in (b) as belonging to a surface that slides behind the near panels and across the far panels, i.e., to the upper-right. This direction is identified with motions of intrinsic terminators. The condition shown in (c) contains the same rightward grating motion but employs the depth-ordering configuration that is complementary to (b). In this case observers typically perceive the grating belonging to a surface that is drifting to the lower right. (After Duncan et al. 2000. Copyright 2000 by the Society for Neuroscience.)

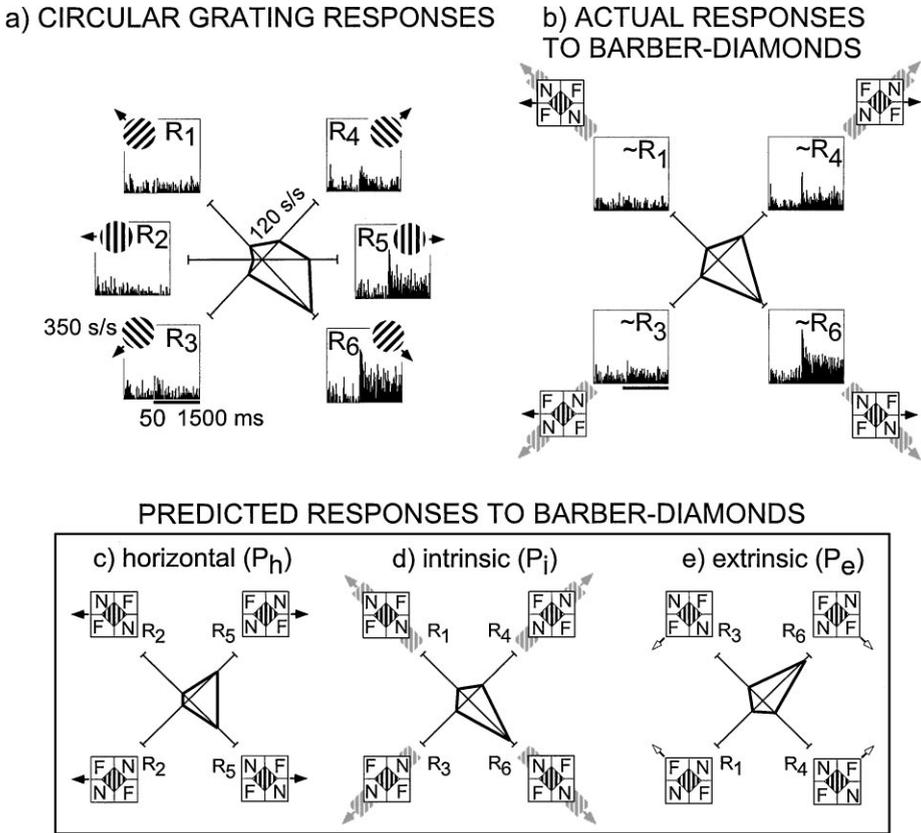


Figure 12 Influence of depth-ordering on direction selectivity of an MT neuron. (a) Neuronal responses to circular gratings moved in each of six directions. These responses (R1–R6) were used to form three different predictions for barber-diamond stimuli (c, d and e). (b) Actual responses to barber-diamond stimuli. Icons illustrate the stimulus configuration for each of four experimental conditions. (c) If depth-ordering in the receptive field surround has no effect, neuronal responses will be influenced solely by the horizontal motion (black arrows) of the grating and, hence, be of the same relative magnitude as responses to circular gratings moved leftward (R2) and rightward (R5), regardless of depth-ordering configuration. (d) If, as expected, neuronal responses are associated with the direction of intrinsic terminator motion (gray arrows), they should be of the same relative magnitude as responses to circular grating moving in the corresponding oblique directions (R1, R3, R4, and R6). (e) If neuronal responses are associated with the direction of extrinsic terminator motion (white arrows), they should be of the same relative magnitude as the intrinsic motion prediction but flipped about the horizontal axis. Observed responses (b) of this neuron to barber-diamond stimuli were more closely correlated with the intrinsic motion prediction than with either the horizontal or extrinsic motion prediction. (From Duncan et al. 2000. Copyright 2000 by the Society for Neuroscience.)

the interior is signaled by the lack of activity of neurons with RFs positioned over intermediate spatial locations. An analogous proposal can be made for our sense of a temporally continuous world (for related discussion see O'Regan 1992). Say your eyes fixate a blue surface, then a red one; neurons sensitive to that chromatic change will respond transiently, and this activation renders a sense of redness even after the response has ended. Taken together, these considerations suggest a brain very stingy with its action potentials, none being wasted on signaling that things have not changed over space or time.

Although it is appealingly economical, arguments and evidence against the sufficiency of this implicit encoding scheme are not hard to find. One empirical argument comes from studies of brightness induction. The interior of a gray patch will appear to undergo a brightness change if the surrounding area is temporally modulated from light to dark. If the frequency of that modulation is varied sinusoidally, it will be discovered that this induction effect disappears above about 2.5 Hz (De Valois et al. 1986, Rossi & Paradiso 1996). This temporal cutoff is much lower than that at which the border between the gray patch and surround appear unchanging. The perception of edges and interior brightness can thus be decoupled and we can conclude that the two cannot be mediated by the same pattern of neuronal activity.

A more general problem with the implicit coding scheme is that contrast measures at the nearest borders are simply not sufficient to support color and brightness perception of interior zones: As discussed below, the visual system integrates diverse types of information from large regions of visual space to infer the reflectance of a given region of visual space (e.g., Land & McCann 1971, Adelson 1993, Brown & MacLeod 1997, Barnes et al. 1999, Purves et al. 1999, Shevell & Wei 1998, Wachtler et al. 2001) (Figure 13*B*). Very complicated nontopographic schemes involving the neurons responding to these diverse information sources could, of course, be advanced, but they hardly seem parsimonious. Fortunately, whether there exist topographic neuronal responses related to brightness perception is not a question whose answer is hostage to theory.

Paradiso and colleagues (Rossi et al. 1996, Rossi & Paradiso 1999, MacEvoy & Paradiso 2001) have examined neuronal correlates of brightness induction. In one study they recorded from retina, thalamic lateral geniculate nucleus, and area V1 of anesthetized cats. The experiments were designed to determine whether neurons respond to brightness changes caused by either "direct" or contextual ("induced") means. Direct brightness changes were produced by varying the intensity of a central patch of light superimposed on the classical RF. Induced brightness changes were caused by holding the classical RF stimulus constant and varying light intensity in the RF surround. A substantial fraction of V1 neurons—unlike those at earlier processing stages—exhibited responses that covaried with the brightness of the classical RF stimulus, regardless of whether those brightness changes were caused directly or induced.

The findings of Paradiso and colleagues and those from recent related studies (Hung et al. 2001) suggest a topographically explicit, rather than implicit, representation of brightness. This conclusion does not refute the observations of Hubel &

Wiesel (1968)—cortical neurons do, in fact, respond to luminance contrast within the classical RF but weakly or not at all to “diffuse light.” Some cortical neurons are, however, sensitive to luminance contrast cues outside the classical field, and that sensitivity permits them to, in effect, encode the ambiguous properties of homogenous regions within that field. Evidence indicates that the observed topographic neuronal representation of uniform image regions develops progressively from border to interior. In human observers this progressive completion can be observed directly and its time course measured (Paradiso & Hahn 1996, Paradiso & Nakayama 1991). Neuronal evidence from “artificial scotoma” experiments (reviewed above) (De Weerd et al. 1995) also supports progressive completion, although it is important to note that these artificial scotomas were fixed (at least transiently) relative to eye position, unlike uniform surfaces under normal viewing conditions.

RECOVERING SURFACE REFLECTANCE: USING CONTEXT TO “SEE BENEATH” THE ILLUMINANT

Having reviewed evidence for an active neuronal filling-in of uniform image regions, we now turn to function: What is completed, and why? We begin by emphasizing a resemblance to the “recovery” of visual information lost owing to occlusion. As we have seen, amodal completion answers the question “What lies behind?” by predicting the content of occluded surfaces. Similarly, completion of uniform image regions identifies what is in front. Although we have portrayed this latter form of completion as propagation of information from edges to interior, the computation required is generally far more complex. As amodal completion must “see beneath” occluding surfaces, so must the mechanisms of foreground surface completion see beneath the light illuminating the scene in order to recover surface reflectance.

To appreciate the magnitude and significance of this problem, recall that the light reaching the retina is a generally a product of two factors: surface reflectance and illumination. Reflectance relates the quantity of light reflected to its wavelength. Reflectance is constant over time (though it may vary over space) for most surfaces and is thus a valuable source of information for object recognition. Illumination is characterized by its intensity as a function of wavelength. Unlike reflectance, however, the illuminant can and often does vary over space and time. Because surface reflectance and illumination properties are confounded in the light reaching the retina, it is impossible to deduce the reflectance of a surface given only the light arising from that surface. Despite this limitation, the primate visual system is quite good at recovering surface reflectance in natural scenes even when faced with significant changes in illumination. (For example, the cover of this book will appear red under many different lighting conditions.) This invariance of color appearance is known as “color constancy.” Brightness constancy is an achromatic analogue of color constancy, which is likely to result from similar neuronal substrates and mechanisms.

Vision scientists have puzzled over color constancy for centuries. Thomas Young (1807) observed, “When a room is illuminated by either the yellow light of a candle, or the red light of a fire, a sheet of writing paper still appears to retain its whiteness.” Helmholtz (2000) also recognized the phenomenon and its functional significance: “Colors are mainly important for us as properties of objects and as means of identifying objects. . . . we constantly aim to reach a judgment on the object colors and to eliminate differences of illumination.” Helmholtz largely attributed the observer’s ability to discount the illuminant, and thereby recover surface reflectance, to prior experience with objects in different lighting conditions. Many of Helmholtz’s contemporaries and followers offered explanations rooted solely in spatial context (e.g., Hering 1964 [1920], Mach 1886).

Contextual Influences on Color Appearance

The influence of context on color appearance has long been a popular topic in the field of visual psychophysics, and there now exists an enormous body of data characterizing human sensitivity. One of the most robust effects of context in this domain—one that has been known for centuries—is the dependency of perceived color on the color of adjacent regions (Figure 13A). From a functional perspective, simple context-induced shifts in color appearance can be viewed as a by-product of a system designed to recover surface reflectance in the face of changing illumination: While an illumination change will alter the spectrum of light arising from a given surface, it does so for every surface so illuminated. Therefore, as noted by the psychologist Hans Wallach (1948), the ratios of spectra from adjacent surfaces generally remain unchanged. Invariant color appearance should thus result from stimuli that possess spectral ratios that are invariant between adjacent image regions. Conversely, in the absence of ratio invariance, perceived color should change, as it indeed does.

Neuronal Correlates of Color Appearance

Until recently, very little was known of the neuronal bases of color appearance. Although there have been several detailed neuropsychological studies of acquired deficiencies of human color vision (achromatopsia) (for review see Zeki 1990), it has been difficult to discover from these studies brain regions specifically involved in representing perceived color, as opposed to low-level color detection and analysis. The past four decades have also seen numerous electrophysiological studies of neurons in the retina, thalamus, and visual cortex that respond to chrominance and/or luminance contrast within their classical RFs (for review see Lennie & D’Zmura 1988). However, most of these studies did not include contextual manipulations that alter color appearance, such as those shown in Figure 13A. Without such manipulations it is impossible to determine whether a neuronal response covaries with the spectral power of the light in the classical RF or, alternatively, with perceived color/brightness. We now turn to a few recent studies that have included appropriate contextual manipulations and have obtained intriguing results.

ACHROMATIC COLOR APPEARANCE (BRIGHTNESS) The complexity of color appearance phenomena is reduced when considering achromatic stimuli, but similar contextual effects and mechanistic interpretations apply. We have already discussed the brightness induction effects reported for neurons in cat area V1 by Paradiso and colleagues (Rossi et al. 1996, Rossi & Paradiso 1999). Not only might these effects underlie perceptual filling-in of the interiors of uniform image regions, but they satisfy one of the criteria for a mechanism that recovers surface reflectance. Specifically, the neuronal response changes when the ratio of intensities changes between adjacent surfaces. If these neurons do, in fact, represent surface reflectance, we should also expect them to exhibit an invariant response under conditions that mimic variations in diffuse illumination, i.e., where contrast remains unchanged. Recent evidence is consistent with this requirement (MacEvoy & Paradiso 2001).

CHROMATIC COLOR APPEARANCE Zeki (1983a,b) was the first to attempt to distinguish between neuronal representations of chromatic spectral composition and perceived color. Zeki adopted a stimulus configuration (Figure 13B) that had been a favorite tool of the color theorist Edwin Land for manipulating color appearance (Land & McCann 1971). The stimulus is known as a “Mondrian” for its resemblance to the style of the painter. These stimuli possess rich chromatic structure and they elicit robust color constancy effects.

In Zeki’s experiments, one of the colored panels in the Mondrian was placed over the classical RF of a cortical neuron, and the remaining panels were placed in the RF surround. When the color of the illuminant changed, the responses of neurons in area V1 changed in accordance with the wavelength of light falling in the classical RF. By contrast, the responses of a subset of neurons in area V4 (see Figure 3) were unaltered by illumination changes and thus paralleled the perceived constancy of the RF stimulus. Zeki suggested, on the basis of these results, that V4 may be the site at which chromatic context is integrated with the RF stimulus in order to achieve a representation of perceived color.

Using simpler stimulus manipulations of the type shown in Figure 13A, Schein & Desimone (1990) also examined the influence of chromatic context on the color selectivity of V4 neurons. Figure 14 illustrates one of their findings. The response of this neuron to the “preferred” chromatic stimulus placed in the classical RF was dependent upon the color of the surround. To a first approximation, these results are consistent with a mechanism in which the output is a function of chromatic contrast between the classical RF stimulus and the surround. But what is the function of such cells? Schein & Desimone speculated that they may signal the average intensity of the illuminant in specific wave-bands. The response of the cell illustrated, for example, should decline as the redness of the illuminant increases. Cells that serve as “wave-band intensity gauges” could, by explicitly recovering the properties of the illuminant, permit the illuminant to be discounted and surface reflectance computed (Land 1986).

Wachtler et al. (1999) used similarly simple stimuli to re-examine the role of area V1 in color appearance. In contrast to Schein & Desimone (1990) and Rossi

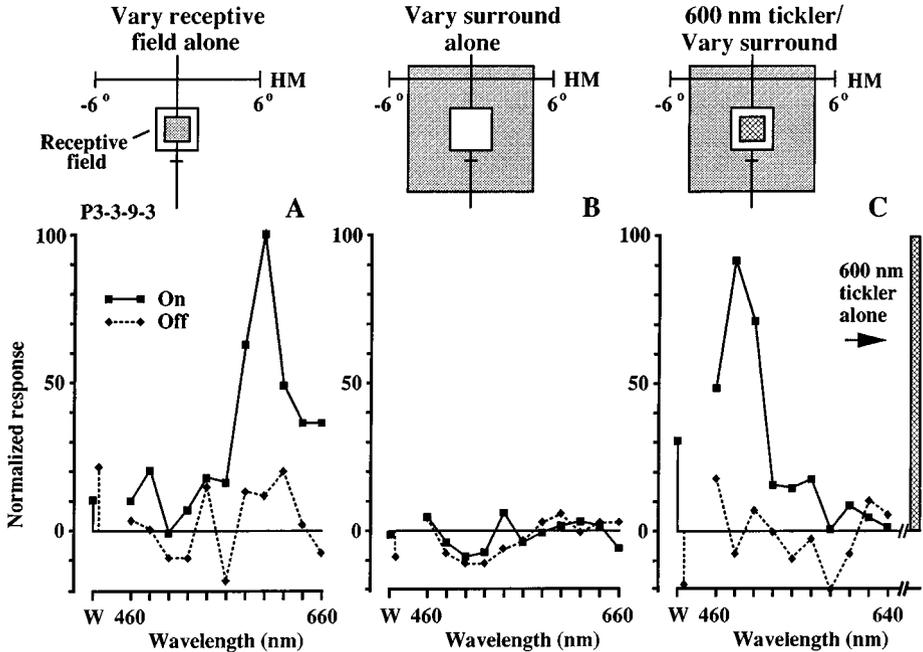


Figure 14 Chromatic selectivity of a V4 neuron is influenced by chromatic context. (*Left*) Wavelength tuning reveals that the neuron responds optimally to a stimulus of 600 nm in the classical receptive field (RF). (*Center*) The cell failed to respond to a chromatic background that extended around but did not intrude upon the classical RF. (*Right*) The chromatic background did, however, modulate the response to the preferred stimulus in the classical RF. To a first approximation, these results are consistent with a mechanism in which the output is a function of chromatic contrast between adjacent stimuli. Moreover, the modulation of response to the classical RF stimulus parallels the context-induced shift in color appearance that occurs for such stimuli. (From Schein & Desimone 1990. Copyright 1990 by the Society for Neuroscience.)

et al. (1996), Wachtler et al. evaluated the effects of color changes in the RF surround on the full chromatic tuning to the classical RF stimulus. For many V1 neurons the effects were striking. For example, introduction of a reddish surround to a red-preferring neuron had the effect of shifting the chromatic tuning curve away from red and toward the green side of color space. Significantly, the same contextual manipulation causes the central square to appear more greenish to an observer (Wachtler et al. 2001). Interestingly, this type of neuronal response shift is similar to what one would expect from double-opponent color-selective neurons, which have been thought to play a role in color constancy (Conway 2001). The important difference, however, is that the contextual manipulations of Wachtler et al. were highly effective even though they were restricted to the RF surround.

These results suggest that at least some V1 neurons use context to achieve a representation of perceived color. Although this conclusion mirrors that of Rossi et al. (1996) regarding the role of V1 in brightness perception, it is at variance with Zeki's (1983a,b) claim that such contextual effects are not seen until V4. This difference is difficult to reconcile, although it is noteworthy that Wachtler et al. studied V1 responses in alert monkeys, whereas Zeki's subjects were anesthetized.

Context and Color: Beyond Spatial Effects

We have thus far considered the simple contextual conditions in which the color appearance of a surface is influenced by light arising from adjacent (or nearby) coplanar surfaces. In most real scenes, however, the light arising from surfaces is determined not only by their reflectance and the properties of the illuminant, but also generally by the angle of the surface relative to the observer and the illumination source. One important consequence is that color ratios between surfaces, which provide evidence of reflectance and illumination, can serve as meaningful cues only if interpreted in the additional context of three-dimensional surface configuration. Striking demonstrations and psychophysical confirmation of this important caveat have been provided by Adelson (1993) and Purves et al. (1999) (see also Mach 1886), but nothing is currently known of the neuronal bases of these effects. Although this is an important area for future research, the number and complexity of contextual cue interactions required present a formidable challenge for controlled neurophysiological experiments.

CONCLUDING REMARKS

Computational considerations and a substantial body of psychophysical and physiological evidence indicate that the visual system employs a well-characterized set of contextual cues to identify the visual scene origins—i.e., the meaning—of visual image features. A common theme in this review is the use of contextual information to recover visual scene properties lost owing to the superimposition of causes in the visual image. We have, for example, discussed the contextual cues used by the visual system to decompose images into multiple overlapping surfaces and/or lighting conditions. Moreover, we have shown that individual neurons at early stages of cortical processing (e.g., V1 and V2) are involved in the detection of these contextual cues and the representation of recovered information.

This recent progress and the attention directed to the role of context represents a critical transition in efforts to understand the neuronal bases of perception. Historically, the reductionism of single-cell physiology has fostered a view of visual neurons as functionally independent windows on the world. In practice, this view has led to an experimental approach that neglects context. The result is that most of the reported responses of visual neurons (e.g., Figure 2) can only be described as neural correlates of the visual image; the full contributions of these neurons to visual scene perception remain unknown.

The transition toward a context-based approach has been inspired, in part, by the discovery of modulatory influences from outside the classical RF. Of even greater importance, perhaps, has been the increased willingness of visual physiologists to turn to experimental psychology for guidance, inspiration, and tools. Embedded in this approach is a renewed focus on function. Accordingly, many of the investigations of neuronal representation reviewed herein are motivated not simply by the question, “How does the visual system work?” but also by the question, “What is the visual system ‘designed’ to do?”

The studies we have reviewed are only the tip of an iceberg. The contextual influences we have omitted include, among others, attention and memory, which have been topics of many neuroscientific studies in recent years. There is also a vast realm of contextual effects that are closely related to those we have reviewed, whose manifestations are termed “perceptual constancies.” These constancies—color (reviewed above), size, viewing angle, etc.—reflect the recovery of invariant attributes of the visual scene in the face of changing image properties. Further identification of neural correlates of perceptual constancies is an important goal for the future.

Finally, in considering future directions, we emphasize that discovery of neuronal correlates is only one step toward understanding the neuronal bases of perception. Evidence of such correlates is now seemingly in abundance but, by contrast, we know very little about how these interesting neuronal properties are implemented. For example, horizontal intrinsic cortical connections are implicated in many phenomena of spatial context, such as filling-in and color appearance effects, but few firm conclusions can be reached. The complexity of cortical wiring and synaptic connectivity/plasticity is indeed daunting. New technologies that enable monitoring of global patterns of neuronal activity and manipulation of that activity within complex circuits nonetheless offer a promising armament for future efforts to reveal the mechanisms underlying contextual influences on visual processing.

ACKNOWLEDGMENTS

We thank Charles Gilbert, Greg Horwitz, Mike Paradiso, and John Reynolds for stimulating discussions and comments on the manuscript. The authors have been supported by grants from the NEI. TDA is an Investigator of the Howard Hughes Medical Institute.

The *Annual Review of Neuroscience* is online at <http://neuro.annualreviews.org>

LITERATURE CITED

- Adelson EH. 1993. Perceptual organization and the judgment of brightness. *Science* 262: 2042–44
- Albright TD. 1984. Direction and orientation selectivity of neurons in visual area MT of the macaque. *J. Neurophysiol.* 52:1106–30
- Allman J, Miezin F, McGuinness E. 1985. Stimulus specific responses from beyond the

- classical receptive field: neurophysiological mechanisms for local-global comparisons in visual neurons. *Annu. Rev. Neurosci.* 8:407–30
- Anderson BL. 1994. The role of partial occlusion in stereopsis. *Nature* 367:365–68
- Bakin JS, Nakayama K, Gilbert CD. 2000. Visual responses in monkey areas V1 and V2 to three-dimensional surface configurations. *J. Neurosci.* 20:8188–98
- Barlow HB. 1953. Summation and inhibition in the frog's retina. *J. Physiol.* 119:69–88
- Barlow HB. 1972. Single units and sensation: a neuron doctrine for perceptual psychology? *Perception* 1:371–94
- Barnes CS, Wei J, Shevell SK. 1999. Chromatic induction with remote chromatic contrast varied in magnitude, spatial frequency, and chromaticity. *Vis. Res.* 39:3561–74
- Baumann R, ver der Zwan R, Peterhans E. 1997. Figure-ground segregation at contours: a neural mechanism in the visual cortex of the alert monkey. *Eur. J. Neurosci.* 9:1290–303
- Beck J, Prazdny K, Ivry R. 1984. The perception of transparency with achromatic colors. *Percept. Psychophys.* 35:407–22
- Bishop PO, Coombs JS, Henry GH. 1973. Receptive fields of simple cells in the cat striate cortex. *J. Physiol.* 231:31–60
- Bosking WH, Zhang Y, Schofield B, Fitzpatrick D. 1997. Orientation selectivity and the arrangement of horizontal connections in tree shrew striate cortex. *J. Neurosci.* 17:2112–27
- Bradley DC, Andersen RA. 1998. Center-surround antagonism based on disparity in primate area MT. *J. Neurosci.* 18:7552–65
- Brown RO, MacLeod DI. 1997. Color appearance depends on the variance of surround colors. *Curr. Biol.* 7:844–49
- Buracas GT, Albright TD. 1996. Contribution of area MT to perception of three-dimensional shape: a computational study. *Vis. Res.* 36:869–87
- Campbell FW, Andrews PR. 1992. Motion reveals spatial visual defects. *Ophthalmic Physiol. Optics* 12:131–32
- Chang G, Stoner GR, Albright TD. 1999. Neuronal correlates for depth-ordering in area V2. *Invest. Ophthalmol. Vis. Sci.* 40(4): S640
- Chang GC, Stoner GR, Albright TD. 2001. Figure-ground selectivity in macaque V2. *Soc. Neurosci. Abstr.* 27:432
- Chino YM, Kaas JH, Smith EL, Langston AL, Cheng H. 1992. Rapid reorganization of cortical maps in adult cats following restricted deafferentation in retina. *Vis. Res.* 32:789–96
- Churchland PS, Ramachandran VI. 1993. Filling-in: why Dennett is wrong. In *Dennett and His Critics: Demystifying Mind*, ed. B Dahlbom, pp. vii, 247. Oxford/Cambridge, MA: Blackwell
- Conway BR. 2001. Spatial structure of cone inputs to color cells in alert macaque primary visual cortex (V-1). *J. Neurosci.* 21:2768–83
- Croner LJ, Albright TD. 1999a. Segmentation by color influences responses of motion-sensitive neurons in the cortical middle temporal visual area. *J. Neurosci.* 19:3935–51
- Croner LJ, Albright TD. 1999b. Seeing the big picture: integration of image cues in the primate visual system. *Neuron* 24:777–89
- Cumming BG, Parker AJ. 1998. Eye movements explain contextual modulation by binocular disparity in primate V1. *Soc. Neurosci. Abstr.* 24:1980
- De Valois RL, Webster MA, De Valois KK, Lingelbach B. 1986. Temporal properties of brightness and color induction. *Vis. Res.* 26: 887–97
- De Weerd P, Desimone R, Ungerleider LG. 1998. Perceptual filling-in: a parametric study. *Vis. Res.* 38:2721–34
- De Weerd P, Gattass R, Desimone R, Ungerleider LG. 1995. Responses of cells in monkey visual cortex during perceptual filling-in of an artificial scotoma. *Nature* 377:731–34
- Dennett DC. 1991. *Consciousness Explained*. New York: Little Brown
- DeYoe EA, Bartlett JR. 1980. Rarity of luxotonic responses in cortical visual areas of the cat. *Exp. Brain. Res.* 39:125–32
- Dresp B. 1993. Bright lines and edges facilitate the detection of small light targets. *Spatial Vis.* 7:213–25
- Duncan RO, Albright TD, Stoner GR. 2000.

- Occlusion and the interpretation of visual motion: perceptual and neuronal effects of context. *J. Neurosci.* 20:5885–97
- Field DJ, Hayes A, Hess RF. 1993. Contour integration by the human visual system: evidence for a local “association field.” *Vis. Res.* 33:173–93
- Fiorani JM, Rosa MG, Gattass R, Rocha-Miranda CE. 2001. Dynamic surrounds of receptive fields in primate striate cortex: a physiological basis for perceptual completion? *Proc. Natl. Acad. Sci. USA* 89:8547–51
- Gallant JL, Connor CE, Van Essen DC. 1998. Neural activity in areas V1, V2 and V4 during free viewing of natural scenes compared to controlled viewing. *Neuroreport* 9:1673–78
- Garraghty PE, Kaas JH. 1992. Dynamic features of sensory and motor maps. *Curr. Opin. Neurobiol.* 2:522–27
- Gerrits HJ, Timmerman GJ. 1969. The filling-in process in patients with retinal scotomata. *Vis. Res.* 9:439–42
- Gerrits HJM, de Haan B, Vendrick AJH. 1966. Experiments with retinal stabilized images. Relations between the observations and neural data. *Vision Res.* 6:427–440
- Gerrits HJ, Vendrik AJ. 1970. Simultaneous contrast, filling-in process and information processing in man’s visual system. *Exp. Brain Res.* 11:411–30
- Gilbert CD, Wiesel TN. 1983. Clustered intrinsic connections in cat visual cortex. *J. Neurosci.* 3:1116–33
- Gilbert CD, Wiesel TN. 1990. The influence of contextual stimuli on the orientation selectivity of cells in primary visual cortex of the cat. *Vis. Res.* 30:1689–701
- Gilbert CD, Wiesel TN. 1992. Receptive field dynamics in adult primary visual cortex. *Nature* 356:150–52
- Gilchrist AL. 1977. Perceived lightness depends on perceived spatial arrangement. *Science* 195:185–87
- Grosf DH, Shapley RM, Hawken MJ. 1993. Macaque V1 neurons can signal ‘illusory’ contours. *Nature* 365:550–52
- Gulyas B, Orban GA, Duysens J, Maes H. 1987. The suppressive influence of moving textured backgrounds on responses of cat striate neurons to moving bars. *J. Neurophysiol.* 57:1767–91
- Hartline HK, Graham CH. 1932. Nerve impulses from single receptors in the eye. *J. Cell Comp. Physiol.* 1:277–95
- He S, Davis WL. 2001. Filling-in at the natural blind spot contributes to binocular rivalry. *Vis. Res.* 41:835–40
- Hering E. 1964 (1920). *Outlines of a Theory of the Light Sense*. Cambridge, MA: Harvard Univ. Press
- Hubel DH, Wiesel TN. 1965. Receptive fields and functional architecture in two nonstriate visual areas (18 and 19) of the cat. *J. Neurophysiol.* 28:228–89
- Hubel DH, Wiesel TN. 1968. Receptive fields and functional architecture of monkey striate cortex. *J. Physiol.* 195:215–43
- Hung CP, Ramsden BM, Chen LM, Roe AW. 2001. Building surfaces from borders in areas 17 and 18 of the cat. *Vis. Res.* 41:1389–407
- James W. 1890. *Principles of Psychology*. New York: Holt
- Kaas JH. 1991. Plasticity of sensory and motor maps in adult mammals. *Annu. Rev. Neurosci.* 14:137–67
- Kahrilas PJ, Doty RW, Bartlett JR. 1980. Failure to find luxotonic responses for single units in visual cortex of the rabbit. *Exp. Brain Res.* 39:11–16
- Kanizsa G. 1979. *Organization in Vision: Essays on Gestalt Perception*. New York: Praeger
- Kapadia MK, Ito M, Gilbert CD, Westheimer G. 1995. Improvement in visual sensitivity by changes in local context: parallel studies in human observers and in V1 of alert monkeys. *Neuron* 15:843–56
- Kapadia MK, Westheimer G, Gilbert CD. 1999. Dynamics of spatial summation in primary visual cortex of alert monkeys. *Proc. Natl. Acad. Sci. USA* 96:12073–78
- Kapadia MK, Westheimer G, Gilbert CD. 2000. Spatial distribution of contextual interactions in primary visual cortex and in visual perception. *J. Neurophysiol.* 84:2048–62
- Kaplan GA. 1969. Kinetic disruption of optical

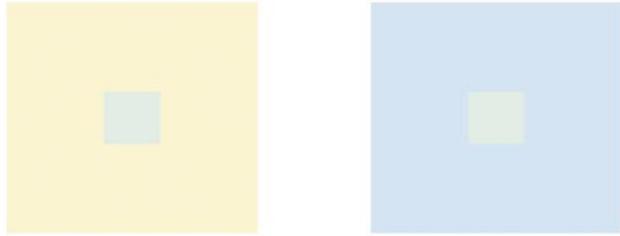
- texture: the perception of depth at an edge. *Percept. Psychophys.* 6:193–98
- Kayama Y, Riso RR, Bartlett JR, Doty RW. 1979. Luxotonic responses of units in macaque striate cortex. *J. Neurophysiol.* 42:1495–517
- Knierim JJ, van Essen DC. 1992. Neuronal responses to static texture patterns in area V1 of the alert macaque monkey. *J. Neurophysiol.* 67:961–80
- Koffka K. 1935. *Principles of Gestalt Psychology*. New York: Harcourt Brace
- Komatsu H, Kinoshita M, Murakami I. 2000. Neural responses in the retinotopic representation of the blind spot in the macaque V1 to stimuli for perceptual filling-in. *J. Neurosci.* 20:9310–19
- Lamme VAF. 1995. The neurophysiology of figure-ground segregation in primary visual cortex. *J. Neurosci.* 15:1605–15
- Land EH. 1986. Recent advances in retinex theory. *Vis. Res.* 26:7–21
- Land EH, McCann JJ. 1971. Lightness and retinex theory. *J. Opt. Soc. Am.* 61:1–11
- Le Gros. 1967. Entopic phenomenon. In *Le Grand, Yves, 1908—Form and Space Vision*. Transl. M Millodot, GG Heath, pp. 146–58. Bloomington: Indiana Univ. Press (From French)
- Lee TS, Nguyen M. 2001. Dynamics of subjective contour formation in the early visual cortex. *Proc. Natl. Acad. Sci. USA* 98:1907–11
- Lennie P, D’Zmura M. 1988. Mechanisms of color vision. *Crit. Rev. Neurobiol.* 3:333–400
- Li CY, Li W. 1994. Extensive integration field beyond the classical receptive field of cat’s striate cortical neurons—classification and tuning properties. *Vis. Res.* 34:2337–55
- Li W, Thier P, Wehrhahn C. 2000. Contextual influence on orientation discrimination of humans and responses of neurons in V1 of alert monkeys. *J. Neurophysiol.* 83:941–54
- Li W, Thier P, Wehrhahn C. 2001. Neuronal responses from beyond the classic receptive field in V1 of alert monkeys. *Exp. Brain Res.* 139:359–71
- Liden L, Pack C. 1999. The role of terminators and occlusion cues in motion integration and segmentation: a neural network model. *Vis. Res.* 39:3301–20
- Liu Z, Kersten D. 1998. 2D observers for human 3D object recognition? *Vis. Res.* 38:2507–19
- Livingstone MS, Hubel DH. 1987. Psychophysical evidence for separate channels for the perception of form, color, movement, and depth. *J. Neurosci.* 7:3416–68
- Locke J. 1690. *An Essay Concerning Human Understanding*. London: Basset
- MacEvoy SP, Paradiso MA. 2001. Lightness constancy in primary visual cortex. *Proc. Natl. Acad. Sci. USA* 98:8827–31
- Mach E. 1886. *Contributions to the Analysis of Sensations*. Trans. CM Williams, 1910. Chicago: Open Court
- Maffei L, Fiorentini A. 1976. The unresponsive regions of visual cortical receptive fields. *Vis. Res.* 16:1131–39
- Mariotte E. 1668. *Nouvelle Découverte Touchant la Veve*. Paris: Frederic Leonard. 27 pp.
- Martin KA, Whitteridge D. 1984. The relationship of receptive field properties to the dendritic shape of neurones in the cat striate cortex. *J. Physiol.* 356:291–302
- McIlwain JT. 1964. Receptive fields of optic tract axons and lateral geniculate cells: peripheral extent and barbiturate sensitivity. *J. Neurophysiol.* 27:1154–73
- Michotte AE, Thinès GO, Crabbé G. 1964. *Les Compléments Amodaux des Structures Perceptives*. Louvain: Inst. Psychol. Univ. Louvain. 56 pp.
- Mountcastle VB, LaMotte RH, Carli G. 1972. Detection thresholds for stimuli in humans and monkeys: comparison with threshold events in mechanoreceptive afferent nerve fibers innervating the monkey hand. *J. Neurophysiol.* 35:122–36
- Müller J. 1838. *Elements of Physiology*. London: Printed for Taylor & Walton. 2 vols.
- Murakami I, Komatsu H, Kinoshita M. 1997. Perceptual filling-in at the scotoma following a monocular retinal lesion in the monkey. *Vis. Neurosci.* 14:89–101
- Nakayama K, Shimojo S. 1990. Toward a neural

- understanding of visual surface representation. *Cold Spring Harbor Symp. Quant. Biol.* LV: 911–24
- Nakayama K, Shimojo S, Silverman GH. 1989. Stereoscopic depth: its relation to image segmentation, grouping, and the recognition of occluded objects. *Perception* 18:55–68
- Nelson JI, Frost BJ. 1978. Orientation-selective inhibition from beyond the classic visual receptive field. *Brain Res.* 139:359–65
- Nelson JI, Frost BJ. 1985. Intracortical facilitation among co-oriented, co-axially aligned simple cells in cat striate cortex. *Exp. Brain Res.* 61:54–61
- Newsome WT, Britten KH, Movshon JA. 1989. Neuronal correlates of a perceptual decision. *Nature* 341:52–54
- O'Regan JK. 1992. Solving the "real" mysteries of visual perception: the world as an outside memory. *Can. J. Psychol.* 46:461–88
- Paradiso MA, Hahn S. 1996. Filling-in percepts produced by luminance modulation. *Vision Res.* 36:2657–63
- Paradiso MA, Nakayama K. 1991. Brightness perception and filling-in. *Vis. Res.* 31:1221–36
- Peterhans E, von der Heydt R. 1989. Mechanisms of contour perception in monkey visual cortex. II. Contours bridging gaps. *J. Neurosci.* 9:1749–63
- Peterson MA. 2002. On figures, grounds, and varieties of surface completion. In *The 31st Carnegie Symposium: Perceptual Organization in Vision: Behavioral and Neural Perspectives*, ed. M Behrmann, R Kimchi, C Olson. Mahwah, NJ: Earlbaum. In press
- Pettet MW, Gilbert CD. 1992. Dynamic changes in receptive-field size in cat primary visual cortex. *Proc. Natl. Acad. Sci. USA* 89: 8366–70
- Polat U, Mizobe K, Pettet MW, Kasamatsu T, Norcia AM. 1998. Collinear stimuli regulate visual responses depending on cell's contrast threshold. *Nature* 391:580–84
- Polat U, Sagi D. 1993. Lateral interactions between spatial channels: suppression and facilitation revealed by lateral masking experiments. *Vis. Res.* 33:993–99
- Polat U, Sagi D. 1994. The architecture of perceptual spatial interactions. *Vis. Res.* 34:73–78
- Purves D, Shimpf A, Lotto RB. 1999. An empirical explanation of the cornsweet effect. *J. Neurosci.* 19:8542–51
- Ramachandran VS. 1992. Blind spots. *Sci. Am.* 266:86–91
- Ramachandran VS. 1994. Phantom limbs, neglect syndromes, repressed memories, and Freudian psychology. *Int. Rev. Neurobiol.* 37:291–333
- Ramachandran VS, Gregory RL. 1991. Perceptual filling in of artificially induced scotomas in human vision. *Lett. Nat.* 350:699–702
- Ramsden BM, Hung CP, Roe AW. 2001. Real and illusory contour processing in area V1 of the primate: a cortical balancing act. *Cereb. Cortex* 11:648–65
- Redies C, Crook JM, Creutzfeldt OD. 1986. Neuronal responses to borders with and without luminance gradients in cat visual cortex and dorsal lateral geniculate nucleus. *Exp. Brain Res.* 61:469–81
- Rockland KS, Lund JS. 1982. Widespread periodic intrinsic connections in the tree shrew visual cortex. *Science* 215:1532–34
- Rossi AF, Desimone R, Ungerleider LG. 2001. Contextual modulation in primary visual cortex of macaques. *J. Neurosci.* 21:1698–709
- Rossi AF, Paradiso MA. 1996. Temporal limits of brightness induction and mechanisms of brightness perception. *Vis. Res.* 36:1391–98
- Rossi AF, Paradiso MA. 1999. Neural correlates of perceived brightness in the retina, lateral geniculate nucleus, and striate cortex. *J. Neurosci.* 19:6145–56
- Rossi AF, Rittenhouse CD, Paradiso MA. 1996. The representation of brightness in primary visual cortex. *Science* 273:1104–7
- Rubin E. 1921. *Visuell Wahrgenommene Figuren; Studien in Psychologischer Analyse*. Kobenhavn: Gyldendalske Boghandel
- Schein SJ, Desimone R. 1990. Spectral properties of V4 neurons in the macaque. *J. Neurosci.* 10:3369–89
- Schmidt KE, Kim DS, Singer W, Bonhoeffer T, Lowel S. 1997. Functional specificity

- of long-range intrinsic and interhemispheric connections in the visual cortex of strabismic cats. *J. Neurosci.* 17:5480–92
- Sheth BR, Sharma J, Rao SC, Sur M. 1996. Orientation maps of subjective contours in visual cortex. *Science* 274:2110–15
- Shevell SK. 1978. The dual role of chromatic backgrounds in color perception. *Vis. Res.* 18:1649–61
- Shevell SK, Wei J. 1998. Chromatic induction: border contrast or adaptation to surrounding light? *Vis. Res.* 38:1561–66
- Shimojo S, Silverman GH, Nakayama K. 1989. Occlusion and the solution to the aperture problem for motion. *Vis. Res.* 29:619–26
- Sillito AM, Grieve KL, Jones HE, Cudeiro J, Davis J. 1995. Visual cortical mechanisms detecting focal orientation discontinuities. *Nature* 378:492–96
- Stanley GB, Li FF, Dan Y. 1999. Reconstruction of natural scenes from ensemble responses in the lateral geniculate nucleus. *J. Neurosci.* 19:8036–42
- Stoner GR, Albright TD. 1993. Image segmentation cues in motion processing: implications for modularity in vision. *J. Cogn. Neurosci.* 5:129–49
- Stoner GR, Albright TD. 1995. Depth from occlusion and non-Fourier motion. *Soc. Neurosci. Abstr.* 21:511
- Stoner GR, Duncan RO, Albright TD. 1998. Area MT responses to contrast-modulated stimuli may encode depth-from-dynamic occlusion. *Invest. Ophthalmol. Vis. Sci.* 39:S904
- Sugita Y. 1999. Grouping of image fragments in primary visual cortex. *Nature* 401:269–72
- Tanaka K, Hikosaka K, Saito H, Yukie M, Fukada Y, Iwai E. 1986. Analysis of local and wide-field movements in the superior temporal visual areas of the macaque monkey. *J. Neurosci.* 6:134–44
- Tong F, Engel SA. 2001. Interocular rivalry revealed in the human cortical blind-spot representation. *Nature* 411:195–99
- Vinje WE, Gallant JL. 2000. Sparse coding and decorrelation in primary visual cortex during natural vision. *Science* 287:1273–76
- von der Heydt R, Peterhans E, Baumgartner G. 1984. Illusory contours and cortical neuron responses. *Science* 224:1260–62
- von Helmholtz H. 2000 (1860/1924). *Treatise on Psychological Optics*, trans. JPC Southall. Rochester, NY: Opt. Soc. Am.
- Wachtler T, Albright TD, Sejnowski TJ. 2001. Nonlocal interactions in color perception: nonlinear processing of chromatic signals from remote inducers. *Vis. Res.* 41:1535–46
- Wachtler T, Sejnowski TJ, Albright TD. 1999. Responses of cells in macaque V1 to chromatic stimuli are compatible with human color constancy. *Soc. Neurosci. Abs.* 25(S7.8):4
- Wallach H. 1935. Uber visuell wahrgenommene Bewegungsrichtung. *Psychol. Forsch.* 20:325–80
- Wallach H. 1948. Brightness constancy and the nature of achromatic colors. *J. Exp. Psychol.* 38:310–24
- Walls G. 1954. The filling-in process. *Am. J. Optom.* 31:329–40
- Wertheimer M. 1961 (1912). *Classics in Psychology*, ed. T Shipley, pp. 1032–89. New York: Phil. Library
- Wundt WM. 1902. *Principles of Physiological Psychology*. Trans. EB Titchener. New York: Macmillan (from German)
- Xiao DK, Marcar VL, Raiguel SE, Orban GA. 1997. Selectivity of macaque MT/V5 neurons for surface orientation in depth specified by motion. *Eur. J. Neurosci.* 9:956–64
- Young T. 1807. *A Course of Lectures on Natural Philosophy and the Mechanical Arts*, 2 vols. London: Johnson
- Zeki S. 1983a. Colour coding in the cerebral cortex: the reaction of cells in monkey visual cortex to wavelengths and colours. *Neuroscience* 9:741–65
- Zeki S. 1983b. Colour coding in the cerebral cortex: the responses of wavelength-selective and colour-coded cells in monkey visual cortex to changes in wavelength composition. *Neuroscience* 9:767–81
- Zeki S. 1990. Colour vision and functional specialisation in the visual cortex. *Discuss. Neurosci.* 6:11–64

-
- Zenger B, Sagi D. 1996. Isolating excitatory and inhibitory nonlinear spatial interactions involved in contrast detection. *Vis. Res.* 36:2497–513
- Zhou H, Friedman HS, von der Heydt R. 2000. Coding of border ownership in monkey visual cortex. *J. Neurosci.* 20:6594–611
- Zipser K, Lamme VA, Schiller PH. 1996. Contextual modulation in primary visual cortex. *J. Neurosci.* 16:7376–89

A



B



Figure 13 Contextual influences on color appearance. (A) The colors of the two central squares are identical, but the one on the right appears greener than the one on the left. This difference in appearance of the central squares is caused by the physical difference between the larger background contexts. This illusion, which has been known for centuries, is caused by the contrast difference between the two image regions and is a product of a mechanism designed to recover the reflectance of surfaces. Neurons in area V1 exhibit context-induced shifts in their sensitivity to receptive field stimuli, which mirror changes in color appearance (Rossi et al. 1996, Wachtler et al. 1999, MacEvoy & Paradiso 2001). (B) “Mondrian” stimulus used to examine influence of context on perceptual and neuronal sensitivity to color (Land & McCann 1971, Zeki 1983a).