

## Intermediate-Level Visual Processing and Visual Primitives

### Internal Models of Object Geometry Help the Brain Analyze Shapes

#### Depth Perception Helps Segregate Objects from Background

#### Local Movement Cues Define Object Trajectory and Shape

#### Context Determines the Perception of Visual Stimuli

Brightness and Color Perception Depend on Context

Receptive-Field Properties Depend on Context

### Cortical Connections, Functional Architecture, and Perception Are Intimately Related

Perceptual Learning Requires Plasticity in Cortical Connections

Visual Search Relies on the Cortical Representation of Visual Attributes and Shapes

Cognitive Processes Influence Visual Perception

### An Overall View

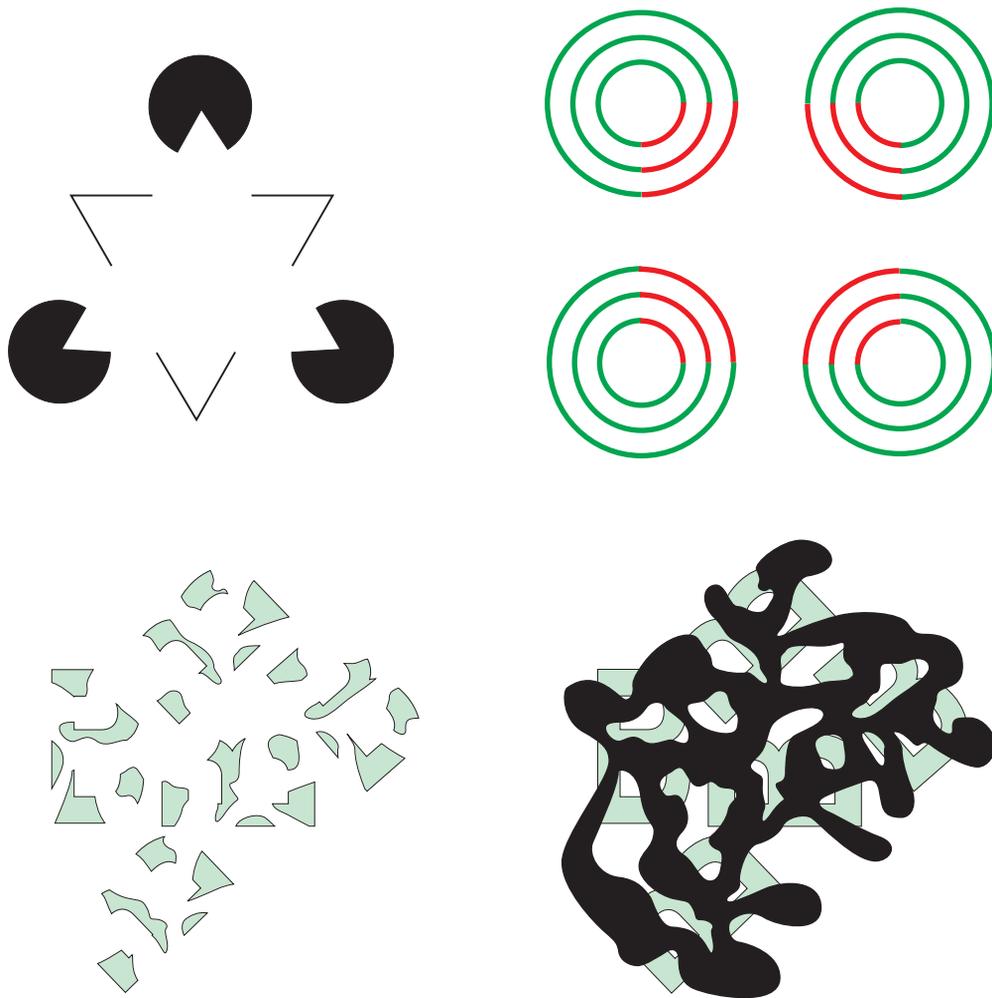
**W**E HAVE SEEN IN THE PRECEDING chapter that the eye is not a mere camera but instead contains sophisticated retinal circuitry that decomposes the retinal image into signals representing contrast and movement. These data are conveyed through the optic nerve to the primary visual cortex, which uses this information to analyze the shape of objects. It first identifies the boundaries of objects, represented by numerous short line segments, each with a specific orientation. The cortex then integrates this information into a representation of specific objects, a process referred to as *contour integration*.

These two steps, local analysis of orientation and contour integration, exemplify two distinct stages of

visual processing. Computation of local orientation is an example of low-level visual processing, which is concerned with identifying local elements of the light structure of the visual field. Contour integration is an example of intermediate-level visual processing, the first step in generating a representation of the unified visual field. At the earliest stages of analysis in the cerebral cortex these two levels of processing are accomplished together.

A visual scene comprises many thousands of line segments and surfaces. Intermediate-level visual processing is concerned with determining which boundaries and surfaces belong to specific objects and which are part of the background (see Figure 25–4). It is also involved in distinguishing the lightness and color of a surface from the intensity and wavelength of light reflected from that surface. The physical characteristics of reflected light result as much from the intensity and color balance of the light that illuminates a surface as from the color of that surface. Determining the actual surface color of a single object requires comparison of the wavelengths of light reflected from multiple surfaces in a scene.

Intermediate-level visual processing thus involves assembling local elements of an image into a unified percept of objects and background. Although determining which elements belong together in a single object is a highly complex problem with the potential for an astronomical number of solutions, the brain has built-in logic that allows it to make assumptions about the likely spatial relationships between elements. In certain cases these inherent rules can lead to the illusion of contours and surfaces that do not actually exist in the visual field (Figure 27–1).



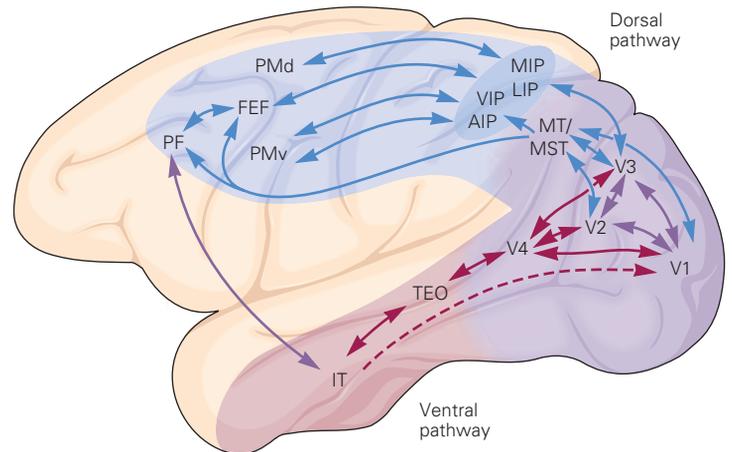
**Figure 27-1** Illusory contours and perceptual fill-in. The visual system uses information about local orientation and contrast to construct the contours and surfaces of objects. This constructive process can lead to the perception of contours and surfaces that do not appear in the visual field, including those seen in illusory figures. In the Kanizsa triangle illusion (**top left**) one perceives continuous boundaries extending between the apices of a white triangle, even though the only real contour elements are those formed by the Pac-Man-like figures and the

acute angles. The inside and outside of the illusory pink square (**top right**) are the same white color as the page, but a continuous transparent pink surface within the square is perceived. As seen in the lower figures, contour integration and surface segmentation can also occur through occluding surfaces. The irregular shapes on the left appear to be unrelated, but when a partially occluding black area is overlaid on them (**right**) they are easily seen as fragments of the letter B.

First, context plays an important role in overcoming ambiguity in the signals from the retina. The way in which a visual feature is perceived depends on everything that surrounds that feature. The perception of a point or a line depends on how that object is perceptually linked to other visual features. Thus the response of a neuron in the visual cortex is context-dependent: It depends as much on the presence of contours and surfaces outside the cell's receptive field as on the attributes within it. Second, the functional properties

of neurons in the visual cortex are highly dynamic and can be altered by visual experience or perceptual learning. Finally, visual processing in the cortex is subject to the influence of cognitive functions, specifically attention, expectation, and "perceptual task," ie, the active engagement in visual discrimination or detection. The interaction between these three factors—visual context, experience-dependent changes in cortical circuitry, and expectation—is vital in the visual system's analysis of complex scenes.

**Figure 27–2** Cortical areas involved with intermediate-level visual processing. Many cortical areas in the macaque monkey, including V1, V2, V3, V4, and middle temporal area (MT), are involved with integrating local cues to construct contours and surfaces and segregating foreground from background. The shaded areas extend into the frontal and temporal lobes because cognitive output from these areas, including attention, expectation, and perceptual task, contribute to the process of scene segmentation. (AIP, anterior intraparietal cortex; FEF, frontal eye fields; IT, inferior temporal cortex; LIP, lateral intraparietal cortex; MIP, medial intraparietal cortex; MST, medial superior temporal cortex; MT, middle temporal cortex; PF, prefrontal cortex; PMd, dorsal premotor cortex; PMv, ventral premotor cortex; TEO, occipitotemporal cortex; VIP, ventral intraparietal cortex; V1, V2, V3, V4, primary, secondary, third, and fourth visual areas.)



In this chapter we examine how the brain's analysis of the local features in a visual scene, or *visual primitives*, proceeds in parallel with the analysis of more global features. Visual primitives include contrast, line orientation, brightness, color, movement, and depth.

Each type of visual primitive is subject to the integrative action of intermediate-level processing. Lines with particular orientations are integrated into object contours, local contrast information into surface lightness, wavelength selectivity into color constancy and surface segmentation, and directional selectivity into object motion. The analysis of visual primitives begins in the retina with the detection of brightness and color and continues in the primary visual cortex with the analysis of orientation, direction of movement, and stereoscopic depth. Properties related to intermediate-level visual processing are analyzed together with visual primitives in the visual cortex starting in the primary visual cortex (V1), which plays a role in contour integration and surface segmentation. Other areas of the visual cortex specialize in different aspects of this task: V2 analyzes properties related to object surfaces, V4 integrates information about color and object shape, and V5—the middle temporal area or MT—integrates motion signals across space (Figure 27–2).

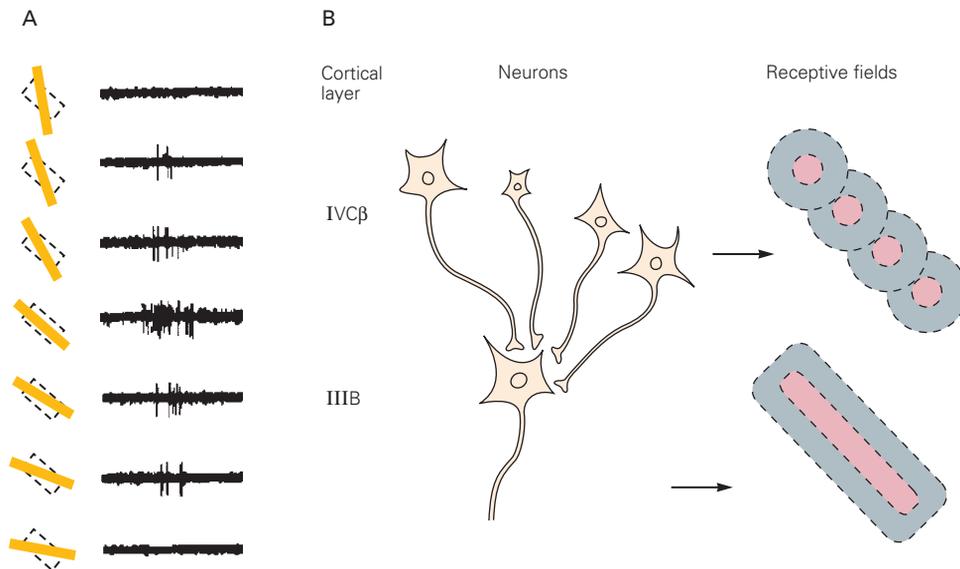
### Internal Models of Object Geometry Help the Brain Analyze Shapes

A first step in determining an object's contour is identification of the orientation of local parts of the contour. This step commences in V1, which plays a critical role in both local and global analysis of form.

Neurons in the visual cortex respond selectively to specific local features of the visual field, including orientation, binocular disparity or depth, and direction of movement, as well as to properties already analyzed in the retina and lateral geniculate nucleus, such as contrast and color. Orientation selectivity, the first emergent property identified in the receptive fields of cortical neurons, was discovered by David Hubel and Torsten Wiesel in 1959.

Neurons in the lateral geniculate nucleus have circular receptive fields with a center-surround organization (see Chapter 25). They respond to the light-dark contrasts of edges or lines in the visual field but are not selective for the orientations of those edges. In the visual cortex, however, neurons respond selectively to lines of particular orientations. Each neuron responds to a narrow range of orientations, approximately  $40^\circ$ , and different neurons respond optimally to distinct orientations. There is now good evidence for the idea, first proposed by Hubel and Wiesel, that this orientation selectivity reflects the arrangement of the inputs from cells in the lateral geniculate nucleus. Each V1 neuron receives input from several neighboring geniculate neurons whose center-surround receptive fields are aligned so as to represent a particular axis of orientation (Figure 27–3).

Two principal types of orientation-selective neurons have been identified. *Simple cells* have receptive fields divided into ON and OFF subregions (Figure 27–4). When a visual stimulus such as a bar of light enters the receptive field's ON subregion, the neuron fires; the cell also responds when the bar leaves the OFF subregion. Simple cells have a characteristic response to a moving bar; they discharge briskly when a bar of light leaves an OFF region and enters an ON region.

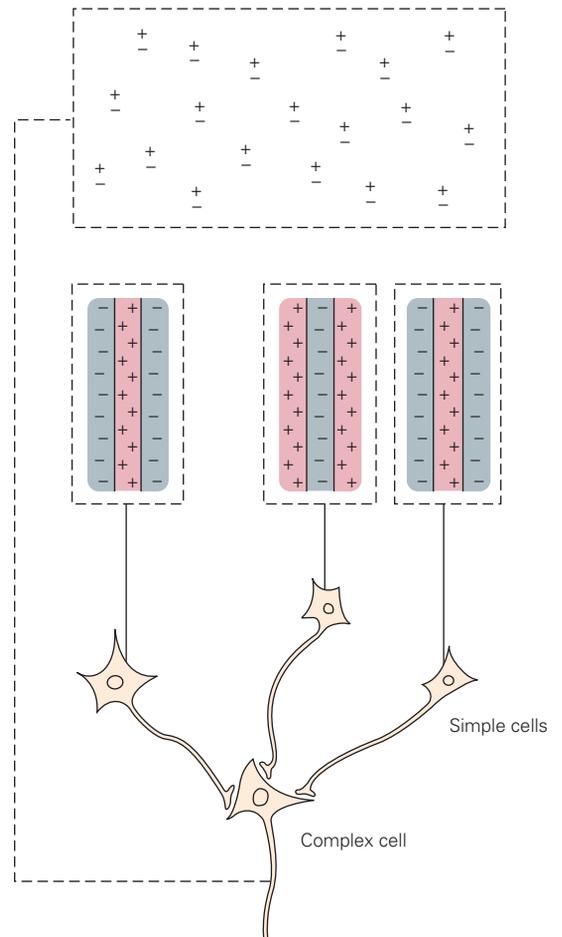


**Figure 27-3** Orientation selectivity and mechanisms.

A. A neuron in the primary visual cortex responds selectively to line segments that fit the orientation of its receptive field. This selectivity is the first step in the brain's analysis of an object's form. (Reproduced, with permission, from Hubel and Wiesel 1968.)

B. The orientation of the receptive field is thought to result from the alignment of the circular center-surround receptive fields of several presynaptic cells in the lateral geniculate nucleus. In the monkey, neurons in layer IVC $\beta$  of V1 have unoriented receptive fields. However, the projections of neighboring IVC $\beta$  cells onto a neuron in layer IIIB create a receptive field with a specific orientation.

**Figure 27-4** Simple and complex cells in the visual cortex. The receptive fields of simple cells are divided into subfields with opposite response properties. In an ON subfield, designated by "+," the onset of a light triggers a response in the neuron; in an OFF subfield, indicated by "-", the extinction of a bar of light triggers a response. Complex cells have overlapping ON and OFF regions and respond continuously as a line or edge traverses the receptive field along an axis perpendicular to the receptive-field orientation.



The responses of these cells are therefore highly selective for the position of a line or edge in space.

*Complex cells*, in contrast, are less selective for the position of object boundaries. They lack discrete ON and OFF subregions and respond similarly to light and dark at all locations across their receptive fields. They fire continuously as a line or edge stimulus traverses their receptive fields.

Moving stimuli are often used to study the receptive fields of visual cortex neurons, not only to simulate the conditions under which an object moving in space is detected but also to simulate the conditions under which stationary objects are tracked by the eyes, which constantly scan the visual environment and therefore move the boundaries of stationary objects across the retina. In fact, visual perception requires eye movement. Visual cortex neurons do not respond to an image that is stabilized on the retina because they require moving or flashing stimuli to be activated: They fire in response to transient stimulation.

Some visual cortex neurons have receptive fields in which an excitatory center is flanked by inhibitory regions. Inhibitory regions along the axis of orientation, a property known as *end-inhibition*, restrict a neuron's responses to lines of a certain length (Figure 27-5). End-inhibited neurons respond well to a line that does not extend into the inhibitory flanks but lies entirely within the excitatory part of the receptive field. Because the inhibitory regions share the orientation preference of the central excitatory region, end-inhibited cells are selective for line curvature and also respond well to corners.

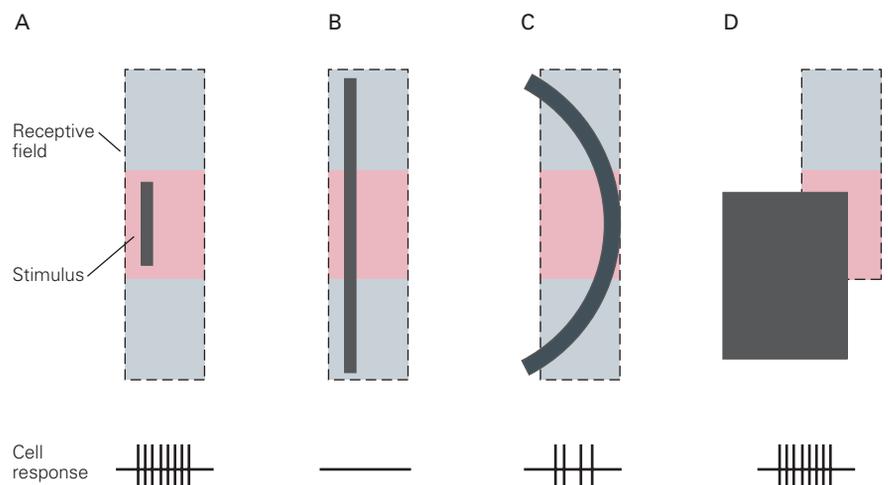
To define the shape of the object as a whole, the visual system must integrate the information on local

orientation and curvature into object contours. The way in which the visual system integrates contours reflects the geometrical relationships present in the natural world (Figure 27-6). As originally pointed out by Gestalt psychologists early in the 20th century, contours that are immediately recognizable tend to follow the rule of good continuation: Curved lines maintain a constant radius of curvature and straight lines stay straight. In a complex visual scene such smooth contours tend to "pop out," whereas more jagged contours are difficult to detect.

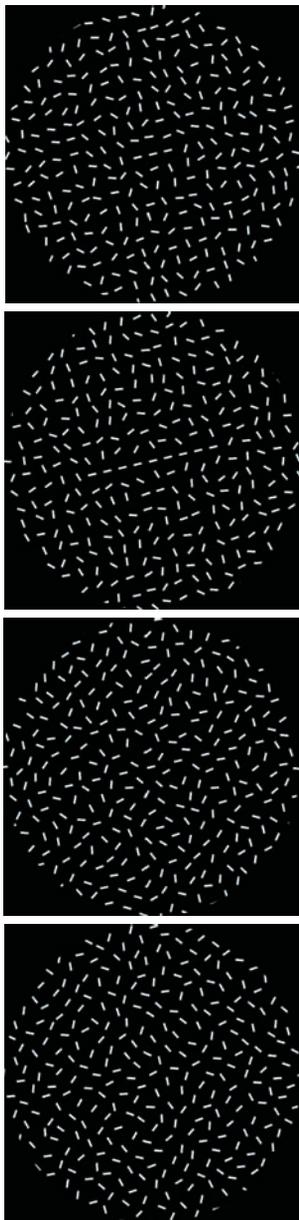
The responses of a visual cortex neuron can be modulated by stimuli that themselves do not activate the cell and therefore lie outside the receptive field's core. This *contextual modulation* endows a neuron with selectivity for more complex stimuli than would be predicted by placing the components of a stimulus at different positions in and around the receptive field. The same factors that facilitate the detection of an object in a complex scene (Figure 27-6A) also apply to contextual modulation. The properties of perceptible contours are reflected in the responses of neurons in the primary visual cortex, which are sensitive to the global characteristics of contours, even those that extend well outside their receptive fields.

Contextual influences over large regions of visual space are likely to be mediated by connections between multiple columns of neurons in the visual cortex that have similar orientation selectivity (Figure 27-6B). These connections are formed by pyramidal-cell axons that run parallel to the cortical surface (see Figure 25-16). The extent and orientation dependency of these horizontal connections provide the interactions that could mediate contour saliency (see Figure 25-14).

**Figure 27-5** End-inhibited receptive fields. Some receptive fields have a central excitatory region flanked by inhibitory regions that have the same orientation selectivity. Thus a short line segment or a long curved line will activate the neuron (A and C) but a long straight line will not (B). A neuron with a receptive field that displays only one inhibitory region in addition to the excitatory region can signal the presence of corners (D).



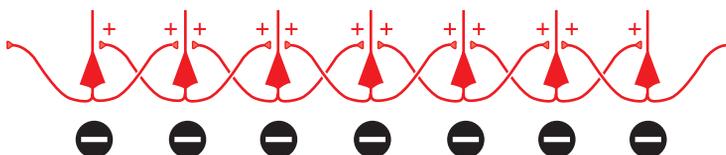
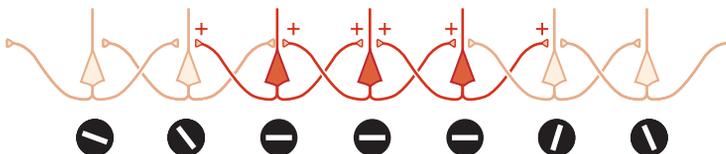
A Visual field



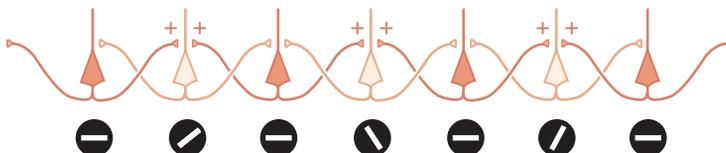
B Laterally connected V1 neurons

Features affecting contour saliency

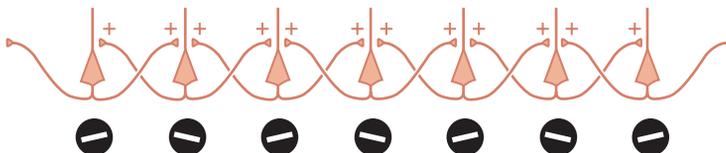
Number of line elements



Spacing of collinear line elements



Smoothness of contour



**Figure 27-6** Contour integration. (Adapted, with permission, from Li W and Gilbert CD 2002.)

A. Contour integration reflects the perceptual rules of proximity and good continuation. Each of the four images here has a straight line in the center, and all four lines have the same oblique orientation. In some images the line pops out more or less immediately, without searching. Factors that contribute to contour saliency include the number of contour elements (compare the first and second frames), the spacing of the elements (third frame), and the smoothness of the contour (bottom frame). When the spacing between contour elements is too

large or the orientation difference between them too great, one must search the image to find the contour.

B. These perceptual properties are reflected in the horizontal connections that connect columns of neurons in the primary visual cortex with similar orientation selectivity. As long as the contour elements are spaced sufficiently close together, excitation can propagate from cell to cell, thus facilitating the responses of V1 neurons. Each neuron in the network then augments the responses of neurons on either side and the facilitated responses propagate across the network.

## Depth Perception Helps Segregate Objects from Background

Depth is another key feature in determining the shape of an object. An important cue for the perception of depth is the difference between the two eyes' views of the world, which must be computed and reconciled by the brain. The integration of binocular input begins in the primary visual cortex, the first level at which individual neurons receive signals from both eyes. The balance of input from the two eyes, a property known as ocular dominance, varies among cells in V1.

These neurons are also selective for depth, which is computed from the relative retinal positions of objects placed at different distances from the observer. An object that lies in the *plane of fixation* produces images at corresponding positions on the two retinas (Figure 27-7). The images of objects that lie in front of or behind the plane of fixation fall on slightly different locations in the two eyes. Individual visual cortex neurons are selective for a narrow range of such disparities. Some are selective for objects lying on the plane of fixation (tuned excitatory or inhibitory cells), whereas others respond only when objects lie in front of the plane of fixation (near cells) or behind that plane (far cells).

Depth plays an important role in the perception of object shape, in surface segmentation, and in establishing the three-dimensional properties of a scene. Objects that are placed near an observer can partially occlude those situated farther away. A surface passing behind an object is perceived as continuous even though its two-dimensional image on each retina represents two surfaces separated by the occluder. When the brain encounters a surface interrupted by gaps displaying appropriate alignment and contrast and lying in the near-depth plane, it fills in the gaps to create a continuous surface (Figure 27-8).

Although the depth of a single object can be established easily, determining the depths of multiple objects within a scene is a much more complex problem that requires linking the retinal images of all objects in the two eyes. The disparity calculation is therefore a global one: The calculation in one part of the visual image influences the calculation for other parts. When the assignment of depth is unambiguous in one part of an image, that information is applied to other parts of the image where there is insufficient information to determine depth, a phenomenon known as disparity capture.

Random-dot stereograms provide a dramatic demonstration of the global nature of disparity analysis. The image presented to each eye appears as noise, but when the images are viewed binocularly the disparity between the random array of dots in the two images

allows an embedded shape to become visible (Figure 27-8C). The calculation underlying this percept is not simple, but requires determining which features shown to the left eye correspond to features seen by the right eye and propagating local disparity information across the image.

Neurons in area V2 display sensitivity to global disparity cues. Even when no contrast boundary exists in a neuron's receptive field, the neuron will respond to illusory contours formed by adjacent line elements (Figure 27-8B). The neuron's response is facilitated when collinear lines appear inside or outside the receptive field. When a perpendicular bar occludes the lines, indicating a break between them, the facilitation disappears. But when the bar is moved to a plane nearer than that of the collinear lines, as would occur if the lines were connected behind the occluder, the facilitation returns.

In addition to binocular disparity, the visual system also uses many monocular cues to discriminate depth. Depth determination through monocular cues, such as size, perspective, occlusion, brightness, and movement, is not difficult. Another cue that originates outside the visual system is vergence, the angle between the optical axes of the two eyes for objects at varying distances. Yet another binocular cue, known as DaVinci stereopsis, is the presence of features visible to one eye but occluded in the other eye's view.

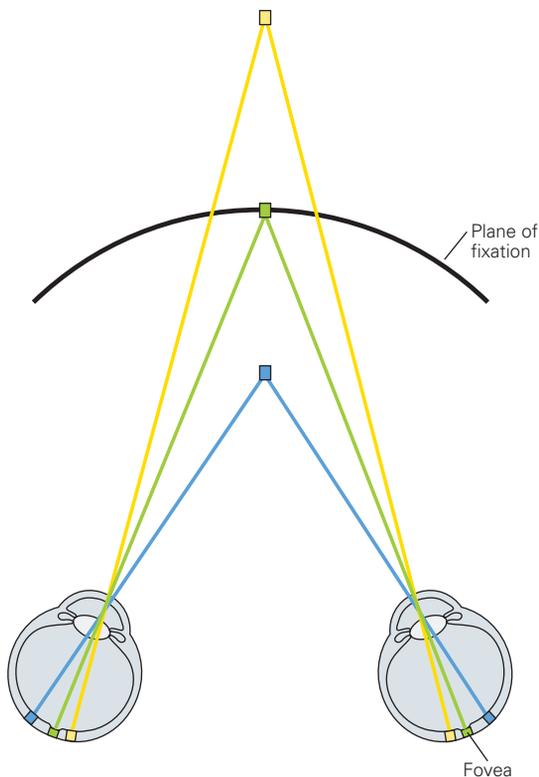
Neurons in areas V1 and V2 also signal foreground-background relationships. A cell with its receptive field in the center of a textured field may respond even when the boundary of that field is distant from the receptive field. This response helps differentiate the object from its background. In parsing an image the brain must identify which edge belongs to which object and differentiate the edge from the background of the object. Some cells in area V2 have the property of "border ownership," firing only when a figure but not the background is to one side of the edge, even when the local edge information is identical in both instances (Figure 27-9).

## Local Movement Cues Define Object Trajectory and Shape

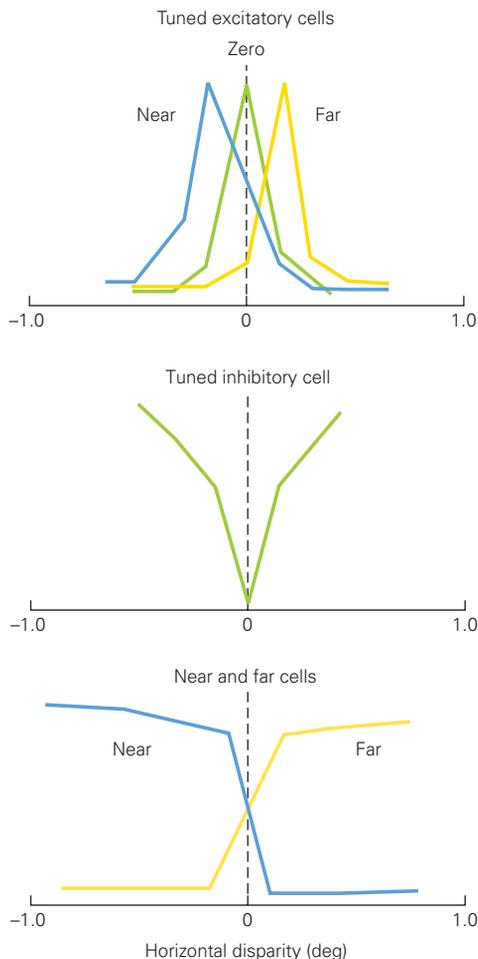
The primary visual cortex determines the direction of movement of objects. Directional selectivity in neurons likely involves sequential activation of regions on different sides of the receptive field.

If an object moving at an appropriate velocity first encounters a region of a neuron's receptive field with long response latencies and then passes into regions with progressively shorter latencies, signals from throughout

A Binocular disparity of retinal images



B Disparity-selective neurons



**Figure 27-7** Stereopsis and binocular disparity.

A. Depth is computed from the positions at which images occur in the two eyes. The image of an object lying in the plane of fixation (**green**) falls on corresponding points on the two retinas. The images of objects lying in front of the plane of fixation (**blue**) or behind it (**yellow**) fall on noncorresponding locations on the two retinas, a phenomenon termed *binocular disparity*.

B. Visual cortex neurons are selective for particular ranges of disparity. Each plot shows the responses of a neuron to

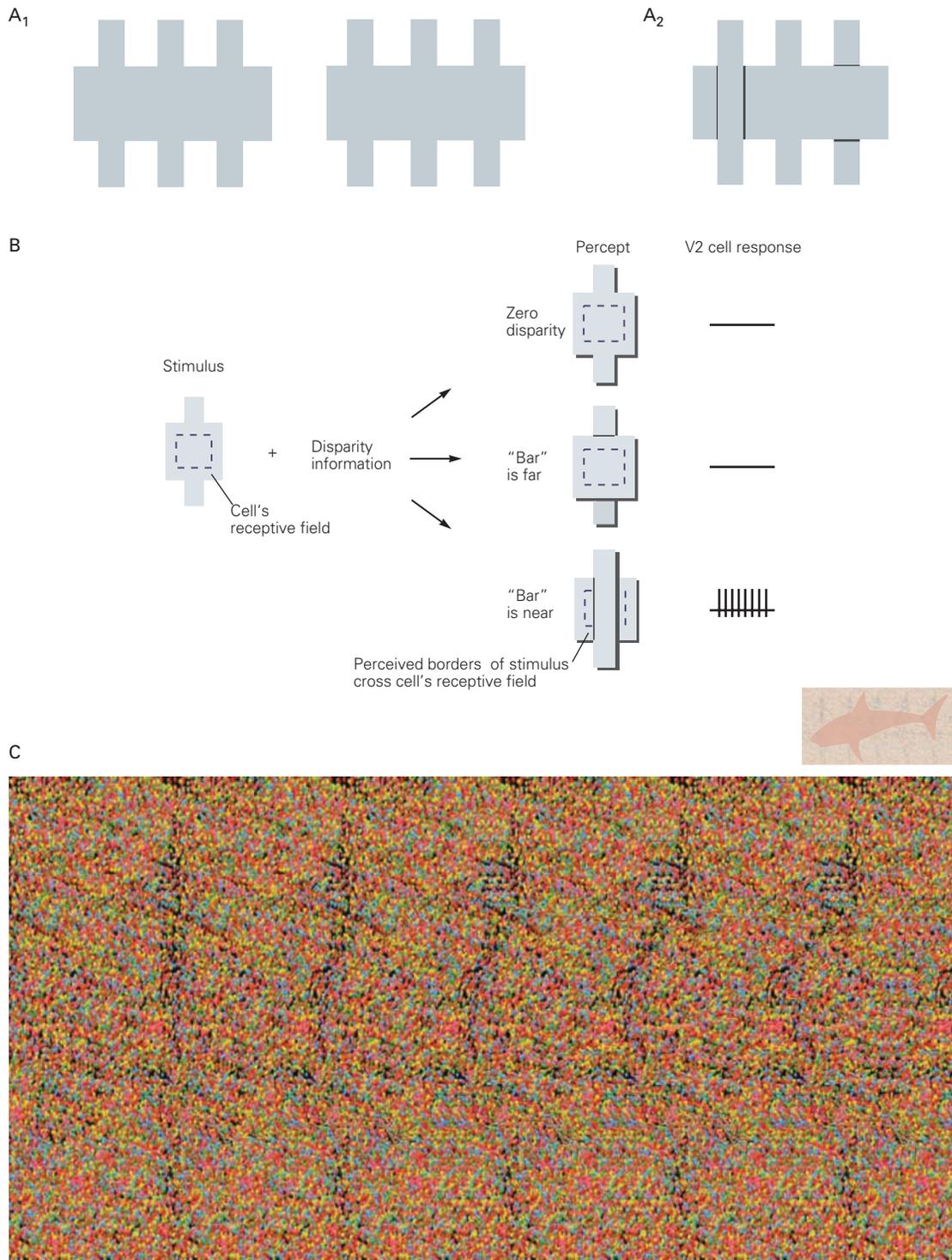
binocular stimuli with different disparities (abscissa). Some neurons are tuned to a narrow range of disparities and thus have particular disparity preferences (tuned excitatory or tuned inhibitory neurons), whereas others are tuned broadly for objects in front of the fixation plane (near cells) or beyond the plane (far cells). (Adapted, with permission, from Poggio 1995.)

the receptive field will arrive at the cell simultaneously and the neuron will fire vigorously. If the object instead moves in the opposite direction, signals from the different regions will not summate and the cell may never reach the threshold for firing (Figure 27-10).

Early in the visual pathways analysis of the movement of an object is limited by the size of the receptive fields of the sensory neurons. Even in the initial cortical areas V1 and V2 the receptive fields of neurons

are small and might encompass only a fraction of an object. Eventually, however, information about the direction and speed of movement of discrete aspects of an object must be integrated into a computation of the movement of a whole object. This problem is more difficult than one might expect.

If one observes a complex shape moving through a small aperture, the part of the object’s boundary within the aperture appears to move in a direction



**Figure 27-8** Global analysis of binocular disparity.

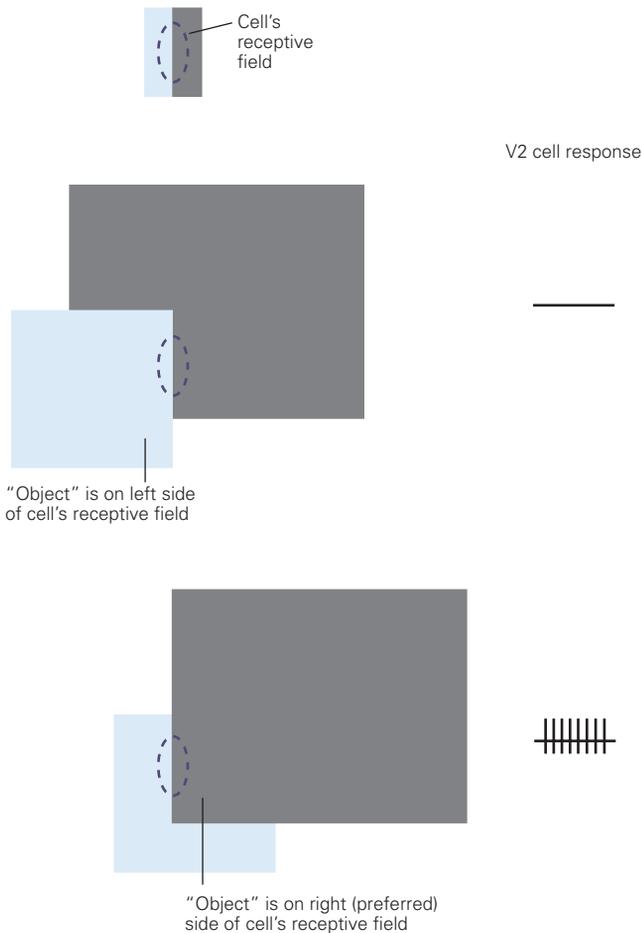
**A.** Depth cues contribute to surface segmentation. Viewing a single image of three gray vertical bars crossing a gray horizontal rectangle, you see a uniform gray area within the rectangle. However, if you fuse the two rectangles in  $A_1$  with diverged eyes, the three vertical bars fall on the two retinas with near, zero, and far disparity, respectively, as portrayed in  $A_2$ . Thus the bar at left appears to hover in front of the rectangle with an illusory vertical edge crossing the rectangle, whereas the bar at right appears to lie behind the edges of the horizontal rectangle.

**B.** A neuron in area V2 responds to illusory edges formed by binocular disparity cues. When the cell's receptive field

is centered in the gray square, the cell does not respond to a vertical bar that has far disparity or the same disparity as the square. When the vertical bar has near disparity, the cell responds as the illusory vertical edge crosses its receptive field. (Reproduced, with permission, from Bakin, Nakayama, and Gilbert 2000.)

**C.** A random-dot stereogram is seen as a random array of colored dots until one diverges or converges the eyes to bring the adjacent dark vertical stripes into register, producing a three-dimensional image of a shark hovering in front of the background noise. This effect stems from systematic disparity for selected sets of dots. (© Fred Hsu/Wikimedia Commons/CC-BY-SA-3.0.)

perpendicular to the boundary's orientation (Figure 27-11A). One cannot detect a line's true direction of movement if the line's ends are not visible. The image of a line appears the same if it is moving slowly along an axis perpendicular to its orientation or more quickly along an oblique axis. This is the quandary presented by the receptive field of a V1 neuron. The visual system's solution is to assume that the movement of a contour is perpendicular to its orientation. Thus an object is first presented to the visual system in countless small pieces with boundaries of different orientations, all of which appear to be moving in different directions and at different velocities (Figure 27-11A).



**Figure 27-9** Border ownership. Cells in area V2 are sensitive to the boundaries of whole objects. Even though the local contrast is the same for the two rectangles within a cell's receptive field, the cell responds only when the boundary is part of a complete surface that lies on the preferred side of the receptive field. (Adapted, with permission, from Zhou, Friedman, and von der Heydt 2000.)

Determining the direction of motion of an object requires resolving multiple cues. This can be demonstrated readily by placing one grating on top of another and moving the two in different directions. The resulting checkerboard pattern appears to move in an intermediate direction between the trajectories of the individual gratings (Figure 27-11B). This percept depends on the relative contrast of the gratings and the area of grating overlap. With large relative contrasts the gratings appear to slide across each other, moving in their individual directions rather than together in a common direction.

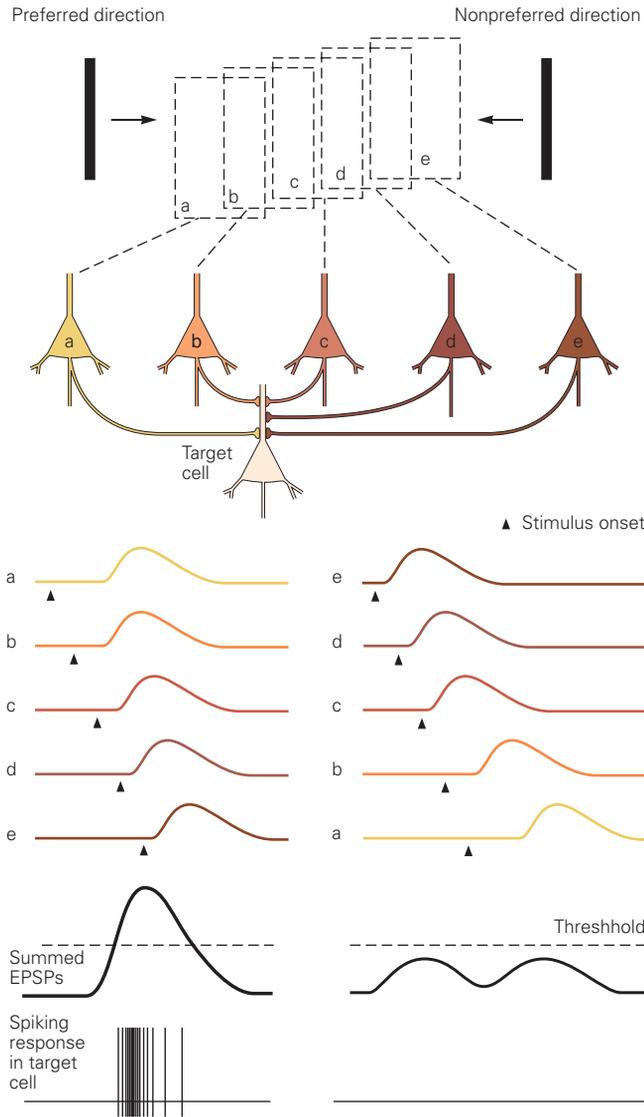
An important determinant of perceived direction is scene segmentation, the separation of moving elements into foreground and background. In a scene with moving objects segmentation is not based on local cues of direction; instead, perception of direction depends on scene segmentation. The barber-pole illusion provides another example of the predominance of global relationships over the perception of simple attributes. The rotating stripes are perceived as moving vertically along the long axis of the pole (Figure 27-11C). The perception of motion in the visual field uses a complex algorithm that integrates the bottom-up analysis of local motion signals with top-down scene segmentation.

Integration of these local motion signals in monkeys has been observed in the middle temporal area (area MT or V5), an area specializing in motion. Remarkably this neural integration mirrors the perceptual effects. The neurons are selective for a particular direction of movement of an overall pattern, rather than to the motion of individual components of the pattern. Their responses also depend on transparency and display the barber-pole effect, sensitivity to the shape and dimensions of the aperture within which the movement is seen.

### Context Determines the Perception of Visual Stimuli

#### Brightness and Color Perception Depend on Context

The visual system attempts to measure the surface characteristics of objects by comparing the light arriving from different parts of the visual field. As a result, the perception of brightness and color is highly dependent on context. In fact, perceived brightness and color can be quite different from what is expected from the physical properties of an object. At the same time, perceptual constancies make objects appear similar even when the brightness and wavelength distribution of



**Figure 27-10** Directional selectivity of movement. The selectivity of a neuron to the direction of movement depends on the response latencies of presynaptic neurons. The response latencies of presynaptic neurons *a* and *b* relative to the onset of a stimulus are somewhat longer than those of neurons *d* and *e*. When a stimulus moves from left to right, neurons *a* and then *b* are activated first, but because their responses are delayed their inputs arrive simultaneously with inputs from neurons *d* and *e* and therefore sum at the target neuron, causing it to fire. In contrast, stimuli moving leftward produce responses that arrive at different times and therefore do not reach threshold. (EPSP, excitatory postsynaptic potential.) (Adapted, with permission, from Priebe and Ferster 2008.)

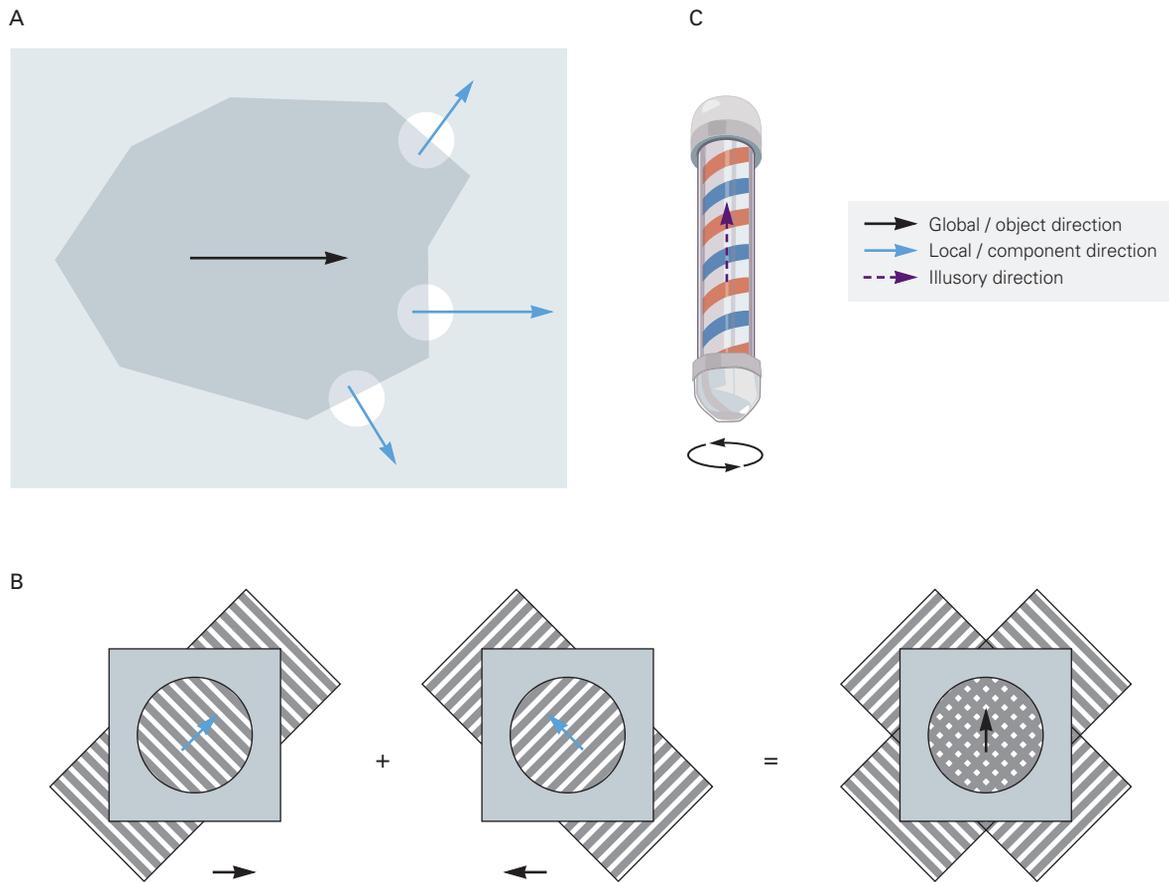
the light that illuminates them changes from natural to artificial light, from sunlight to shadow, or from dawn to midday (Figure 27-12A).

As we move about or as the ambient illumination changes, the retinal image of an object—its size, shape, and brightness—also changes. Yet under most conditions we do not perceive the object itself to be changing. As we move from a brightly lit garden into a dimly lit room, the intensity of light reaching the retina may vary a thousandfold. Both in the room’s dim illumination and in the sun’s glare we nevertheless see a white shirt as white and a red tie as red. Likewise, as a friend walks toward you she is seen as coming closer; you do not perceive her to be growing larger even though the image on your retina does expand. Our ability to perceive an object’s size and color as constant illustrates again the fundamental principle of the visual system: It does not record images passively, like a camera, but instead uses transient and variable stimulation of the retina to construct representations of a stable, three-dimensional world.

Another example of contextual influence is color induction, whereby the appearance of a color in one region shifts toward that in an adjoining region. Shape also plays an important role in the perception of surface brightness. Because the visual system assumes that illumination comes from above, gray patches on a folded surface appear very different when they lie on the top or bottom of the surface, even when they are in fact the same shade of gray (Figure 27-12B).

The responses of some neurons in the visual cortex correlate with perceived brightness. Most visual neurons respond to surface boundaries; the center-surround structure of the receptive fields of retinal ganglion cells and geniculate neurons is suited to capturing boundaries. Most such cells do not respond to the interior parts of surfaces, for uniform interiors produce no contrast gradients across receptive fields. However, a small percentage of neurons do respond to the interiors of surfaces, signaling local brightness, texture, or color. Their responses are influenced by context: As the brightness of surfaces outside a cell’s receptive field change, the cell’s response changes, even when the brightness of the surface within the receptive field remains fixed.

Because most neurons respond to surface boundaries and not to areas of uniform brightness, the visual system calculates the brightness of surfaces from information about contrast at the edges of surfaces. The brain’s analysis of surface qualities from boundary information is known as perceptual fill-in. If one fixates the boundary between a dark disk and a surrounding bright area for a few seconds, the disk will “fill in” with the same brightness as the surrounding area.



**Figure 27-11** The aperture problem and barber-pole illusion.

**A.** Although an object moves in one direction, each component edge when viewed through a small aperture appears to move in a direction perpendicular to its orientation. The visual system must integrate such local motion signals into a unified percept of a moving object.

**B.** Gratings are used to test whether a neuron is sensitive to local or global motion signals. When the gratings are superimposed and moved independently in different directions, one

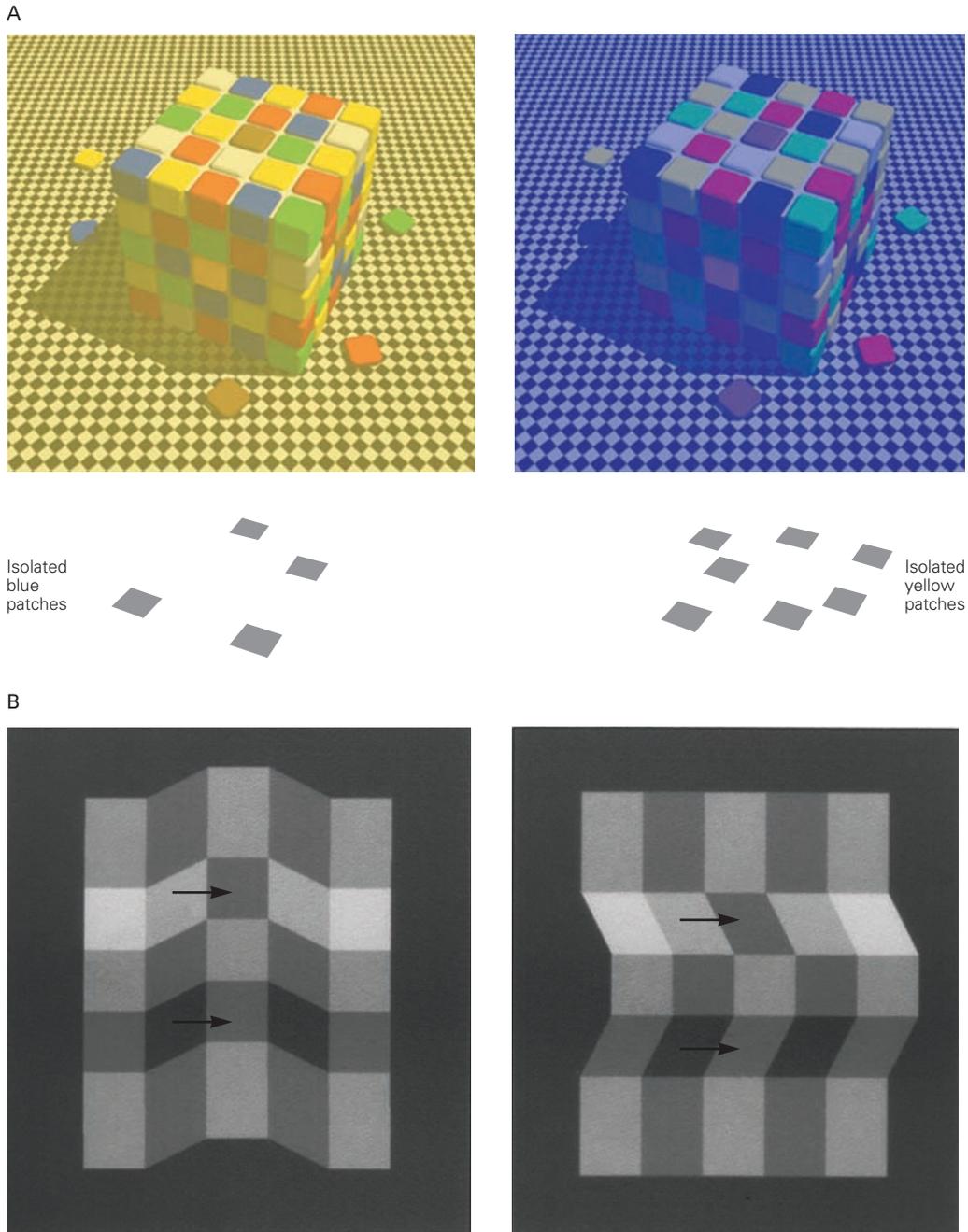
does not see the two gratings sliding past each other but rather a plaid pattern moving in a single, intermediate direction. Neurons in the middle temporal area of monkeys are responsive to such global motion rather than to local motion.

**C.** Motion perception is influenced by surrounding segmentation cues, as seen in the barber-pole illusion. Even though the pole rotates around its axis, one perceives the stripes as moving vertically.

This occurs because the cells that respond to edges fire only when the eye or stimulus moves. They gradually cease to respond to a stabilized image and no longer signal the presence of the boundary. Neurons with receptive fields within the disk gradually begin to respond in a fashion similar to those with receptive fields in the surrounding area, demonstrating short-term plasticity in their receptive-field properties.

An object's color always appears more or less the same despite the fact that under different conditions of illumination the wavelength distribution of light reflected from the object varies widely. To identify an object we must know the properties of its surface

rather than those of the reflected light, which are constantly changing. Computation of an object's color is therefore more complex than analyzing the spectrum of reflected light. To determine a surface's color the wavelength distribution of the incident light must be determined. In the absence of that information surface color can be estimated by determining the balance of wavelengths coming from different surfaces in a scene. Some neurons in V4 respond similarly to different illumination wavelengths if the perceived color remains constant. By being responsive to the light across an extensive surface, these neurons are selective for surface color rather than wavelength.



**Figure 27-12** Color and brightness perception depend on contextual cues.

A. The perception of surface color remains relatively stable under different illumination conditions and the consequent changes in the wavelengths of light reflected from the surface. The yellow squares on the left and right cubes appear similar despite the fact that the wavelengths of light coming from the two sets of surfaces are very different. In fact, if the blue squares on the top of the left cube and the yellow squares on the top of the right cube are isolated from their contextual squares, their colors appear identical. (Reproduced, with permission, from R. Beau Lotto at [www.lottolab.org](http://www.lottolab.org).)

B. Brightness perception is also influenced by three-dimensional shape. The four gray squares indicated by arrows all have the same luminance. In the left illustration the apparent brightnesses are similar. At the right, however, the apparent brightnesses are different. The visual system has an inherent expectation that illumination comes from above (the position of the sun relative to us), which leads to the perception that the surface below the fold appears brighter than the surface of the same luminance that lies above. (Reproduced, with permission, from Adelson 1993.)

### Receptive-Field Properties Depend on Context

The distinction between local and global effects—between stimuli that occur within a receptive field and those beyond—poses the problem of how the receptive field itself is defined. Because the original characterization of the receptive fields of visual cortex neurons did not take into account contextual influences, some investigators now distinguish between “classical” and “nonclassical” receptive fields.

However, even the earliest description of the sensory receptive field allowed for the possibility of influences from portions of the sensory surface outside the narrowly defined receptive field. In 1953 Steven Kuffler, in his pioneering observations on the receptive-field properties of retinal ganglion cells, noted that “not only the areas from which responses can actually be set up by retinal illumination may be included in a definition of the receptive field but also all areas which show a functional connection, by an inhibitory or excitatory effect on a ganglion cell. This may well involve areas which are somewhat remote from a ganglion cell and by themselves do not set up discharges.”

A more useful distinction contrasts the response of a neuron to a simple stimulus, such as a short line segment, with its response to a stimulus with multiple components. Even in the primary visual cortex neurons are highly nonlinear; their response to a complex stimulus cannot be predicted from their responses to a simple stimulus placed in different positions around the visual field. Their responses to local features are instead dependent on the global context within which the features are embedded. Contextual influences are pervasive in intermediate-level visual processing, including contour integration, scene segmentation, and the determination of object shape and surface properties.

### Cortical Connections, Functional Architecture, and Perception Are Intimately Related

Intermediate-level visual processing requires sharing of information from throughout the visual field. The interconnections within the primary visual cortex and the relationship of these connections to the functional architecture of this area suggest that they mediate contour integration.

Cortical circuits include a plexus of long-range horizontal connections, running parallel to the cortical surface, formed by the axons of pyramidal neurons. Horizontal connections exist in every area of the cerebral cortex, but their function varies from one area to

the next depending on the functional architecture of each area. In the visual cortex these connections mediate interactions between orientation columns of similar specificity thus integrating information over a large area of visual cortex that represents a great expanse of the visual field (see Figure 25–14).

The combination of this like-to-like rule of connections and the fact that the horizontal connections link distant locations in the visual field suggest these connections have a role in contour integration. Contour integration and the related property of contour saliency reflect the Gestalt principle of good continuation. Contour integration and saliency are mediated by the horizontal connections in V1 (see Figure 27–6).

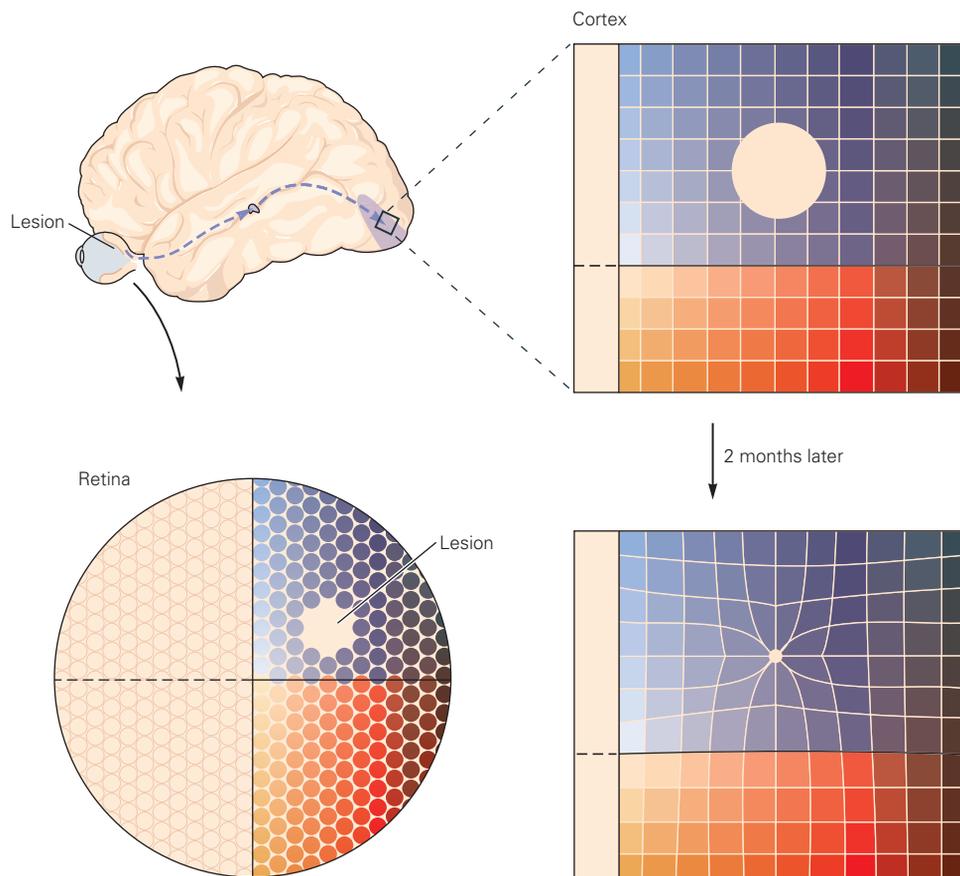
A final feature of cortical connectivity important for visuospatial integration is feedback projection from higher-order cortical areas. Feedback connections are as extensive as the feed-forward connections that originate in the thalamus or at earlier stages of cortical processing. Little is known about the function of these feedback projections. They likely play a role in mediating the top-down influences of attention, expectation, and perceptual task, all of which are known to affect early stages in cortical processing.

### Perceptual Learning Requires Plasticity in Cortical Connections

The synaptic connections in ocular-dominance columns are adaptable to experience only during a critical period in development (see Chapter 57). This suggests that the functional properties of visual cortex neurons are fixed in adulthood. Nevertheless, many properties of cortical neurons remain mutable throughout life. For example, changes in the visual cortex can occur following retinal lesions.

When focal lesions occur in corresponding positions on the two retinas, the corresponding part of the cortical map, referred to as the lesion projection zone, is initially deprived of visual input. Over a period of several months, however, the receptive fields of cells within this region shift from the lesioned part of the retina to the functioning area surrounding the lesion. As a result, the cortical representation of the lesioned part of the retina shrinks while that of the surrounding region expands (Figure 27–13).

The plasticity of cortical maps and connections did not evolve as a response to lesions. Instead, plasticity is the neural mechanism for improving our perceptual skills. Many of the attributes analyzed by the visual cortex, including stereoscopic acuity, direction of movement, and orientation, become sharper with practice. Hermann von Helmholtz stated in 1866 that



**Figure 27-13** Adult cortical plasticity. When corresponding positions in both eyes are lesioned, the cortical area receiving input from the lesioned areas—the lesion projection zone—is initially silenced. The receptive fields of neurons in the lesion projection zone eventually shift from the area of the lesion to

the surrounding, intact retina. This occurs because neurons surrounding the lesion projection zone sprout collaterals that form synaptic connections with neurons inside the zone. As a result, the cortical representation of the lesioned part of the retina shrinks while that of the surrounding retina expands.

“the judgment of the senses may be modified by experience and by training derived under various circumstances, and may be adapted to the new conditions. Thus, persons may learn in some measure to utilize details of the sensation which otherwise would escape notice and not contribute to obtaining any idea of the object.” This perceptual learning is a variety of implicit learning that does not involve conscious processes (see Chapter 65).

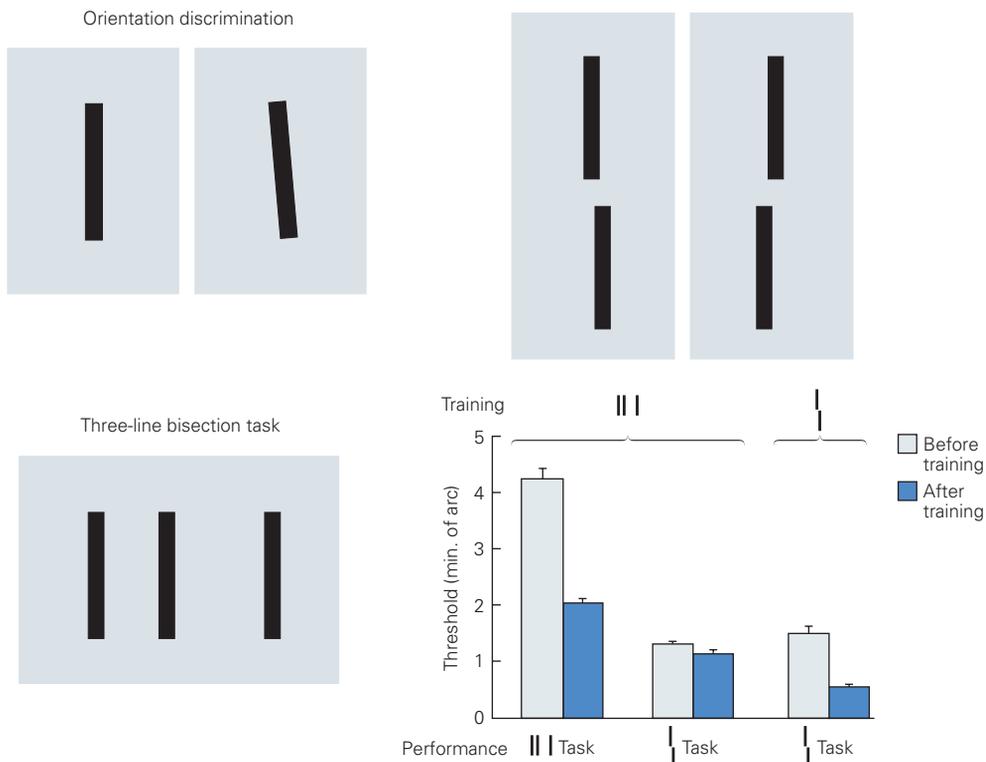
Perceptual learning involves repeating a discrimination task many times and does not require error feedback to improve performance. Improvement manifests itself as a decrease in the threshold for discriminating small differences in the attributes of a target stimulus or in the ability to detect a target in a complex environment. Several areas of visual cortex,

including the primary visual cortex, participate in perceptual learning.

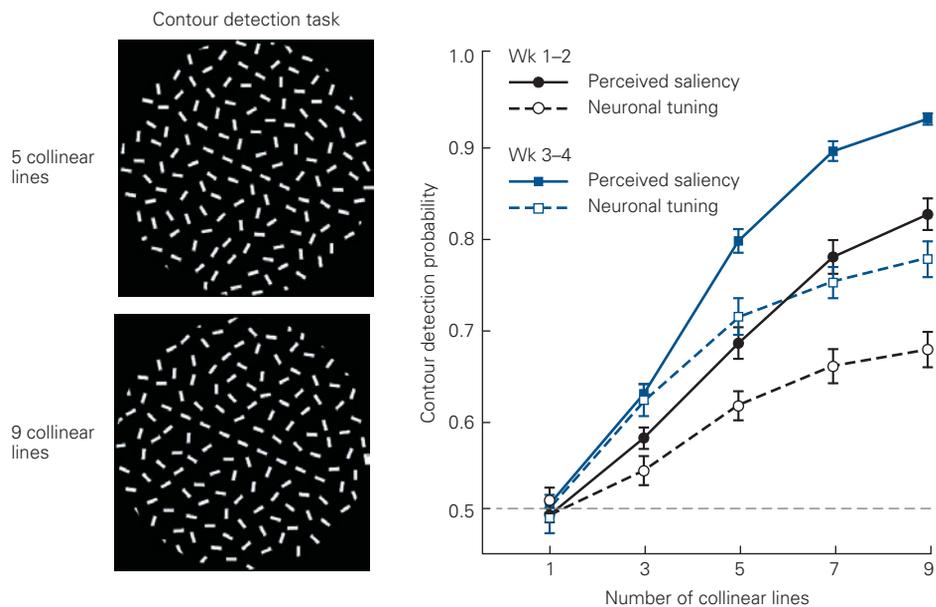
An important aspect of perceptual learning is its specificity: Training on one task does not transfer to other tasks. For example, in a three-line bisection task the subject must determine whether the centermost of three parallel lines is closer to the line on the left or the one on the right. The amount of offset from the central position required for accurate responses decreases substantially after repeated practice (Figure 27-14A).

The learning in this task is specific to the location in the visual field and to the orientation of the lines. This specificity suggests that early stages of visual processing are responsible, for in the early stages receptive fields are smallest, visuotopic maps are most precise, and orientation tuning is sharpest. The learning is also

A Perceptual learning is task-specific



B Neuronal responsiveness changes during training



**Figure 27-14** Perceptual learning. Perceptual learning is a form of implicit learning. With practice one can learn to discriminate smaller differences in orientation, position, depth, and direction of movement of objects.

A. The improvement is seen as a reduction in the amount of change required to reliably detect a tilted line or one positioned to the left or right of a nearly collinear line (vernier task). Perceptual learning is highly specific, so that training on a three-line bisection task leads to substantial improvement in that task (left pair of bars in the bar graph) without affecting performance on the vernier discrimination task (central pair of bars).

However, training specifically on vernier discrimination does enhance performance on that task (right pair of bars).

B. The responses of neurons in V1 parallel perceptual learning. Subjects can detect collinear line segments embedded in a random background more easily as the number of segments is increased. The responses of neurons in V1 grow correspondingly stronger with the increase in the number of line segments. After practice, a line with fewer segments stands out more easily, and with this improvement the responses in V1 also increase. (Reproduced, with permission, from Crist et al. 2001; and Li et al. 2008.)

specific for the stimulus configuration. Training on three-line bisection does not transfer to a vernier discrimination task in which the context is a line that is collinear with the target line (see Figure 27–14A).

The response properties of neurons in the primary visual cortex change during the course of perceptual learning in a way that tracks the perceptual improvement. An example of this is seen in contour saliency. With practice, subjects can more easily detect contours embedded in complex backgrounds. Detection improves with contour length, and the responses of neurons in V1 increase as well. With practice, subjects improve their ability to detect shorter contours and V1 neurons become correspondingly more sensitive to shorter contours (see Figure 27–14B).

**Visual Search Relies on the Cortical Representation of Visual Attributes and Shapes**

The detectability of features such as color, orientation, and shape is related to the process of visual search. Certain objects emerge or “pop out” from others in a complex image because the visual system processes

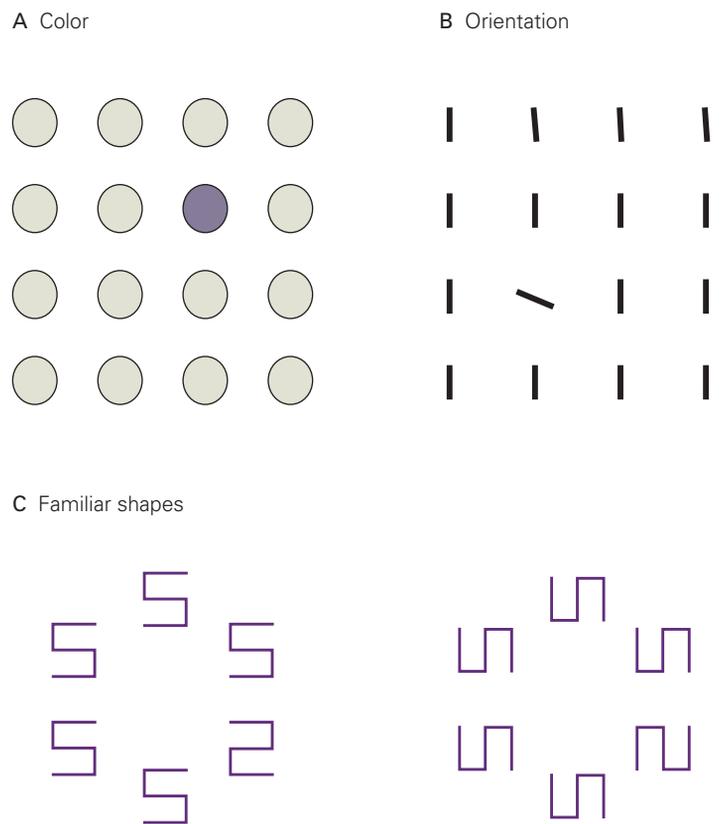
simultaneously, in parallel pathways, the features of the target and the surrounding distractors (Figure 27–15). When the features of a target are complex, the target can be identified only through careful inspection of an entire image or scene (see Figure 21–5).

The pop-out phenomenon can be influenced by training. A stimulus that initially cannot be found without effortful searching will pop out after training. The neuronal correlate of such a dramatic change is not certain. Parallel processing of the features of an object and its background is possible because feature information is encoded within retinotopically mapped areas at multiple locations in the visual cortex. Pop-out probably occurs early in the visual cortex. The pop-out of complex shapes such as numerals lends support to the idea that early in visual processing neurons can represent, and be selective for, shapes more complex than line segments with a particular orientation.

**Cognitive Processes Influence Visual Perception**

Scene segmentation—the parsing of a scene into different objects—involves a combination of bottom-up

**Figure 27-15** One object in a complex image stands out under certain conditions.  
 A. A differently colored object pops out.  
 B. A differently oriented line also pops out.  
 C. More complex shapes can pop out when they are very familiar, such as the numeral 2 embedded in a field of 5s. Rotating the image by 90° renders the elements of the figure less recognizable, making it more difficult to find the one figure that differs from the rest. (Reproduced, with permission, from Wang, Cavanagh, and Green 1994.)



processes that follow the Gestalt rule of good continuation and top-down processes that create object expectation.

One strong top-down influence is spatial attention, which can change focus without any movement of an observer's eyes. Spatial attention is object-oriented in that it is distributed over the area occupied by the attended object, allowing the visual cortex to analyze the shape and attributes of objects one at a time.

Attentional mechanisms can solve the superposition problem. For us to recognize an object in a scene that includes multiple objects, we must determine which features correspond to which objects. Our sense that we identify multiple objects simultaneously is illusory. Instead, we serially process objects in rapid succession by shifting attention from one to the next. The results of each analysis build up the perception of a complex environment populated with many distinct objects. A dramatic demonstration of the importance of attention in object recognition is *change blindness*. If a subject rapidly shifts between two slightly different views of the same scene, he will not be able to detect the absence of an important component of the scene in one view without considerable scrutiny (see Figure 29–3).

Another top-down influence is perceptual task. At early stages in visual processing the properties of the same neuron vary with the type of visual discrimination being performed. Object identification itself involves a process of hypothesis testing in which internal representations of objects are compared with information arriving from the retina. This process is reflected in studies of visual imagery: Early stages in processing such as the primary visual cortex are activated when one imagines scenes in the absence of visual input.

## An Overall View

Intermediate-level visual processing is concerned with parsing the visual world into contours and surfaces that belong to objects and segregating these elements from the background. This is the most challenging job that the visual system must perform. When confronted with a complex visual environment, we could assemble local features into a potentially enormous number of distinct objects. Nonetheless, we quickly classify the local features into a set of objects that can be matched with internal representations of object shape and identity that are stored in the brain from earlier experiences.

This global integration is simplified by applying rules of perceptual grouping that were described by

Gestalt psychologists and are apparently implemented by circuits beginning in the primary visual cortex. Global integration involves analysis of local attributes that depends on the properties of sensory neurons: Selectivity for local orientation supports the analysis of extended contours, directional sensitivity underlies the determination of object motion, disparity selectivity implements global stereopsis, and contrast sensitivity mediates color constancy. The process of integration is not simply a bottom-up one but is influenced by information arriving from higher-order areas of the visual cortex. Attention, expectation, and perceptual task influence how we segment the visual world.

Intermediate-level vision is a product of lateral connections between functional columns of neurons in a cortical area and the convergence of feed-forward signals with feedback information from higher-order areas. Vision therefore is not simply a feed-forward mechanism that assembles shapes in stages with increasing complexity. The underlying processes are highly dynamic on short time scales. The strategies that we use to interpret visual scenes also involve experience-dependent changes in the cortical circuits in which we constantly store information about shapes that we experience throughout life.

---

Charles D. Gilbert

## Selected Readings

- Albright TD, Stoner GR. 2002. Contextual influences on visual processing. *Annu Rev Neurosci* 25:339–379.
- Gilbert CD, Sigman M. 2007. Brain states: top-down influences in sensory processing. *Neuron* 54:677–696.
- Gilbert CD, Sigman M, Crist R. 2001. The neural basis of perceptual learning. *Neuron* 31:681–697.
- Li W, Piech V, Gilbert CD. 2004. Perceptual learning and top-down influences in primary visual cortex. *Nat Neurosci* 7:651–657.
- Li W, Piech V, Gilbert CD. 2006. Contour saliency in primary visual cortex. *Neuron* 50:951–962.
- Priebe NJ, Ferster D. 2008. Inhibition, spike threshold, and stimulus selectivity in primary visual cortex. *Neuron* 57:482–497.

## References

- Adelson EH. 1993. Perceptual organization and the judgment of brightness. *Science* 262:2042–2044.

- Bakin JS, Nakayama K, Gilbert CD. 2000. Visual responses in monkey areas V1 and V2 to three-dimensional surface configurations. *J Neurosci* 20:8188–8198.
- Crist RE, Li W, Gilbert CD. 2001. Learning to see: experience and attention in primary visual cortex. *Nat Neurosci* 4:519–525.
- Cumming BG, DeAngelis GC. 2001. The physiology of stereopsis. *Annu Rev Neurosci* 24:203–238.
- Ferster D, Miller KD. 2000. Neural mechanisms of orientation selectivity in the visual cortex. *Annu Rev Neurosci* 23:441–471.
- He ZJ, Nakayama K. 1994. Apparent motion determined by surface layout not by disparity or three-dimensional distance. *Nature* 367:173–175.
- Hubel DH, Wiesel TN. 1968. Receptive fields and functional architecture of monkey striate cortex. *J Physiol* 195:215–243.
- Li W, Gilbert CD. 2002. Global contour saliency and local colinear interactions. *J Neurophysiol* 88:2846–56.
- Li W, Piech V, Gilbert CD. 2008. Learning to link visual contours. *Neuron* 57:442–451.
- Movshon JA, Adelson EH, Gizzi MS, Newsome WT. 1985. The analysis of moving visual patterns. In: C Chagas, R Gattass, CG Gross (eds.). *Study Group on Pattern Recognition Mechanisms* pp. 67–86, Vatican City: Pontificia Academia Scientiarum.
- Nakayama K. 1996. Binocular visual surface perception. *Proc Natl Acad Sci U S A* 93:634–639.
- Nakayama K, Joseph JS. 2000. Attention, pattern recognition and popout in visual search. In: R Parasuraman (ed.). *The Attentive Brain*. Cambridge, MA: MIT Press.
- Poggio GE. 1995. Mechanisms of stereopsis in monkey visual cortex. *Cereb Cortex* 5:193–204.
- Purves D, Lotto RB, Nundy S. 2002. Why we see what we do. *Am Sci* 90:236–243.
- Wang Q, Cavanagh P, Green M. 1994. Familiarity and popout in visual search. *Percept Psychophys* 56:495–500.
- Zhou H, Friedman HS, von der Heydt R. 2000. Coding of border ownership in monkey visual cortex. *J Neurosci* 20:6594–6611.