Wolfe, J. M., Kluender, K. R. & Levi, D. M. (2018) *Sensation & Perception, Fifth Edition*, NY: Sinauer, Ch. 6: Space Perception and Binocular Vision, p. 189-203

accommodation The process by which the eye changes its focus (in which the lens gets fatter as gaze is directed toward nearer objects).

convergence The ability of the two eyes to turn inward, often used in order to place the two images of a feature in the world on corresponding locations in the two retinal images (typically on the fovea of each eye). Convergence reduces the disparity of that feature to zero (or nearly zero).

divergence The ability of the two eyes to turn outward, often used in order to place the two images of a feature in the world on corresponding locations in the two retinal images (typically on the fovea of each eye). Divergence reduces the disparity of that feature to zero (or nearly zero).

Accommodation and Convergence

Like a camera, the eyes need to be focused to see objects at different distances clearly. As we learned in Chapter 2, the human eye focuses via a process called **accommodation**, in which the lens gets fatter as we direct our gaze toward nearer objects (see Figure 2.3). We also need to point our eyes differently to focus on objects at different distances. As the schematic eyeballs in **Figure 6.23** move from the red dot to the blue dot, they rotate inward—a process called **convergence** (Figure 6.23A); refocusing on the red dot would require rotation outward, which is known as **divergence** (Figure 6.23B).

Focus cues can in principle enable us to see depth because we see the image from various points across the pupil, and if we could monitor our state of accommodation and/or the extent to which our eyes were converged, we could use this information as a cue to the depth of the object we were trying to bring into focus: the more we have to converge and the more the lens has to bulge in order to focus

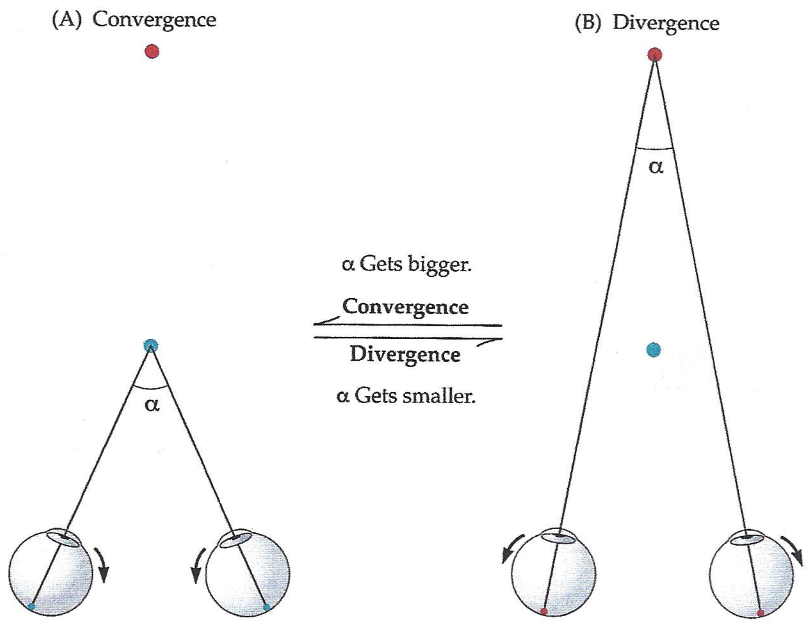

Figure 6.23 Vergence. (A) As we shift focus from a far to a near point, our eyes converge. (B) As we go from near to far, the eyes diverge. The size of the angle (labeled α) is a cue to depth.

on the object, the closer it is. In fact, we do use this information. For example, Hoffman and Banks (2010) showed that depth discrimination improves when the focus is correct. However, when we focus on objects more than about 2–3 meters away, the lens is as thin as it can get and the eyes are diverged about as much as possible, so neither cue provides much useful information. But careful studies have shown that the visual system takes advantage of both cues for objects closer than this limit. Convergence is used more than accommodation (Fisher and Ciuffreda, 1988; Owens, 1987). Moreover, in principle these cues can tell us the *exact* distance to an object. However, humans are not particularly precise about measuring the exact angles shown in Figure 6.23. Chameleons, on the other hand, do use the absolute metrical depth information from convergence to catch prey insects with their sticky tongues. Harkness (1977) showed this by fitting a chameleon with glasses that distorted the angle of convergence. The result was that the poor chameleon flicked out its tongue to the wrong distance and missed its intended dinner.

Binocular Vision and Stereopsis

As defined earlier, the term *binocular disparity* refers to differences between the images falling on our two retinas, and *stereopsis* refers to the impression of three-dimensionality—of objects "popping out in depth"—that most humans get when they view real-world objects with both eyes. Like the accounts of other depth cues, the story of the route from binocular disparity to stereopsis is a story of the visual system exploiting the regularities of projective geometry to recover the three-dimensional world from its projections—this time, onto a pair of two-dimensional surfaces. We will illustrate the translation from disparity to stereopsis using the situation shown in **Figure 6.24A**, in which the viewer (let's call him Bob) is facing a scene that includes four colored crayons at different depths. Suppose that Bob is focusing his gaze on the red crayon, as shown in **Figure 6.24B**. The two lines in this figure trace the paths of the light rays that reflect off the red crayon and onto Bob's two retinas. (Similar experiences with crayon scenes are also demonstrated in **Web Activity 6.2: Binocular Disparity**.)

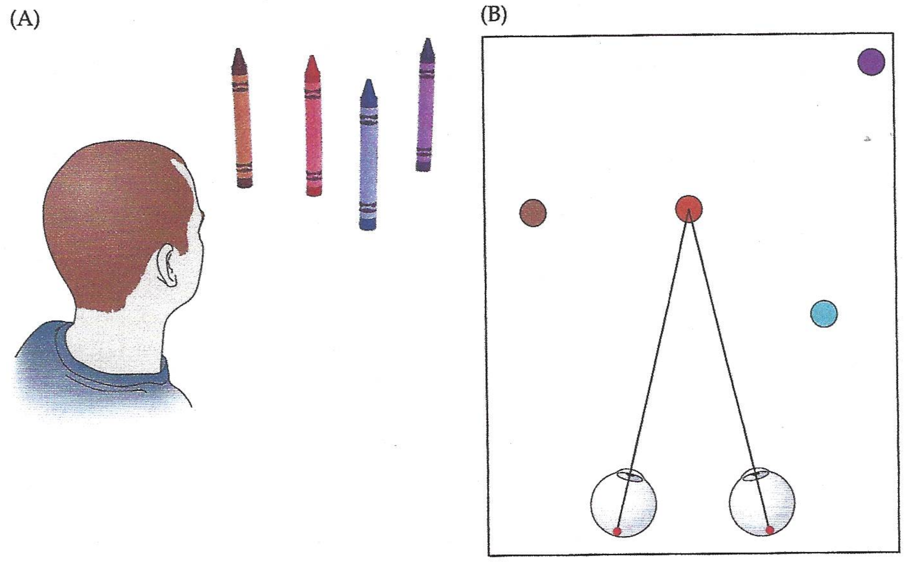

Figure 6.24 This simple visual scene illustrates how geometric regularities are exploited by the visual system to achieve stereopsis from binocular disparity. (A) The viewer, Bob, is assumed to be fixing his gaze on the red crayon. (B) This top view traces the rays of light bouncing off the red crayon onto Bob's retinas.

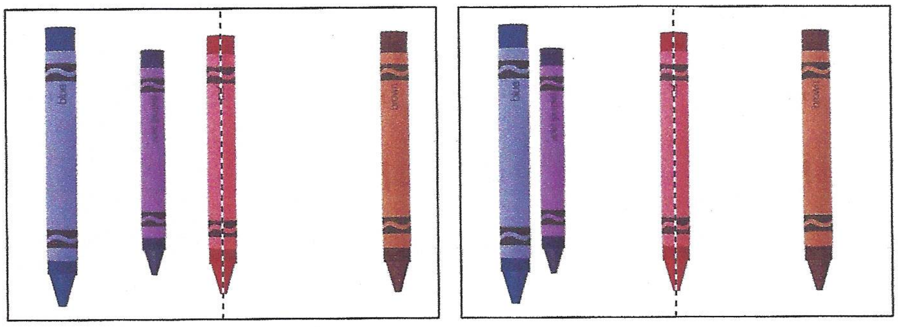

Figure 6.25 The overlapping portions of the images falling on Bob's left and right retinas. Because the retinal image is reversed, the blue and purple crayons on the right side of the scene in Figure 6.24 project to the left side of each retina, whereas the brown crayon on the left side of the scene projects to the right side of each retina. The size differences between the retinal images of the crayons in the two retinas are exaggerated in this figure compared with the differences we would observe if we saw this scene in the real world.

Because the visual system is designed so that the object of our gaze *always* falls on the fovea, the rays from the red crayon fall on the fovea in each of Bob's eyes. **Figure 6.25** shows the retinal image for the crayons in each eye. The red crayon is in the center of both images. We've added a dashed vertical line in front of this crayon in each image, to emphasize the fact that this is the location of the fovea.

Now consider the retinal images of the blue crayon. As you saw in Chapter 2, the optics of the eye reverse left-right and up-down (see Figure 2.2A). Thus, the blue crayon on the right side of the scene in Figure 6.24 falls on the left side in each of the two retinal images shown in Figure 6.25. In our imaginary scene, the blue crayon is placed so that the monocular retinal images of that crayon are formed at the same distance from the fovea in both eyes. We say that this crayon's images fall on **corresponding retinal points**. The same can be said of the images of the red crayon, which fall on the two foveas.

In fact, any object lying on the **Vieth-Müller circle**—the imaginary circle that runs through the two eyeballs and the object on which Bob is fixated—should project to corresponding retinal points. This imaginary circle is drawn in gray in **Figure 6.26**. Objects that fall on corresponding retinal points are said to have zero binocular disparity. If the two eyes are looking at one spot (such as the red crayon), then there will be a surface of zero disparity running through that spot. That surface is known as the **horopter**. Any object placed on that imaginary surface in the world will form images on corresponding retinal points. As it happens, the horopter and the Vieth-Müller circle are not quite the same. If you are *extremely* fond of rather complicated geometry, you may want to pursue this topic in one of the following sources: I. P. Howard and Rogers, 1995 or 2001; or Tyler, 1991. Otherwise, the important point is that there is a surface of zero disparity whose position in the world depends on the current state of convergence of the eyes.

Objects that lie on the horopter are seen as single objects when viewed with both eyes. Objects significantly closer to or farther away from the surface of zero disparity form images on decidedly noncorresponding points in the two eyes, and we see two of each of those objects. This double vision is known as **diplopia**. Objects that are close to the horopter but not quite on it can still be seen

corresponding retinal points Two monocular images of an object in the world are said to fall on corresponding points if those points are the same distance from the fovea in both eyes. The two foveas are also corresponding points.

Vieth-Müller circle The location of objects whose images fall on geometrically corresponding points in the two retinas. If life were simple, this circle would be the horopter, but life is not simple.

horopter The location of objects whose images lie on corresponding points. The surface of zero disparity.

diplopia Double vision. If visible in both eyes, stimuli falling outside of Panum's fusional area will appear diplopic.

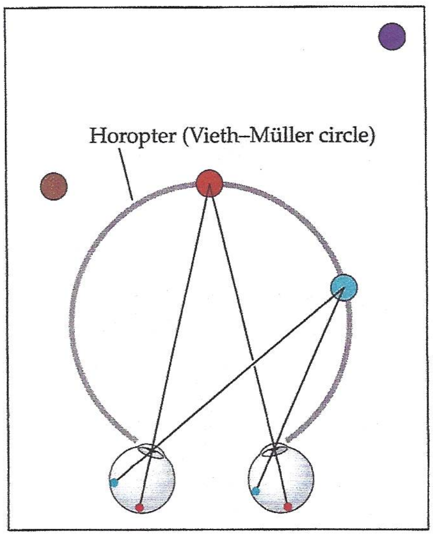

Figure 6.26 Bob is still gazing at the red crayon. This view from above traces the light rays reflecting from the red and blue crayons onto Bob's retinas. The blue crayon projects to *corresponding retinal points*—positions that are equidistant from and on the same side of the fovea. The same would be true of any object falling on the gray curve shown in the figure. (The horopter and Vieth-Müller circle are not exactly the same, but they would be very similar in this case.)

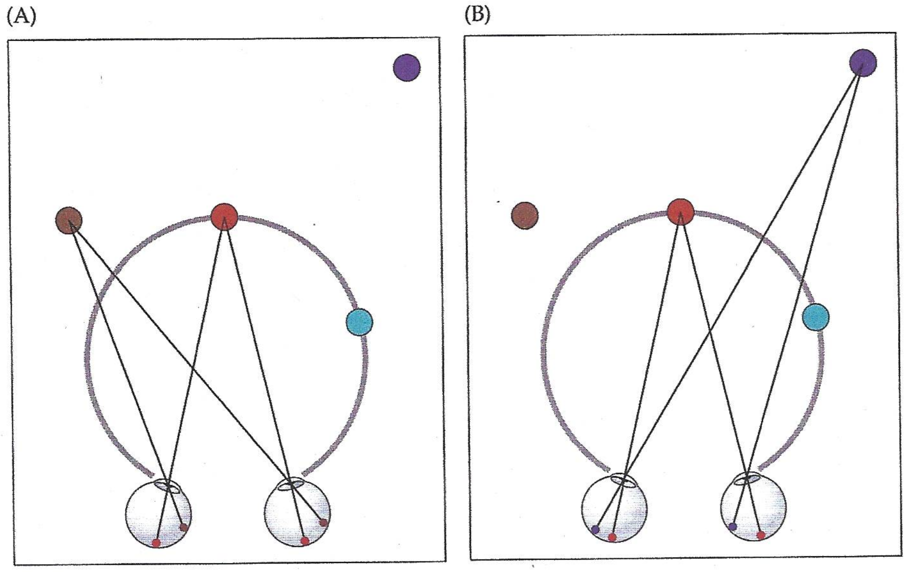

Figure 6.27 Light rays projecting from the brown (A) and purple (B) crayons onto Bob's retinas as he continues to gaze at the red crayon.

as single objects. This region of space in front of and behind the horopter, within which binocular single vision is possible, is known as **Panum's fusional area** (Panum, 1940). You can check this ability quite simply, by holding a red crayon (or pen) directly in front of you with your left hand, at a distance of about 20 centimeters (cm), and keeping both eyes on it. Now hold a blue crayon (or pen) about 5 cm to either side of the red one with your right hand, and slowly move it nearer to your eyes and then farther away, while maintaining careful fixation on the red one. You should initially see the blue crayon/pen as single, when it is about the same distance from you as the red one, because it is within Panum's fusional area. However, when it falls outside Panum's area, it will appear double. Panum's area provides a little room for small errors in eye alignment, while still maintaining single vision.

Panum's fusional area The region of space, in front of and behind the horopter, within which binocular single vision is possible.

Armed with this terminology, let's return to Bob and his crayons. Consider the retinal images of the brown crayon, lying just off the horopter. As Figure 6.24 and the view from above in **Figure 6.27A** show, rays of light bouncing off this crayon do *not* fall on corresponding retinal points: the crayon's image is farther away from the fovea on the left retina than on the right retina. Relative to the horopter, this crayon forms retinal images with a nonzero binocular disparity. The purple crayon is even farther off the horopter (**Figure 6.27B**); it forms retinal images that are even more disparate (**Figure 6.28**).

The geometric regularity that the visual system uses to extract metrical depth information from binocular disparity should now be growing clear. The larger the disparity, the greater the distance in depth of the object from the horopter.

The direction in depth is given by the *sign* (that is, "crossed" or "uncrossed") of the disparity, as illustrated in **Figure 6.29**. Suppose that Bob is looking at a red crayon with his eyes converged so that the red crayon falls on the fovea in each eye. A closer, blue crayon will form images on noncorresponding, disparate points. On the left retina, blue will lie to the left of red. Because the image is reversed, this means that, viewed from the left

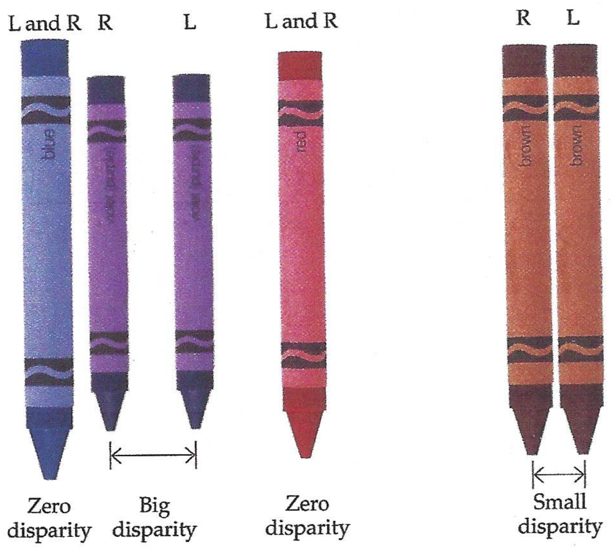

Figure 6.28 Superposition of Bob's left (L) and right (R) retinal images of the crayons in Figure 6.25, showing the relative disparity for each crayon. Size differences are ignored here. The red and blue crayons sit on the horopter and have zero disparity. They form retinal images in corresponding locations. The brown crayon forms images with a small binocular disparity. The purple crayon, farther from the horopter, has larger binocular disparity.

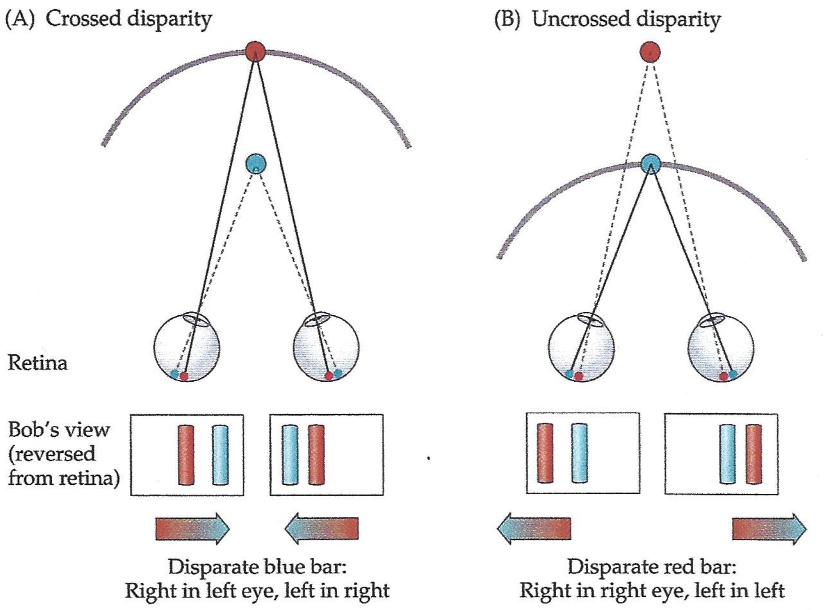

Figure 6.29 Crossed and uncrossed disparity. (A) Here Bob is foveating the red crayon. In the Bob's-eye views, the closer, blue object is seen to the right in the left eye and to the left in the right eye. In this scenario the disparity is crossed. (B) Here Bob has shifted his gaze and his horopter to the blue crayon. In the Bob's-eye views, the farther, red object is seen to the left in the left eye and to the right in the right eye. In this scenario the disparity is uncrossed

eye, blue is to the right of red. Viewed from the right eye, blue is to the left. Right in left, and left in right. This is known as **crossed disparity** (Figure 6.29A), and crossed disparity always means "in front of the horopter." In Figure 6.29B, Bob is looking at the blue crayon, and the red is seen to the left with the left eye and to the right with the right. That's **uncrossed disparity**, and uncrossed disparity always means "behind the horopter." Note that if we change our fixation, the horopter is now at a different location in space. Stereopsis is a relative depth cue that provides very-high-resolution depth information for objects that are close to the horopter.

Figure 6.30 illustrates the relationship between binocular disparity and perceived relative depth. In the middle of the graph, just to either side of the

crossed disparity The sign of disparity created by objects in front of the plane of fixation (the horopter). The term *crossed* is used because images of objects located in front of the horopter appear to be displaced to the left in the right eye and to the right in the left eye.

uncrossed disparity The sign of disparity created by objects behind the plane of fixation (the horopter). The term *uncrossed* is used because images of objects located behind the horopter will appear to be displaced to the right in the right eye and to the left in the left eye.

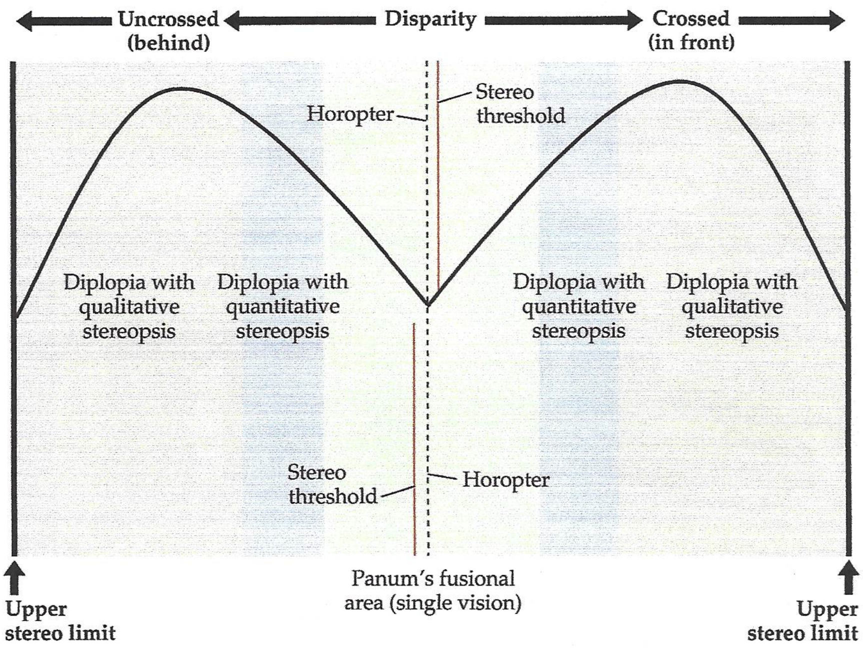

Figure 6.30 In this figure, disparity increases from zero at the horopter (dashed line in the middle of the figure). Going in either the crossed or uncrossed direction, we first find the smallest disparity that would support stereopsis (stereo threshold). Next, there is a range of single vision with quantitative stereo (green), a range of diplopia or "double vision" with quantitative stereopsis (blue), and diplopia with qualitative (just near or far) stereopsis (orange). Finally, there is an upper stereo limit, the disparity beyond which stereoscopic processing does not occur. The black curve gives a feeling for the relative size of the depth impression for that disparity. (After Ogle, 1952; Wilcox and Allison, 2009.)

stereoscope A device for simultaneously presenting one image to one eye and another image to the other eye. Stereoscopes can be used to present dichoptic stimuli for stereopsis and binocular rivalry.

horopter, are the lower limits of stereopsis—the smallest crossed and uncrossed disparities that can be detected (i.e., the stereo thresholds). Farther from the center in either direction, disparity increases. The limit of Panum's fusional area (the green zone in Figure 6.30) marks the end of the zone of single vision. Beyond that the images are diplopic. Items will look doubled but, interestingly, they are still seen in relative depth. When the disparity is not too large (the blue zones in Figure 6.30), the depth information from stereopsis is quantitative. Here stereopsis still provides an accurate estimate of the relative depth. As disparity increases beyond this range into the purple zones, it still provides usable depth order information, but it is now qualitative rather than quantitative. An item will appear to be in front or behind, but that is about all that stereopsis can provide in this range. Finally, once disparities are larger than the upper disparity limit, there is no longer any useful depth information.

Stereoscopes and Stereograms

Interestingly, although scientists had studied the geometry of binocular vision for millennia (the geometer Euclid was at it in the third century BCE), not until the nineteenth century was binocular disparity properly recognized as a depth cue. In the 1830s, Sir Charles Wheatstone invented a device called the **stereoscope** (**Figure 6.31**) that presented one image to one eye and a different image to the other eye. The stereoscope confirmed that the visual system treats binocular disparity as a depth cue, regardless of whether the disparity is produced by actual or simulated images of a scene.

For the average citizen at the time, the stereoscope was not science; it was home entertainment. The Wheatstone stereoscope held two different images in two different places. In the 1850s, however, David Brewster and Oliver Wendell Holmes invented viewers (**Figure 6.32A**) that held a card with a double image like that shown in **Figure 6.32B**. The double images were captured by cameras with two lenses separated by about 2.5 inches (≈ 63 cm), the distance between the average human's eyes. This arrangement allows stereo cameras to take a pair of pictures that mimic the images produced by the projective geometry of human binocular vision. Photographers traveled the world with these stereo cameras, capturing far-off scenes in a way that enabled a London schoolchild to

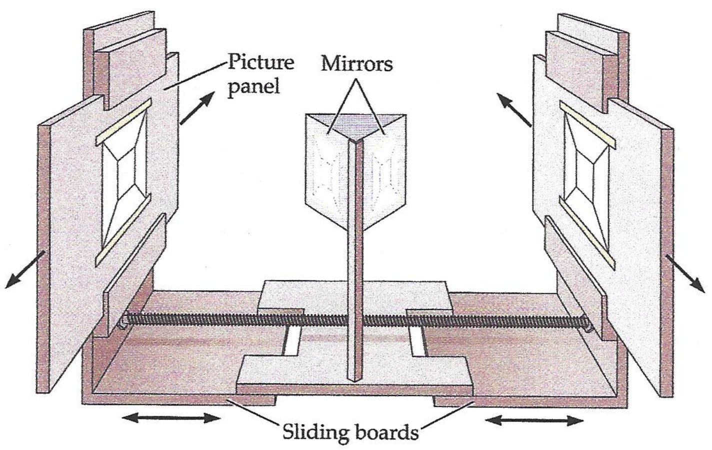

Figure 6.31 Wheatstone's stereoscope. The viewer would bring her nose up to the vertical rod at the center of the apparatus so that each eye was looking at the image reflected in one of the two mirrors. (After Wheatstone, 1838.)

Figure 6.32 Stereopsis for the masses. (A) This Holmes stereoscope—among others—brought stereo photos into many mid-nineteenth-century homes. (B) A stereo photo of South African Light Horse, a scouting regiment of the British Army, on Adderly Street in Cape Town, South Africa, in 1900. If you can free-fuse (explained later in this section), you will be able to see this scene jump out in depth.

see, for example, a vivid three-dimensional image of the British Army in Cape Town, South Africa. (For a guide to more of these historical images, as well as many other stereograms, see **Web Essay 6.2: Stereo Images on the Web**.)

A stereoscope is very helpful, but you don't actually need one to experience stereopsis. You can teach yourself a technique known as **free fusion**, which John Frisby (1980) called "the poor man's stereoscope." **Figure 6.33** contains two almost identical pictures. If you cross your eyes hard, you should see four sets of squares. This is the phenomenon of double vision (diplopia), described in the previous section. Two of those sets of squares are seen with the left eye, and two with the right. The trick is to relax just a bit until you see just three sets of squares. The far-left set is seen only in the left eye; the far-right set, only in the right; but the middle set is the fusion of two sets, one seen by the left eye and another seen by the right eye. This fusion of the separate images seen by the two eyes makes stereopsis possible.

Achieving the perception of three, instead of four or two, sets of squares in Figure 6.33 is the first step toward free fusion. The second step is to bring the middle set into focus. Convergence and accommodation, discussed earlier, normally

free fusion The technique of converging (crossing) or diverging the eyes in order to view a stereogram without a stereoscope.

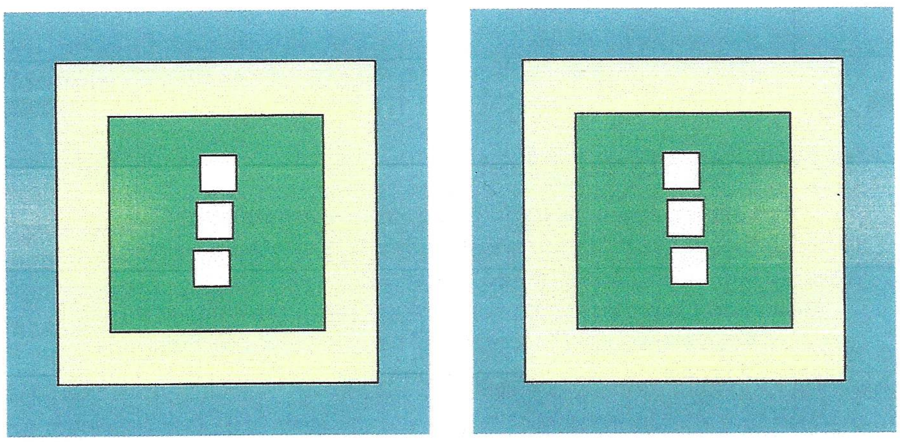

Figure 6.33 Try to converge (cross) or diverge (uncross) your eyes so that you see exactly three big blue squares here, rather than the two on the page. If you succeed, you will probably be able to see that the three little white squares lie at different depths in the middle set of big squares.

stereoblindness An inability to make use of binocular disparity as a depth cue. This term is typically used to describe individuals with vision in both eyes. Someone who has lost one (or both) eyes is not typically described as "stereoblind."

work in lockstep, so crossing your eyes automatically leads your ciliary muscles to make your lenses more spherical (unless you are presbyopic—see Chapter 2). Similar problems (in the opposite direction) will occur if you diverge your eyes. To see the middle set of squares clearly, you have to decouple accommodation and convergence. This is hard to do, but if you can manage it, then the image will come into focus and the three white squares will appear to lie at different depths in the middle set of squares. When you view them normally, notice that the white squares in the left and right panels look misaligned in opposite directions. Those are the monocular views. When you free-fuse, the opposite misalignments become the binocular disparity, and your visual system converts that disparity into a perception of depth.

The depth that you see depends on whether you converged or diverged your eyes. We described crossing, or converging, the eyes. It is also possible to free-fuse the images in Figure 6.33 by *diverging* your eyes. Divergence requires focusing on a point beyond the plane of the page so that the image of the left-hand set of squares falls on the left fovea and the image of the right-hand set falls on the right fovea. Because the images falling on the two retinas in the divergence method are reversed compared with the convergence situation, the disparities are reversed and the perceived depth will be reversed. That is, if you converge, the top square will be the farthest back. If you diverge, it will appear closest to you. Either converging or diverging will produce a clear stereoscopic effect, so give it a try.

Before we go on, we should note that approximately 3–5% of the population lacks stereoscopic depth perception—a condition known as **stereoblindness**. Stereoblind individuals might be able to achieve the perception of three sets of squares in Figure 6.33, but the little white squares will not pop out in depth. Stereoblindness is usually a secondary effect of childhood visual disorders such as strabismus, in which the two eyes are misaligned. (See the Sensation & Perception in Everyday Life box "Recovering Stereo Vision" below.) If you had such a visual disorder during childhood and/or you've been diagnosed with stereoblindness, we apologize, but you just won't perceive depth in the stereograms presented here and on the website. That said, many people who try and fail to see depth in stereograms have "normal" vision (wearing glasses doesn't count as "abnormal" in this case). Those people just need practice, so don't give up. **Web Activity 6.3: Stereoscopes and Stereograms** provides more stereograms for practice, and **Web Essay 6.2: Stereo Images on the Web** leads to another website with more tips for free-fusing.

FURTHER DISCUSSION of strabismus can be found in Chapter 3 on page 94.

Sensation & Perception in Everyday Life

Recovering Stereo Vision

As noted earlier, about 3–5% of the population are stereoblind, usually as a result of early childhood visual disorders. Can stereo vision be recovered later in life? Meet Stereo Sue and Binocular Bruce.

Susan Barry, a professor of neurobiology, had strabismus as an infant and had been stereoblind essentially all her life. Her book *Fixing My Gaze* (2009) provides a fascinating, informative, and beautifully written

Sensation & Perception in Everyday Life *(continued)*

account of her acquisition of stereopsis following vision therapy. It describes her transformative journey from the many visual, social, and psychological challenges of a turned eye (a squint or strabismus) early in life, to the sudden enrichment of her perceptions of the world following successful unconventional visual therapy begun at 48 years of age. (An earlier article about her visual recovery was published in *The New Yorker* under the title of "Stereo Sue" by Oliver Sacks.)

Barry vividly recounts how acquiring stereoscopic vision led to a dramatic improvement of her perception of depth, or the appreciation of the "space between objects." A particularly valuable insight is her argument for the inability of people with normal vision to appreciate the visual experience of being stereoblind. Naively one might think that this experience could be duplicated simply by closing one eye so all information about depth was conveyed by monocular cues. Not so, however, Barry argues: the monocular experience of a typically reared person who closes one eye has been informed by a lifetime of experience with stereoscopic vision and so is far different from that of a person who is stereoblind. As a result, Barry's new stereoscopic vision brought much more to her life than just depth perception: objects became clearer, motion perception became more veridical, and her movement around the world became more confident. Even more poignant is her vivid description of the enhanced sense of touch she had developed over the years and its key role in informing her newly acquired sense of stereo vision.

Barry did not simply "recover" stereopsis, but rather had to relearn to see with stereo vision. As blind or deaf individuals often describe, individuals deprived of a sense are not just "missing" a sense. Rather, they have developed an entirely different way of sensing the world. Upon sensory restitution, a fascinating but rather disturbing experience unfolds as the brain has to adapt to a new way of functioning.

Even more dramatic is the experience of Binocular Bruce (the late Bruce Bridgeman), a very perceptive vision scientist who had been stereo-deficient all his life. Remarkably, he recovered stereopsis after watching the 3D movie *Hugo* (Bridgeman, 2014). Whether this sort of immersive experience, with very large disparities along with many other depth cues, will be a generally effective treatment for abnormal stereopsis remains to be tested. However, these case studies, along with lab studies of perceptual learning that have resulted in the recovery of stereopsis (Ding and Levi, 2011; Vedamurthy et al., 2015), call into question the notion that has been the received wisdom, that recovery of stereopsis can only occur during early childhood. The idea, dating back to the early twentieth century, has been that there is a "critical period" of development when the visual system is still plastic and capable of change. After that, it was thought, our basic visual capabilities are fixed. This led a number of practitioners to tell Susan Barry and her mother that "nothing could be done" about her vision (one suggested that she might need a psychiatrist). Since binocular neurons are present in the visual cortex of primates within the first week of life (see the "Development of Binocular Vision and Stereopsis" section below), Barry surmises that some of the innate wiring of her binocular connections remained intact and that vision therapy taught her to move her eyes into position for stereo vision, "finally giving these neurons the information they were wired to receive" (Barry, 2009).

Random Dot Stereograms

For 100 years or so after the invention of the stereoscope, it was generally supposed that stereopsis occurred relatively late in the processing of visual information. The idea was that the first step in free-fusing images such as those in **Figure 6.34** would be to analyze the input as a face. We would then use the slight disparities between the left-eye and right-eye images of the nose, eyes, chin, and other objects and parts to enrich the sense that the nose sticks out in front of the face, that the eyes are slightly sunken, and so on.

Bela Julesz, a Hungarian radar engineer who spent most of his career at Bell Labs in New Jersey, thought the conventional wisdom might be backward. He theorized that stereopsis might be used to *discover* objects and surfaces in

Figure 6.34 A stereo photograph of a woman's face.

random dot stereogram (RDS) A stereogram made of a large number (often in the thousands) of randomly placed dots. Random dot stereograms contain no monocular cues to depth. Stimuli visible stereoscopically in random dot stereograms are Cyclopean stimuli.

Cyclopean Referring to stimuli that are defined by binocular disparity alone. Named after the one-eyed Cyclops of Homer's *Odyssey*.

the world. Why would this be useful? Julesz thought that stereopsis might help reveal camouflaged objects. A mouse might be the same color as its background, but out in the open it would be in front of the background. A cat that could use stereopsis to break the mouse's camouflage would be a more successful hunter. (Cats do have stereopsis, by the way [Blake, 1988; R. Fox and Blake, 1971].) To prove his point, Julesz (1964, 1971) made use of **random dot stereograms (RDSs)**. An example is shown in **Figure 6.35**. If you can free-fuse these images, you will see a pair of squares, one sticking out like a bump, the other looking like a hole in the texture (which one is the bump and which one is the hole depends, again, on whether you converge or diverge your eyes).

The important point about RDSs is that we cannot see the squares in either of the component images. We cannot see the squares using any monocular depth cues. These are shapes that are defined by binocular disparity alone. Julesz called such stimuli **Cyclopean**, after the one-eyed Cyclops of Homer's *Odyssey*. Wheatstone showed with his stereoscope that binocular disparity is a necessary condition for stereopsis. Julesz demonstrated with RDSs that disparity is *sufficient* for stereopsis. To understand how RDSs are made, see **Web Activity 6.3: Stereoscopes and Stereograms**. And to learn something about 3D movies and games see **Web Essay 6.3: Stereo Movies, TV, and Video Games**.

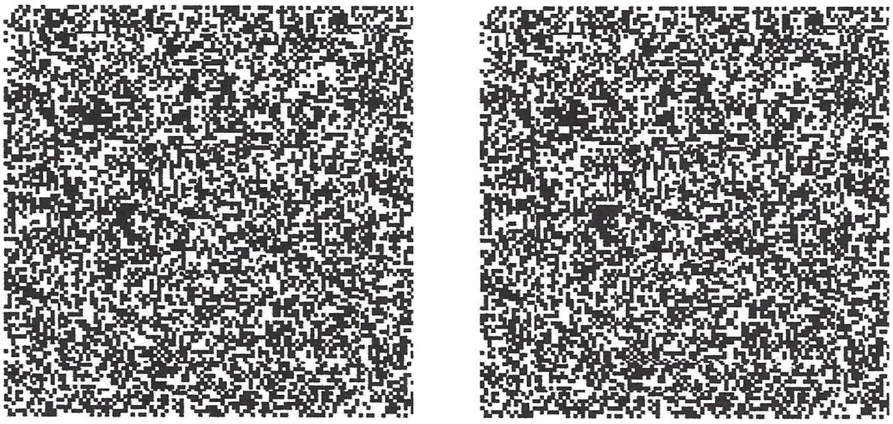

Figure 6.35 If you can free-fuse this random dot stereogram, you will see two rectangular regions: one in front of the plane of the page, the other behind the page. Which is which depends on whether you converge or diverge in order to fuse the two squares.

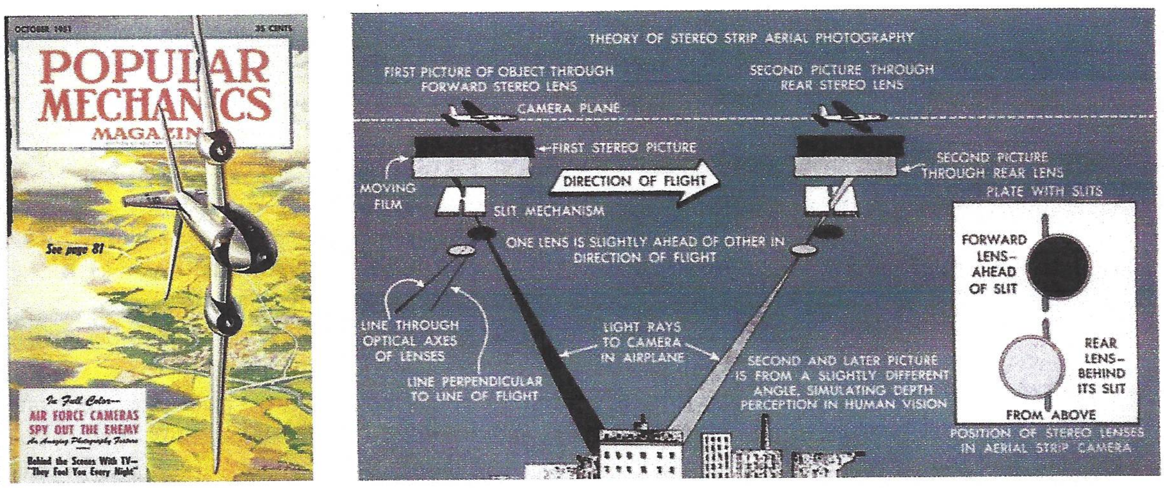

Figure 6.36 It is possible to make effective stereoscopic images of terrain by taking two aerial pictures from two, quite widely separated viewpoints. (From Goddard, 1951; Hearst Communications Inc.., reprinted with permission.)

Using Stereopsis

Stereopsis has been put to work in a number of fields. The military has known for a long time that you can get more information out of aerial surveillance if your view of the ground is stereoscopic. However, if you've ever looked out the window from thousands of meters in the air, you may have noticed the ground looks rather flat. This is due to yet more geometry. Stereopsis can provide useful information about metric depth only for distances up to 40 meters (Palmisano et al., 2010). With eyes a few centimeters apart, you don't get adequate disparity from more distant targets. What you need are eyes separated by hundreds of meters. This can be done if you have a plane and a special camera. **Figure 6.36** reprints a figure from a 1951 issue of *Popular Mechanics* in which Colonel George W. Goddard showed the public how images taken from two vantage points produced stereo images during the Korean War.

We can also use stereopsis to have a better look inside the body. **Figure 6.37** shows a stereo view of a mammogram, an X-ray of the breast of the sort used to detect breast cancer. In order to create another free-fusion demonstration, we've

Figure 6.37 This stereo mammogram was created by taking X-rays of a woman's breast from two viewpoints. If you can free-fuse, you will see the structures in the image separate in depth, making it easier to decide whether the filament is part of the breast tissue or is a different structure. (Courtesy of David Getty.)

correspondence problem In reference to binocular vision, the problem of figuring out which bit of the image in the left eye should be matched with which bit in the right eye. The problem is particularly vexing when the images consist of thousands of similar features, like dots in random dot stereograms.

printed the one image of the breast twice, to the left and the to right of a second image of the same breast, placed in the middle of this figure. The two images are taken from slightly different viewpoints, creating a binocular disparity if one view is presented to one eye and the other view to the other eye. Thus, if you free-fuse the images so that you see four breast images, one of the two center images will have the correct disparities (depending on whether you diverge or converge your eyes to free-fuse). The correct one will show the white wire—a marker for the surgeon—on top of the breast. You can see that the breast tissue is characterized by a network of intersecting structures. This is like our earlier example of looking at the little branches of a tree while lying on your back with just one eye open. It can be very hard to tell which line-like structures actually intersect and which ones lie at different depths. This turns out to be important when reading a mammogram. A starburst structure might be a sign of cancer, but not if it is an accidental viewpoint (see Chapter 4) of structures at different depths in the breast that just happen to form a suspicious pattern in a two-dimensional projection. Stereopsis can disambiguate this situation. If you can free-fuse these images, you will see the texture in three dimensions, and you will be able to determine how different structures relate to each other in depth. Stereoscopic displays are beginning to be used in radiology (Held and Hui, 2011), and they can reduce the error rate in these important tasks (Getty, D'Orsi, and Pickett, 2008).

Is stereopsis useful in everyday life? In people with normal binocular vision, visually guided hand movements are significantly impaired when viewing is restricted to one eye (Fielder and Mosely, 1996), likely owing to the fact that binocular depth thresholds are about a factor of 10 better than monocular thresholds (McKee and Taylor, 2010). These results are mirrored in patients with amblyopia ("lazy eye") for whom many observed visuomotor deficits are due to impaired stereopsis, and in particular impaired visual feedback control of movements, rather than visual acuity loss (Grant and Moseley, 2011). Loss of stereopsis may also result in unstable gait, especially reduced accuracy when a change of terrain (e.g., steps) occurs, and difficulties for children in playing some sports.

FURTHER DISCUSSION of stereo sensation can be found in Chapter 10 (sound localization; pages 316–321) and Chapter 14 (binaral rivalry in olfaction; page 481).

Stereoscopic Correspondence

If you successfully free-fused the random dot patterns in Figure 6.35, you solved a truly daunting problem. Even if you didn't, if you have normal binocular vision, you are solving the **correspondence problem** all the time. The correspondence problem is the problem of figuring out which bit of the image in the left eye should be matched with which bit in the right eye. **Figures 6.38** and **6.39** use an extremely simple situation to illustrate why correspondence is so tricky. There are, of course, just three dots in Figure 6.38. Figure 6.39A traces the paths of the rays of light from the printed circles on the page to the images on the viewer's retinas. The retinal images of the circles are labeled to make it clear which image on the left retina corresponds to which image on the right retina, but your visual system has no such labels. All it knows about is the retinal images, as shown in Figure 6.39B. Figure 6.39C shows another possible geometric interpretation of the situation: if the left retinal image of circle 2 is matched to the right retinal image of circle 1, and the left retinal image of circle 3 is matched to the right retinal image of circle 2, you will perceive *four* circles, with the inner pair of

Figure 6.38 Is this a simple picture or a complicated computational problem?

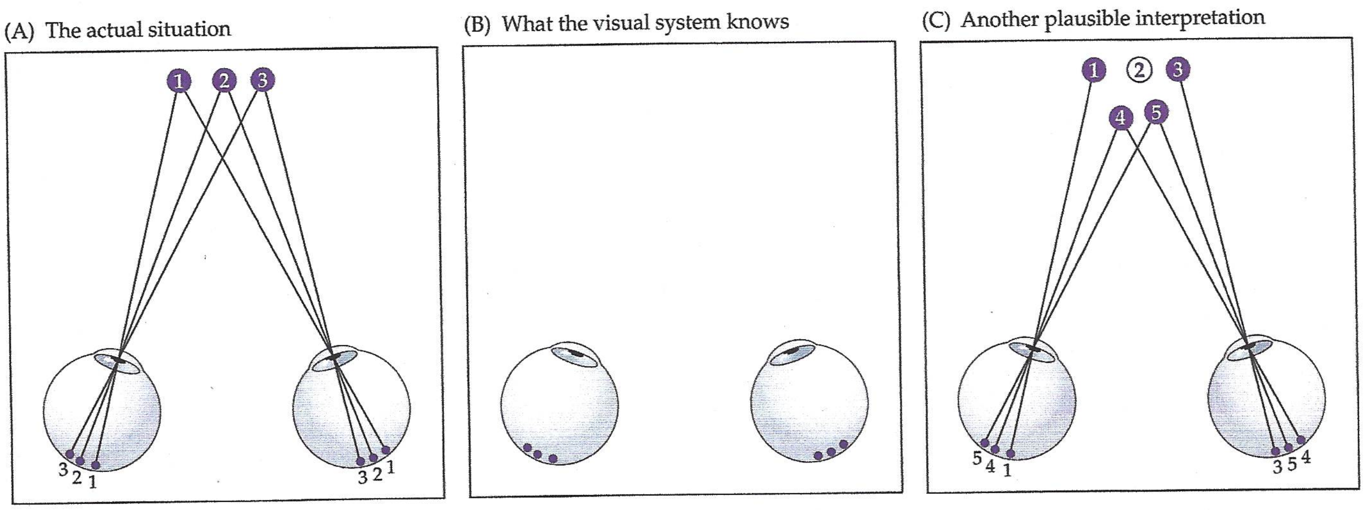

Figure 6.39 Interpreting the visual information from the three circles in Figure 6.38 (A, B). It would require careful placement, but four dots in the world could produce three dots in each eye as in (C).

circles perceived as floating in front of the outer pair. In fact, you may be able to experience this for yourself if you can cross your eyes correctly.

With only three elements in the visual scene, it isn't hard to imagine how the visual system might achieve the proper correspondence: first match the two circles whose images fall on the foveas; then match the two images to the left of the foveas with each other; then match the two images to the right of the foveas. Before the introduction of random dot stereograms, a similar logic seemed reasonable for more complex scenes, too. Go back to the face in Figure 6.34. Our visual systems could solve the correspondence problem by first finding the parts of the two faces and then matching nose to nose, mouth to mouth, and so forth. The RDS in Figure 6.35, however, contains thousands of identical black and white dots falling on each retina. How can we be sure that the dot in the center of the fovea of one eye corresponds to the dot in the center of the other eye? Even if we knew that, could we really match each dot in the right eye with just one dot in the left eye? If there were a little dirt on the page, would the whole process collapse? How in the world does our visual system succeed in making the proper matches?

Matching thousands of left-eye dots to thousands of right-eye dots in Figure 6.35 would require a lot of work for any computational system. However, the problem is simpler if we look at a blurred version of the stereogram. Blurring leaves only the low-spatial-frequency information. **Figure 6.40** shows the low spatial frequencies of the stereogram from

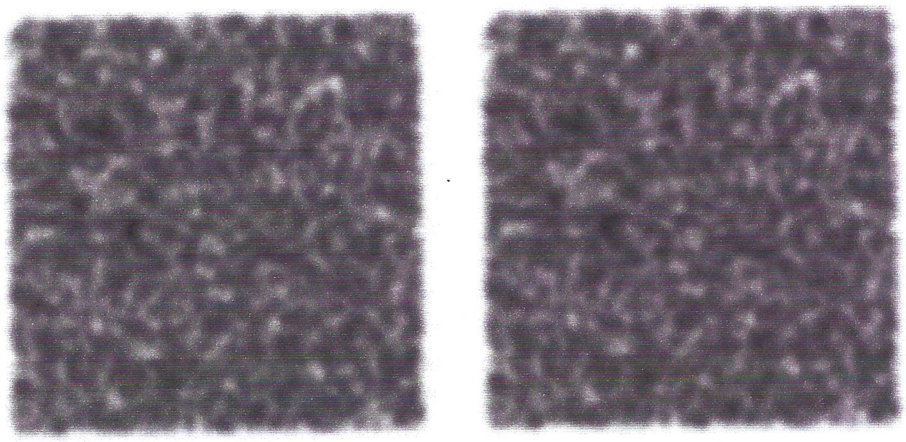

Figure 6.40 A low-spatial-frequency filtered version of the stereogram in Figure 6.35.

uniqueness constraint In reference to stereopsis, the observation that a feature in the world is represented exactly once in each retinal image. This constraint simplifies the correspondence problem.

continuity constraint In reference to stereopsis, the observation that, except at the edges of objects, neighboring points in the world lie at similar distances from the viewer. This is one of several constraints that have been proposed as helpful in solving the correspondence problem.

Figure 6.35. Now, rather than thousands of dots, we have just a few large blobs. Now you could imagine a process that, for example, matched the black blob in the upper left corner of the left image with the very similar blob in the right image. Crude matches of this sort could act as anchors, allowing the visual system to fill in the finer (high-spatial-frequency) matches from there.

In addition to starting with low-spatial-frequency information, David Marr and Tomaso Poggio (1979) suggested two more heuristics for solving the correspondence problem. They called these the uniqueness and continuity constraints. The **uniqueness constraint** acknowledges that a feature in the world is represented exactly once in each retinal image. Working in the opposite direction, the visual system knows that each monocular image feature (e.g., a nose or a dot) should be paired with exactly one feature in the other monocular image. Notice that Figure 6.39C would not violate uniqueness. Each dot in the world would be represented exactly once in each retinal image. The odd thing is that two dots in the real world could be represented by the same dot in the retinal image. The **continuity constraint** holds that, except at the edges of objects, neighboring points in the world lie at similar distances from the viewer. Accordingly, disparity should change smoothly at most places in the image. (These constraints are difficult to illustrate on a static page, but **Web Activity 6.4: Stereoscopic Correspondence** provides dynamic explanations.) With those constraints, the correspondence problem is not entirely solved, but it is made much more tractable. There are not so many possible solutions. However, recent work suggests that identifying correct matches may not be the optimal strategy (Goncalves and Welchman, 2017). Rather, they suggest that the brain uses "what not detectors" that sense dissimilar features in the two eyes. These suppress unlikely interpretations of the scene and facilitate stereopsis by providing evidence against interpretations that are incompatible with the true structure of the scene.

The Physiological Basis of Stereopsis and Depth Perception

Now that we know something about the theoretical basis of stereopsis, we can ask how it is implemented by the human brain. The most fundamental requirement is that input from the two eyes must converge onto the same cell. As noted in Chapter 3, this convergence does not happen until the primary visual cortex, where most neurons can be influenced by input from both the left and right eyes—that is, they are binocular (Hubel and Wiesel, 1962). A binocular neuron has two receptive fields, one in each eye. In binocular primary visual cortex neurons, the receptive fields in the two eyes are generally very similar, sharing nearly identical orientation and spatial-frequency tuning, as well as the same preferred speed and direction of motion (Hubel and Wiesel, 1973). Thus, these cells are well suited to the task of matching images in the two eyes.

Many binocular neurons respond best when the retinal images are on corresponding points in the two retinas, thereby providing a neural basis for the horopter. However, many other binocular neurons respond best when similar images occupy slightly *different* positions on the retinas of the two eyes (Barlow, Blakemore, and Pettigrew, 1967; Pettigrew, Nikara, and Bishop, 1968). In other words, these neurons are tuned to a particular binocular disparity, as diagramed in **Figure 6.41**.

Recall the distinction, from earlier in this chapter, between metrical and nonmetrical depth cues. Stereopsis can be used both metrically and nonmetrically. Nonmetrical stereopsis might just tell you that a feature lies in front of or behind the plane of fixation. Gian Poggio and his colleagues (Poggio and Talbot, 1981) found disparity-tuned neurons of this sort in V2 (which stands for "visual

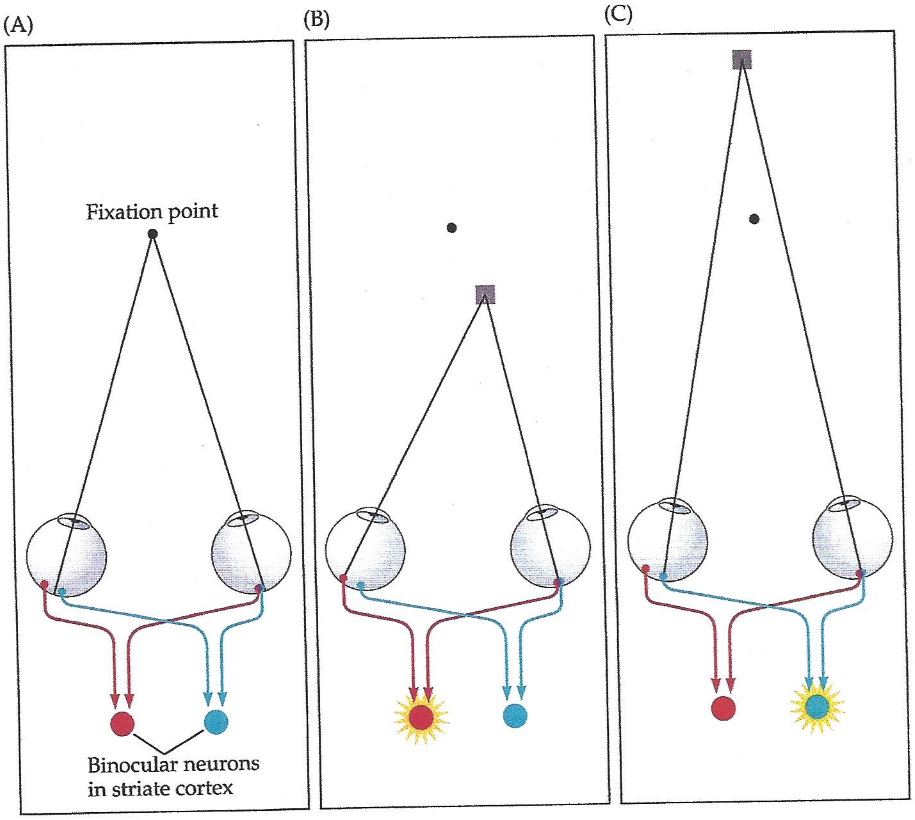

Figure 6.41 In these simplified diagrams of receptive fields for two binocular-disparity–tuned neurons in primary visual cortex, the red neuron "sees" stimuli falling on the red receptive fields, and the blue neuron responds to stimuli falling on the blue receptive fields (these receptive fields overlap on the right retina). (A) The overall picture, showing the fixation point in relation to the two retinas. (B) The red neuron responds best to a stimulus closer to and slightly to the right of the fixation point. (C) The blue neuron responds best if its preferred stimulus is behind and slightly to the left of fixation.

area 2") and some higher cortical areas. Some neurons responded positively to disparities near zero—that is, to images falling on corresponding retinal points. Other neurons were broadly tuned to a range of crossed (near) or uncrossed (far) disparities. On the other hand, stereopsis can also be used in a very precise, metrical manner. Indeed, stereopsis is a "hyperacuity" like Vernier acuity (see **Web Essay 3.1: Hyperacuity**), with thresholds smaller than the size of a cone. Both of these forms of stereopsis have their uses, and functional magnetic resonance imaging (fMRI) data suggest that the dorsal *where* pathway is most interested in metrical stereopsis, while the ventral *what* pathway makes do with more categorical, near-versus-far information (Preston et al., 2008) (see Chapter 4).

The neural bases of other depth cues have also been investigated. For example, when we discussed motion parallax earlier, we suggested moving your head back and forth while looking into the branches of a tree in order to create a more vivid impression of the depth relationships among the branches and twigs. To exploit that cue properly, you need to know how your head is moving (see Chapter 12) and how items in the visual field are moving (see Chapter 8). Nadler, Angelaki, and DeAngelis (2008) looked for the neural substrate of parallax in the middle temporal area (area MT) of the brain of macaque monkeys. As we will see in Chapter 8, this area is very important in the perception of motion. Nadler et al. set up an apparatus where the monkey was moved from side to side while items on the screen also moved. If the monkey was integrating signals about its head movement with the motion signals, then the objects on the screen should have been seen in depth. Otherwise, they would have been seen as just moving in the plane of the screen. It turns out that cells in area MT can signal the sign of depth (near or far) based on this motion parallax signal alone.