

Chapter 14

Illusions and Ambiguous Shapes: Epiphenomena of Brain Mechanisms

In the early days of radio, it was often possible to guess the method of signal amplification and the kind of power supply used in a receiver by characteristic whistles, howls, hums, and low-frequency throbbing sounds called motorboating. These audio effects would occur under certain conditions as intrusive epiphenomena associated with particular kinds of circuits and were telltale clues to their design. In the same way, illusions and other anomalous perceptions may provide clues to some fundamental mechanisms in the human brain and lend selective support to particular candidates among competing theories of brain design. For example, it was shown in chapter 5 that the postulated neuronal wiring scheme connecting the 3-D retinoid to the mosaic cell array is able to account for size constancy in human visual perception, a normal process with obvious ecological utility. But it was also demonstrated on the basis of the same brain circuitry that accurate predictions would be made about a variety of visual illusions involving the perceived size and shape of retinal after images as a function of the distance and slant of fixated surfaces—phenomena with no apparent utility. In this chapter, we consider other examples of anomalous visual phenomena and show how they can be explained by the natural operations of the hypothesized brain mechanisms.

Seeing More Than Is There

If a vertically oriented slit in an otherwise occluding screen is fixated while a visual pattern is moved back and forth behind it, much more of the pattern can be seen at a given instant than is physically present within the aperture of the slit. For example, with a screen aperture of 0.1 degree of visual angle in width, it is possible to see a complete figure having a true width of 2.0 degrees (20 times as wide the aperture) if it is in horizontal oscillatory motion behind the masking screen. This phenomenon of anorthoscopic perception was reported

as long ago as 1862 by Zöllner and by Helmholtz in 1867. More recently, it has been called the seeing-more-than-is-there effect (McCloskey and Watkins 1978). Efforts to understand this interesting phenomenon continue, and it remains the subject of active investigation (Casco and Morgan 1984; Parks 1965, 1970; Rock et al. 1987; Shimojo and Richards 1986).

The retinoid system can solve a number of fundamental problems in the realm of cognitive ecology. Moreover, it can be shown that the same putative brain mechanism is able to account in a straightforward fashion for the seeing-more-than-is-there (SMTT) effect—as an epiphenomenon of retinoid dynamics, with no apparent ecological utility.

Explaining SMTT

The ability of the retinoid system to capture a fleeting image in short-term memory and translate its neuronal representation in an egocentric spatial frame is central to the physical explanation of SMTT. On this account, the SMTT phenomenon occurs in the standard experimental paradigm when the successive pattern segments, directly observed through a narrow aperture, evoke autaptic excitation that is swept across a retinoid surface driven by signals from the visual system's motion detection cells. This explanation depends on the detection of the velocity and direction of the stimulus pattern on the basis of information available in the screen aperture and the short-term memory properties of the retinoid that enable it to assemble a unified representation of the stimulus from the sequence of partial inputs that are registered in its aperture region and then shifted stepwise across adjacent autaptic cells.

The significant factors in the SMTT paradigm are shown in schematic form in figure 14.1. The stimulus pattern (depicted in the figure as a triangle) is moving to the right, with its leading edge just entering the aperture in an otherwise occluding screen. The symbols in figure 14.1 are to be interpreted as follows:

S = complete stimulus pattern behind occluding screen.

s = input from stimulus contour appearing in aperture.

μ = autaptic cell in aperture field of retinoid excited by stimulus segments.

μ = autaptic cell in retinoid beyond aperture field.

ι = excitatory interneuron in retinoid.

ρ = retinoid shift control cell (shift right).

ing cell (ρ) and the spike frequency of its donor autaptic cell ($\dot{\mu}$ or μ). Thus

$$F_i \propto (\rho, \dot{\mu}) \quad \text{or} \quad F_i \propto (\rho, \mu)$$

Given any arbitrary pulse rate from a shift control cell (say ρ), variation in the actual velocity of translation of an excitation pattern across the surface of a retinoid will be a positive function of the discharge frequency (F) of those retinoid cells carrying the pattern. This follows from the fact that the latency of discharge in each successive cell on the translation route of the image is inversely related to the F of its donor cell. If the translation velocity of an assembled retinoid image is equal in direction and velocity to the stimulus pattern S (in degrees of visual angle per second), then it is assumed that the perceived proportions of the pattern will correspond to the veridical proportions of the stimulus. If, however, translation velocity on the retinoid is less than stimulus velocity, then the stimulus pattern will be perceived as shorter along its axis of motion in relation to its other dimensions. As a corollary, it follows that if translation velocity on the retinoid should somehow be greater than stimulus velocity, the stimulus pattern would be perceived as longer along its axis of motion.

The dynamic properties of the retinoid system not only explain the sheer occurrence of this phenomenon but also predict certain systematic changes in SMTT perception as experimental parameters vary. In a series of SMTT experiments, Trehub and Pollatsek (1986) made predictions about the effects of change in stimulus velocity, aperture width, and ambient illumination on the basis of the SMTT-retinoid paradigm presented above.

General Method

Visual stimuli were presented on a CRT screen that was partially masked by a sheet of gray posterboard with a narrow, vertically oriented aperture in its center. The height of the aperture was fixed and was greater than the height of the stimulus pattern. The width of the aperture was adjustable and always much less than the width of the stimulus, so that only a small segment of the stimulus pattern could be seen by the subject at any given moment. The subject was provided with a hand-held controller. The controller had a potentiometer knob that allowed the subject to vary the rate of stimulus motion across the face of the CRT from 1.0 degree per second to a maximum velocity of 12.5 degrees per second at the viewing distance used in this study. The subject could press a button on the controller to signal

when a perceptual criterion was met. After instructions were read, all experimental trials were controlled by a computer. Stimulus presentation, recording, and analysis of the subjects' responses were performed automatically according to a preprogrammed protocol.

Experiment 1

Hypothesis 1 Given a horizontally oscillating figure in the SMTT paradigm, the perceived width of the figure will contract as stimulus velocity increases.

This follows from the fact that the dwell-time of S on $\dot{\mu}$, hence $F_{\dot{\mu}}$, will decrease with an increase of stimulus velocity, causing the rate of translation over the retinoid to decrease relative to the true stimulus velocity.

Hypothesis 2 For any given stimulus velocity, the perceived width of the figure will increase as aperture width increases.

This follows from the fact that the proportion $\Sigma\dot{\mu}/\Sigma\mu$ will increase as the aperture for direct stimulation increases. Since $F_{\dot{\mu}} > F_{\mu}$, the velocity of translation over the retinoid will be greater for the larger aperture.

Procedure The subject was seated approximately 58 centimeters from the CRT. A laterally oscillating triangle on the face of the CRT was exposed to the subject through the narrow aperture in the masking screen. The triangle disappeared behind the screen at the terminus of both left and right excursions. The height of the triangle was fixed at 0.7 degree in visual angle, and the base of the triangle varied randomly on each trial within a range of 0.8 degree to 1.9 degrees in visual angle. Incident illumination on the CRT mask was set at approximately 8.0 foot-candles. The subject was instructed to fixate the center of the aperture on each stimulus presentation and to adjust the controller knob until the width of the perceived triangle was equal to its height. As soon as this occurred, the subject was to press the controller button, which recorded the current rate setting, cleared the CRT screen, and terminated the trial.

Data were collected from four subjects under two aperture conditions in five 20-trial blocks for each condition. Two subjects were first tested on five blocks with a 0.1 degree aperture, followed by five blocks with a 0.2 degree aperture. During each block of trials, there was an interstimulus interval of 5 seconds following each trial response of the subject. A rest interval of 90 seconds was given after each block of trials.

Results In this experiment, the independent variables were the height-width ratios of the stimulus patterns and the width of the viewing aperture. The dependent variable was the stimulus velocity set by the subject to give the perception of height-width equivalence. At each velocity setting required to meet the criterion of equivalence, the ratio of perceived width to true width was computed and expressed as a contraction ratio. Figure 14.2 shows the results plotted separately for each of the four subjects. Hypothesis 1 was confirmed by a statistical test of the response curves obtained from each of the subjects under the two aperture conditions.

To test hypothesis 2, we analyzed the stimulus velocity settings selected by the subject to achieve the contraction ratios required to perceive height-width equivalence for the 0.1 degree versus 0.2 degree viewing aperture. If the subject sees a wider triangle given a wider aperture, then the subject must increase the velocity of the stimulus to perceive equivalence. Hypothesis 2 was confirmed by a statistical test of the differences in the response curves obtained from each of the subjects under the two aperture conditions (figure 14.2).

Experiment 2

Hypothesis 3 For a given aperture, the perceived width of the stimulus figure will increase as ambient illumination increases.

This follows from the assumption that ambient illumination increases both F_{μ} and F_{ν} . Thus velocity of translation over the retinoid should be greater with higher general illumination.

Procedure The apparatus, instructions, and procedure were the same as in experiment 1, except the width of the aperture was fixed at 0.2 degree and data were collected under two different intensities of ambient illumination. Two subjects were tested with 8.0 foot-candles of incident illumination followed by 2.0 foot-candles. For the remaining subjects, the order of illumination conditions was reversed.

Results In this experiment, the independent variables were the height-width ratios of the stimulus patterns and the level of ambient illumination. The dependent variable was the stimulus velocity set by the subject to give the perception of height-width equivalence. Figure 14.3 shows the results plotted separately for each of the four subjects. Hypothesis 3 was not confirmed in this test since only one of the subjects showed the predicted effect of an increase in the perceived width of the figure under higher ambient illumination.

Experiment 3

Hypothesis 4 In the SMTT paradigm with stimulus parameters held constant, the absolute number of stimulus sweeps required to squelch the SMTT effect will increase as ambient illumination increases (one sweep equals one complete excursion of the figure behind the aperture).

This follows from the assumption that the small interneurons (ι) in the retinoid will fatigue under prolonged driving. Cells of this type have a relatively high ratio of surface membrane to internal volume and are therefore more easily fatigued than are the larger autaptic neurons (Shepherd 1974). Thus, F_i will decrease to the point that it no longer adds sufficient excitation to the target autaptic cells (μ) for them to reach discharge threshold. When this occurs, pattern translation over μ will cease, and SMTT will be squelched. Since the theoretical model assumes that ambient illumination (λ) adds biasing excitation to μ , it follows that SMTT squelching should occur later (more successive stimulus sweeps will be required) under higher ambient illumination. This prediction is opposite to what might be expected given the reduction in stimulus contrast on the face of the CRT screen when ambient illumination increases.

Procedure The apparatus in experiment 3 was the same as used in experiments 1 and 2. The screen aperture was fixed at 0.2 degree in visual angle. The stimulus presented on each trial was a triangle with a height of 0.7 degree and a base of 1.7 degrees in visual angle. The triangle oscillated horizontally at a constant rate of 10.0 degrees per second and disappeared beyond the aperture on both left and right excursions. In this test, however, the subject could not vary the rate of stimulus motion. The subject was instructed to fixate the center of the aperture and to press the controller button as soon as the perceived pattern shifted from a horizontally oscillating triangle to a short, vertically oscillating line segment above a short, stable line segment (the veridical retinal input). When the button was pressed, the number of stimulus sweeps required for the perceptual shift (squelch of SMTT) was automatically recorded. The screen was then cleared and, after a 5-second interval, a warning buzzer sounded and was followed a second later by another trial.

Each subject was given 10 consecutive trials under each of two intensities of ambient illumination. Two subjects were tested first with 8.0 foot-candles, followed by 2.0 foot-candles. The order of illumination was reversed for the other two subjects. A rest interval of 5 minutes was allowed between the illumination conditions.

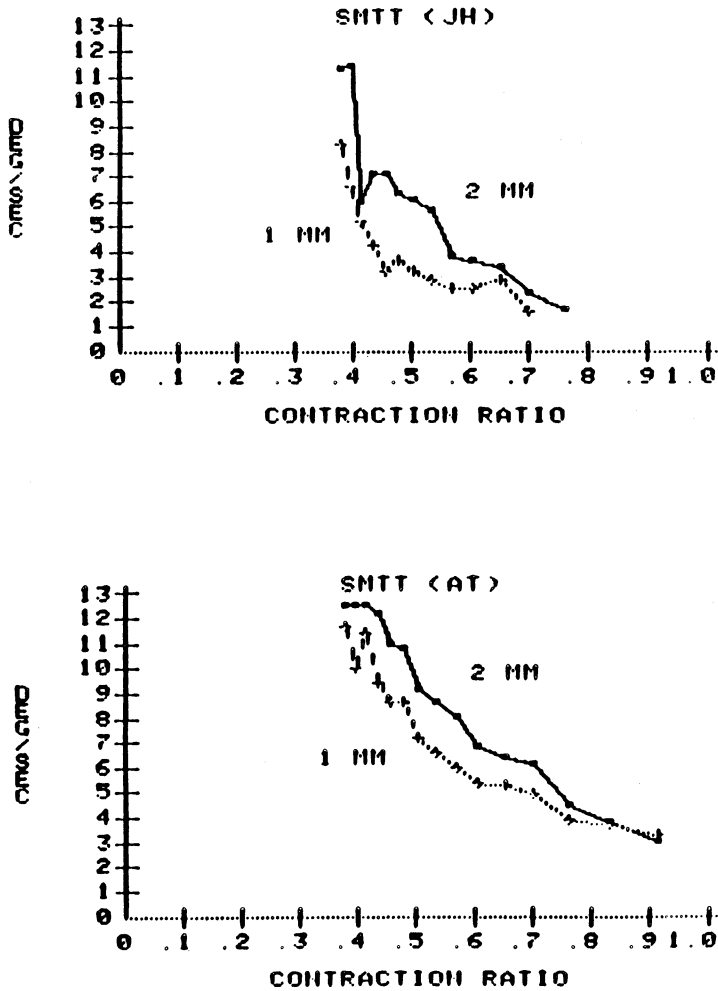


Figure 14.2
Experiment 1: Adjusted stimulus rates plotted against contraction ratios needed to achieve a perceived width-to-height match for each of four subjects. Aperture width = 0.1 degree (1 MM) and 0.2 degree (2 MM). Source: Trehub and Pollatsek 1986.

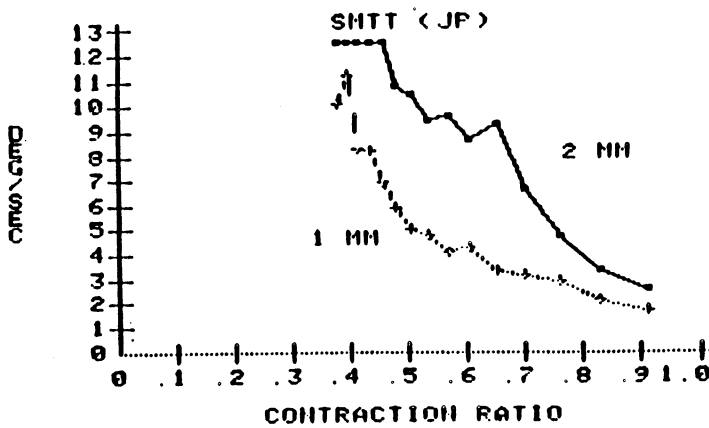
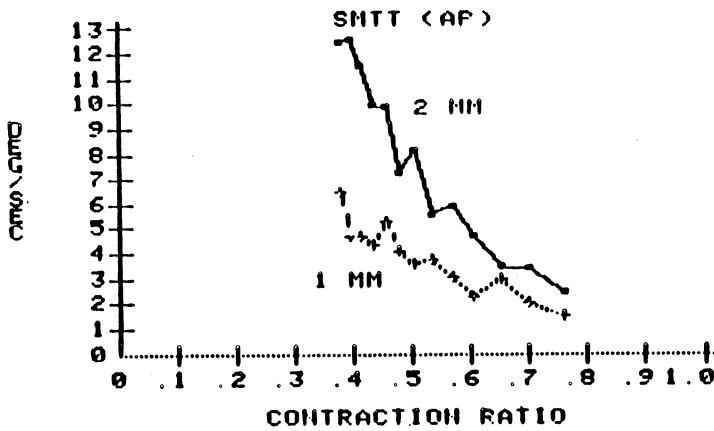


Figure 14.2 (continued)

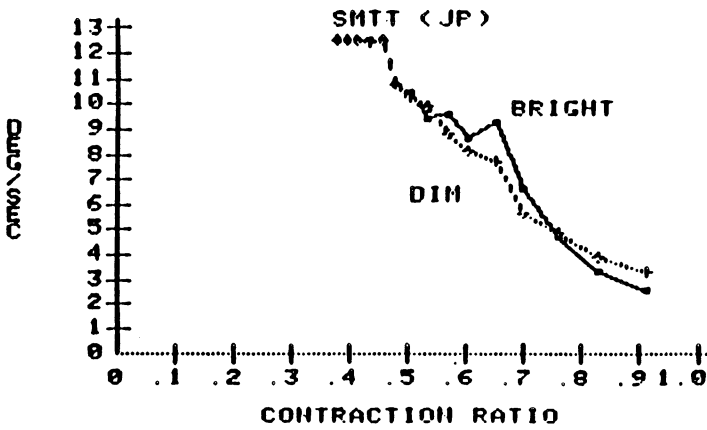
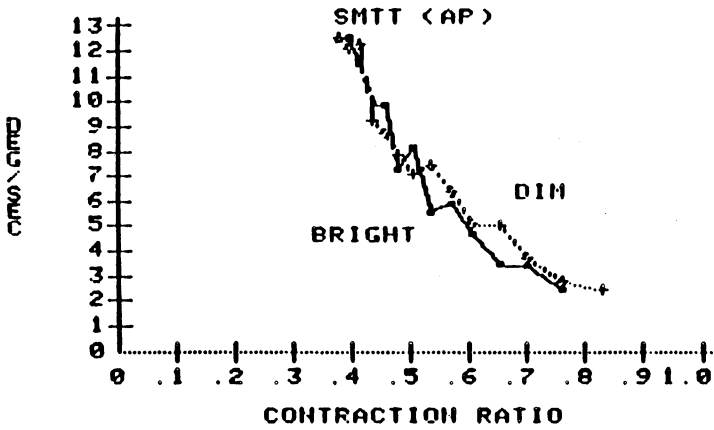


Figure 14.3

Experiment 2: Adjusted stimulus rates plotted against contraction ratios needed to achieve a perceived width-to-height match for each of four subjects. Ambient illumination = 2.0 foot-candles (DIM) and 8.0 foot-candles (BRIGHT). Source: Ibid.

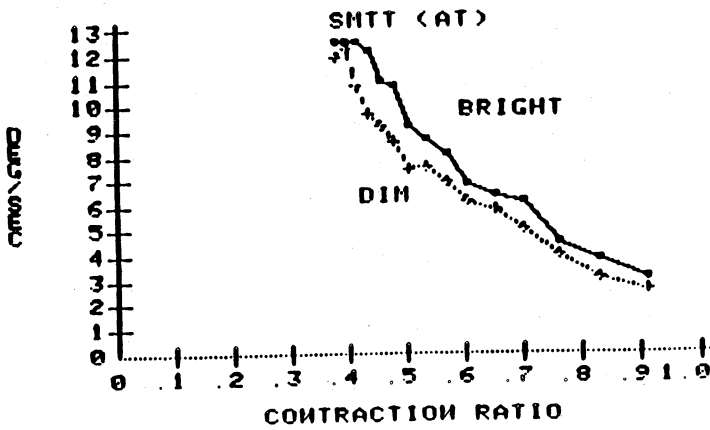
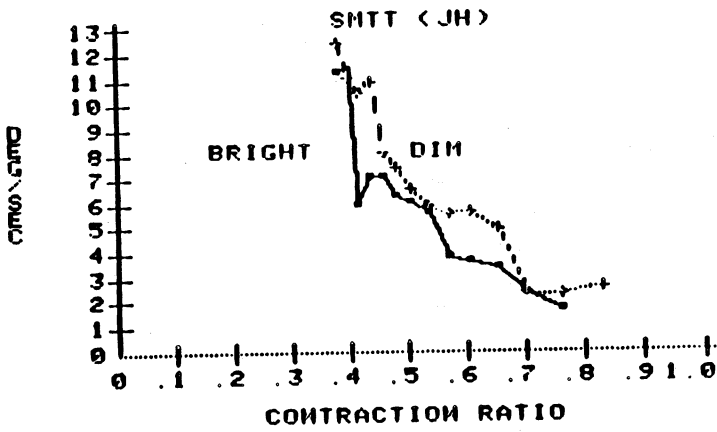


Figure 14.3 (continued)

Results In this experiment, the independent variable was the level of ambient illumination. The height-width ratio of the stimulus pattern, as well as its velocity, were held constant. The dependent variable was the number of successive stimulus sweeps required to squelch the SMTT effect. Figure 14.4 shows the results plotted separately for each of the four subjects. Hypothesis 4 was confirmed by a statistical test of the differences in the response curves obtained from each of the subjects under the two illumination conditions.

Comments

SMTT provides a strong argument against a static representation in visual short-term memory and argues for a dynamic form of postretinal storage (McCloskey and Watkins 1978). The retinoid system (Trehub 1977) offers a biologically plausible model of what such a dynamic storage mechanism would look like. It explains the SMTT effect as an epiphenomenon of the retinoid system and makes several confirmed predictions about the effects of changing parameters in the standard paradigm.

The retinoid model predicted an effect of ambient illumination on perceived width in SMTT (experiment 2, hypothesis 3) that was not confirmed. According to the model, an increase in the level of ambient illumination should increase the perceived width of a stimulus pattern by adding excitatory bias to autaptic cells in the retinoid network. This, in turn, should increase, to some extent, the rate of translation of the autaptic representation of the figure on each of its sweeps across the viewing aperture. While this assumption of the model might be incorrect, it is also possible that the difference in ambient illumination in the experiment was too small to exert a reliably detectable influence on translation velocity over the retinoid at each sweep. In contrast, the prediction of resistance to SMTT squelching under higher ambient illumination (experiment 3, hypothesis 4) was confirmed, and this depends on the same assumption that an increase in ambient illumination will add excitatory bias to autaptic cells in the retinoid network. However, the effect observed in experiment 3 does not depend on the ambient light intensity at each single sweep but on a cumulative influence over the entire trial (typically thirty to several hundred sweeps). Thus, it appears that ambient illumination has a real but relatively weak influence on retinoid dynamics.

A New Visual Illusion

In the standard SMTT demonstration, a rectangular aperture provides the observer with a continuous sequence of fragmentary views of a moving stimulus pattern. Because of the rectangular shape of the slit in the occluding screen, the width of the stimulus field is uniform along the vertical axis. This results in the direct stimulation (via retinal afferents) of a uniform number of autaptic cells (μ) on each horizontal plane along the vertical axis of the retinoid. Suppose that the aperture were triangular in shape. In this case, the number of autaptic cells that would be directly stimulated (μ) by the oscillating figure would be much greater at the base of the triangular aperture than at its vertex. Since the number of μ collinear with the horizontal direction of stimulus motion will change as a function of the triangular shape of the aperture and since $F_{\dot{\mu}} > F_{\mu}$, the rate of horizontal translation of a retinoid representation should increase from the vertex to the base of the aperture. This property of retinoid dynamics implies that if a rigid shape were to be presented as an SMTT stimulus behind a triangular aperture instead of a rectangular one, the observer should experience an illusion of a figure swinging in pendular fashion as though pivoting near the vertex of the triangle rather than moving in simple reciprocating translation (the veridical situation).

The predicted pendulum illusion was confirmed (Trehub 1985). It can be clearly seen with simple materials. Near the top of a strip of white cardboard, draw a heavy black circle approximately 25 millimeters in diameter with a vertical line through its center. On a sheet of gray cardboard, cut out an isosceles triangle approximately 30 millimeters high and 15 millimeters wide. Holding the gray cardboard screen at normal reading distance, position the circle on the white strip behind the triangular aperture and slide it laterally back and forth at a rate of about 2 cycles per second, taking care to move the circle beyond the aperture in each direction. If the center of the cutout is fixated, an egg-shaped figure will be perceived swinging like a pendulum within the aperture (figure 14.5). The apparent contraction of the figure along the axis of motion occurs because the rate of image translation on the retinoid is less than the veridical rate of stimulus motion. Moreover, in accordance with the retinoid model, the rate lag is greater near the vertex (narrower aperture) than near the base (wider aperture), which accounts for the egg-shaped transformation of the circle.

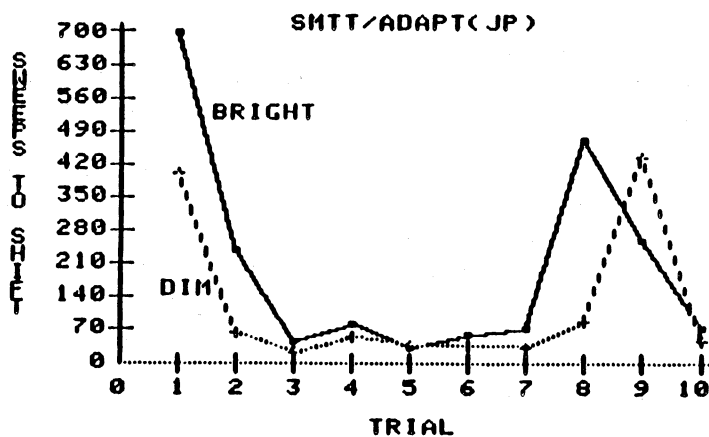
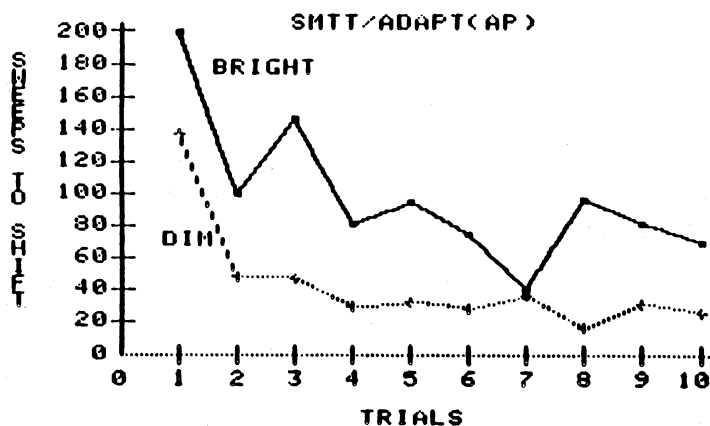


Figure 14.4

Experiment 3: Plot of the number of successive stimulus sweeps required to squelch the SMTT effect over 10 successive trial blocks for each of four subjects. Ambient illumination = 2.0 foot-candles (DIM) and 8.0 foot-candles (BRIGHT). Source: Ibid.

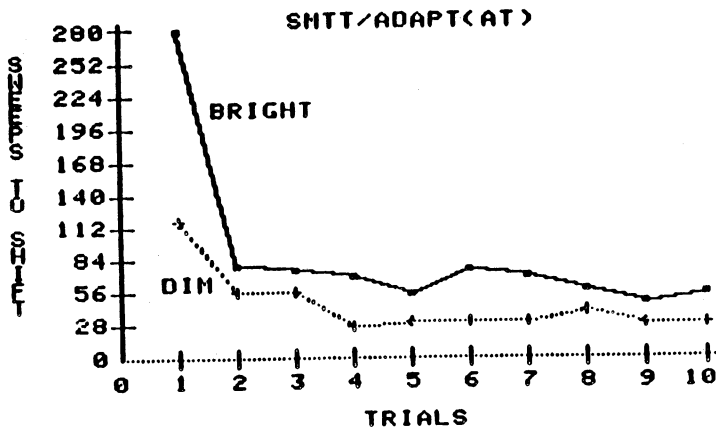
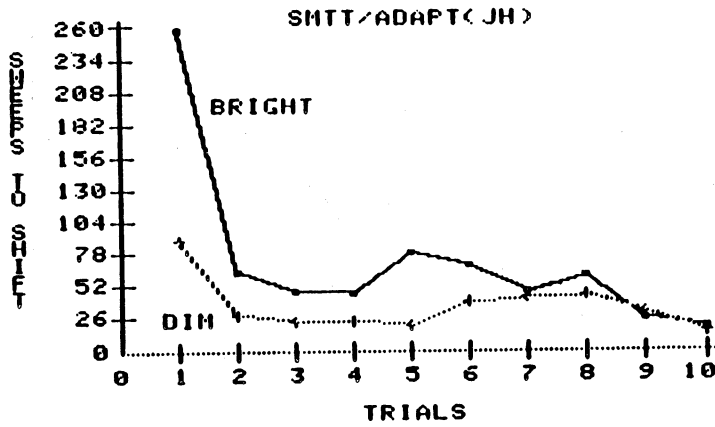


Figure 14.4 (continued)

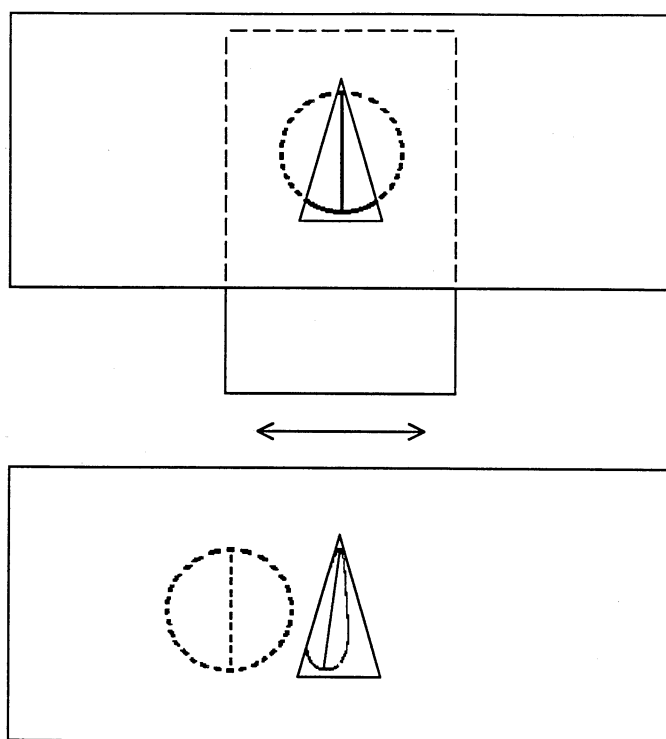


Figure 14.5

Top: Opaque screen with triangular aperture through which a part of the stimulus pattern (circle with vertical line) can be seen. *Bottom:* Appearance of pendulum illusion when stimulus is moved back and forth at approximately 2 cycles per second.

The Moon Illusion

Surely the oldest and one of the most striking and puzzling illusions—a false perception on a truly cosmic scale—is the so-called moon illusion (Reimann 1902, Schur 1925, Holway and Boring 1940, Boring 1943). In this illusion, the moon, or any other celestial body, when viewed at or near the horizon is perceived as much larger than when it is at its zenith, though its true visual angle (~ 0.5 degree) remains unchanged. The kind of process that might account for this phenomenon, even in the most general terms, has long been a subject of vigorous dispute (Boring 1962, Rock and Kaufman 1962, Wallach 1962, Reed 1984). The debate has remained in the arena of psychophysics, and, to date, no theoretical explanation of the illusion has been proposed at the level of neuronal structure in the visual system.

Here I will show how a plausible anisotropy of the 3-D retinoid system provides a simple biophysical explanation of the moon illusion.

Explaining the Moon Illusion

A guiding premise motivating the models that I have proposed is that the structure of the human cognitive brain has been shaped by evolution to cope with the ecologically significant demands of the human environment. In this process, we can imagine a principle of neuronal economy at work. We would expect limited genotypic neuronal resources to be allocated for specific kinds of cognitive representation and computation in rough accordance with their importance to the survival of the species. In the visual system of lower animals, there are clear examples of this principle (Ingle 1968, Fite 1976, Spinelli 1987). If the location of objects in particular regions of egocentric space is a matter of little ecological importance, then we might expect a reduction in neuronal resources devoted to the representation of those regions, the savings invested (by evolutionary competition) in more useful brain structures.

Can regions of human egocentric space be characterized as having more or less ecological value? If so, how might this fact shape those genotypic visual mechanisms that must be involved in the moon illusion? While we cannot assign precise ecological value to sharply defined extrapersonal regions, there are considerations that persuasively suggest the relative ecological value of broadly specified regions in egocentric space. For example, the space within the immediate reach of an individual is critically important for a wide variety of vital behaviors. Beyond that, space within the range of quick natural locomotion contains objects and features that may demand immediate response. As the regional locus of terrestrial space increases in distance from the observer, the objects contained tend to diminish in ecological urgency. Thus, the brain's reduction in distance discrimination with the increasing distance of visual targets entails no serious loss of the ability to cope in the natural environment. In addition, with respect to the location of objects with ecological significance, as egocentric distance increases, the maximum vertical angle of regard tends to decrease. For example, an apple on a tree branch that is within reaching distance may be directly overhead, a vertical sighting angle of 90 degrees; but any earthbound object at a horizontal distance of 1000 feet would have to be 1000 feet tall at the point of fixation to demand an angle of regard of even 45 degrees.

If evolution has conserved neuronal resources in mechanisms for 3-D viewing of horizon-limited terrestrial space, why should it not devote even fewer biological resources for the 3-D discrimination of

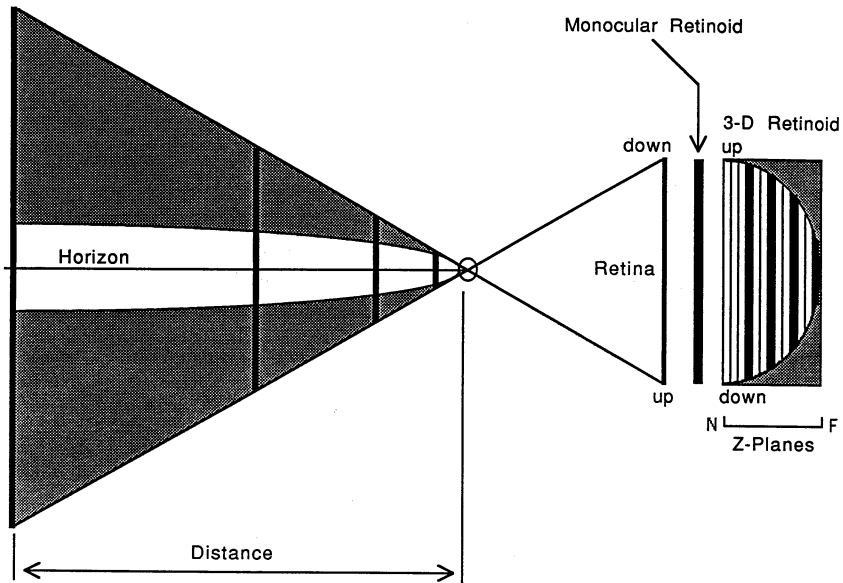


Figure 14.6

Schematic illustration of anisotropic representation of egocentric space within 3-D retinoid. Shading indicates the region in egocentric space that is represented in the 3-D retinoid on progressively nearer Z-planes as the vertical angle of regard departs from the horizon and as distance increases. Shaded area in 3-D retinoid indicates absent neuronal substrate for spatial representation.

objects in extraterrestrial space? After all, these are phenomena that, except for the briefest (modern) period of human experience, were beyond the scope of possible contact and demanded only philosophical speculation. I argue that egocentric space is anisotropically represented in the brain's 3-D retinoid system so that representations of objects within regions of increasing distance are progressively collapsed onto nearer Z-planes as the elevation in their lines of sight departs from the plane of the normal horizon (figure 14.6). Thus, valuable neuronal tissue is conserved for more important matters.

Recall the connectivity scheme from the 3-D retinoid to the mosaic cell array that was discussed in chapter 5 (figure 5.8). Principal cells in the nearest Z-plane are mapped to their corresponding retinotopic coordinates in the mosaic cell array while the axons of cells in the more distant Z-planes diverge to project to increasingly more eccentric coordinates on the mosaic cell array in accordance with the visual distance they represent. This provides a neuronal circuit for magnifying (or reducing) images in rough compensation for the change in retinal size at different viewing distances.

Since the distance of the moon from an observer on earth, as well as its retinal size, remain effectively unchanged whether it is viewed at the horizon or at its zenith, one might not expect a mechanism for size constancy to play a role in the moon illusion. However, the hypothesized brain architecture predicts size constancy (Emmert's law) (Boring 1940) for all vertical angles of regard at all distances only if it is assumed that the structure of the 3-D retinoid is functionally isotropic over all Z-planes. If the combined effect of viewing distance and angle of regard results in the shift of an excitation pattern of constant size from a distance Z-plane in the 3-D retinoid to a nearer Z-plane, then the axonal projection of the image on the mosaic cell array will shrink correspondingly in size. I propose that this is precisely what happens when the moon is seen after it has departed from the horizon (figure 14.7). Because of the great distance of the moon, it is always represented on the most distant Z-plane in the 3-D retinoid that can provide a neuronal medium for the moon's current angular position in egocentric space. As the moon rises above the horizon, the most distant Z-planes that have a substrate to represent its increasing elevation correspond to progressively nearer regions of 3-D space. Thus, the image projected to the mosaic cell array is that of a rising moon diminishing in size.

Comments

The neuronal architecture of the 3-D visual system described can not only explain the moon illusion but can account as well for at least two other previously unexplained classical psychophysical observations. Even before the writings of Ptolemy, many observers reported that the "dome of the sky" appeared to be flattened from horizon to horizon (Smith 1738). The vertical distance of the sky is perceived to be closer than its horizontal distance. But notice that this perception is just what one should expect if the properties of the 3-D retinoid model accurately reflect the corresponding visual-cognitive mechanisms in the human brain. Since egocentric space is represented within the limits of the 3-D retinoid, the outer representational bounds of the retinoid structure will define the outer bounds of perception of the limitless sky. Thus the anisotropy of retinoid space that is characterized by representation on closer Z-planes as the vertical angle of regard is elevated will result in a corresponding anisotropy of spatial perception when the sky is scanned. The sky above will appear closer than the sky beyond the horizon.

One might raise the objection that the phenomenon is experienced even in a starless and cloudless (empty) sky. What then would be the targets of visual regard? And if there were no visual stimuli repre-

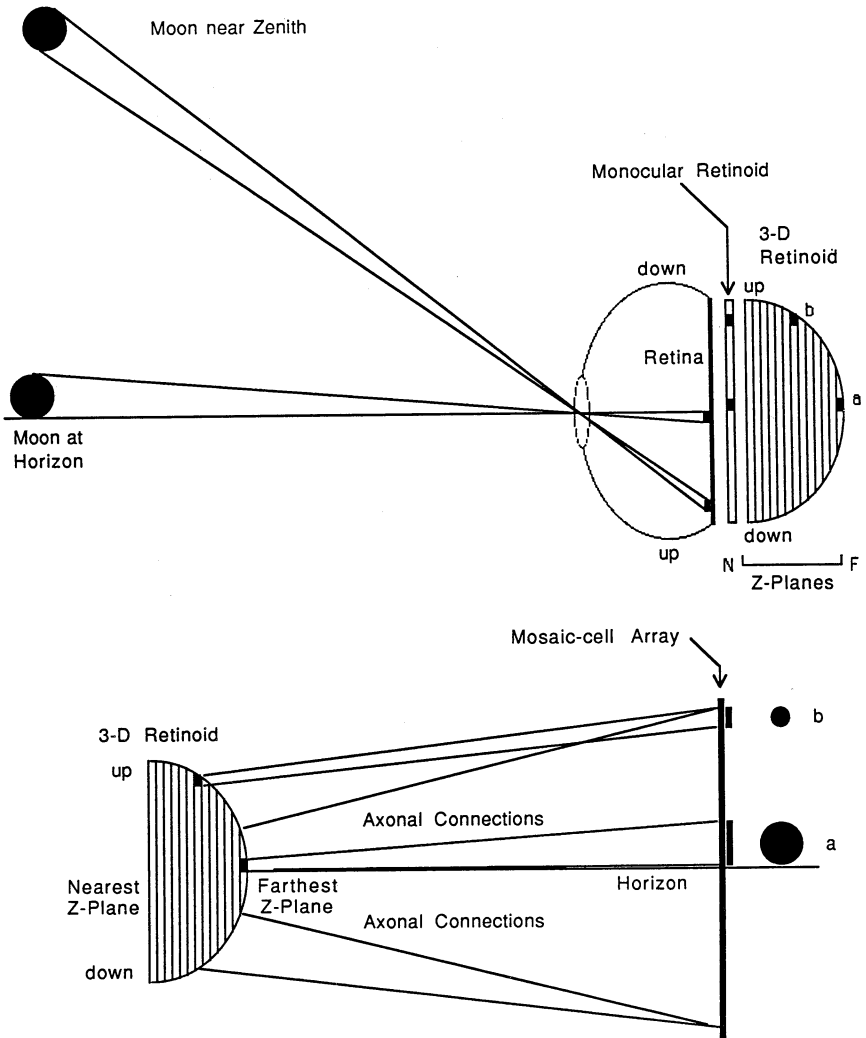


Figure 14.7

Moon illusion. *Top:* N = near; F = far. Moon at horizon is represented on farthest Z-plane (a). Moon near zenith is represented on nearer Z-plane (b), which is the farthest Z-plane available for the representation of an object at its distance and elevation. *Bottom:* Expanding axonal connections from Z-planes in 3-D retinoid to mosaic cell array that normally preserve size constancy evoke a larger representation of the horizon moon (a) than the zenith moon (b).

sented in the anisotropic space of the 3-D retinoid system, how could such a brain structure shape our perception of the sky? But notice that even when we scan an empty sky, binocular convergence relaxed and lens accommodation at infinity, the locus of our attention moves in accordance with our visual search. In the retinoid model, shifts in attention are real neuronal events—excursions of the self-locus (autaptic cell discharge) over selected coordinates of retinoid space. When we visually scan an empty sky, neuronal excitation provided by the corresponding excursion of the self-locus traces the outer bounds of retinoid space—closer as the line of sight approaches the vertical axis, more distant as the line of sight approaches the horizontal axis (the horizon).

The same anisotropic properties of the retinoid model and the circuitry for size constancy explain the results of an early psychophysical investigation of perceived size and angle of regard reported by Schur (1925). Working in a large, dark room, she presented two projected disks of light to subjects, one in a horizontal direction and the other above the subject and always at the same distance as the horizontal disk. Both stimuli were presented at distances varying from 3 to 33 meters. It was found that as the viewing distance increased, there was an increase in the perceived size of the stimulus in front of the observer relative to the stimulus above. The advantage in the ratio of perceived size for the frontal stimulus relative to the overhead stimulus was 1.16 to 1 at 3 meters, 1.71 to 1 at 22 meters, and 1.87 to 1 at 33 meters. These results are consistent with my assumption of an anisotropy in the representation of visual distance within the brain (figure 14.6), which is related to elevation of gaze and is based on the argument of evolutionary economy.

Other Visual Illusions

Many other kinds of visual illusions can be explained by the intrinsic properties of the retinoid system and accessory mechanisms. A few more examples will show how a variety of other illusions that have had no satisfactory physical explanation will occur as natural consequences of the structure and dynamics of the proposed brain model.

The Waterfall Illusion

When we look at a stationary pattern of vertical bars under ordinary conditions, we see no more than the stimulus itself—a stable pattern of vertical bars. If we look at the same stimulus after prolonged viewing of a pattern moving, for example, to the left, the stationary bars are perceived as streaming to the right. The illusion of motion of the

stationary test pattern is an aftereffect of a prior moving stimulus and is always in a direction opposite to that of the prior inducing motion. This phenomenon is commonly called the waterfall illusion. (If the motion of the inducing stimulus is upward, the stationary stimulus is seen as streaming down like a waterfall.)

According to the retinoid hypothesis, the illusion of motion is experienced because the neuronal representation of the stimulus pattern evoked on the retinoid actually moves a small distance in the direction perceived, though the stimulus itself is stationary. The reason that it happens is a straightforward consequence of retinoid structure, the known properties of the ocular-motor system, and the physiology of nerve cells.

Even while we are fixating a visual target, our eyes are in constant motion, which can be characterized as a high-frequency, low-amplitude tremor. The tremor movements occur with a magnitude of less than 0.5 minute of arc and at a rate of up to 150 cycles per second (Ratliff and Riggs 1950; Riggs, Armington, and Ratliff 1954). Movements of the eye normally drive shift control cells in the retinoid, with the result that a retinoid image is translated in the direction of eye movement. However, the micro-oscillations of the eye change direction at high frequency; thus, under normal conditions, their directional influences on the retinoid offset each other during fixation so that there is no false representation of stimulus motion. But if one fixates a stimulus that is in constant unidirectional motion for a prolonged period, the small interneurons within the retinoid that are part of the activated translation circuit become selectively fatigued (Shepherd 1979) and a significant bias is induced in the dynamic properties of the retinoid. The bias consists of a loss of efficacy for translation in the direction served by the fatigued interneurons. It is this bias, caused by the moving induction stimulus, that produces the waterfall illusion.

Figure 14.8 illustrates what happens in the retinoid when the illusion is induced. After shift-left interneurons have been fatigued by prolonged fixation of a pattern in constant motion to the left, the stationary test pattern appears to stream to the right. This occurs because the phase of eyeball tremor that would normally drive the retinoid representation of the pattern in offsetting motion to the left does not elicit a corresponding response from the fatigued shift-left interneurons. At the same time, right-going movements of the oscillating eyes are effective in eliciting small successive right-going translation of the retinoid image, which decays as it traverses more autaptic cells beyond the microfields of direct excitation by stimulus contours.

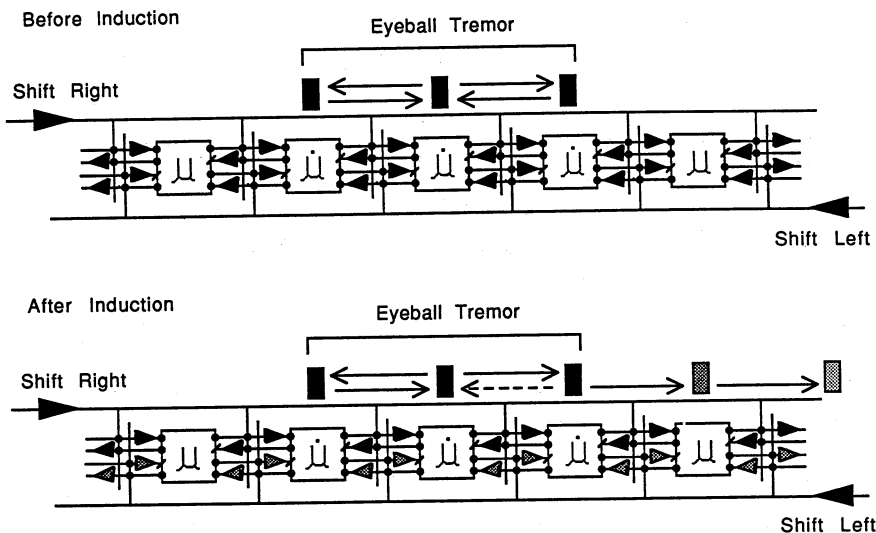


Figure 14.8

Waterfall illusion. *Top*: Normal eyeball tremor shifts a stationary stimulus back and forth over adjacent autaptic cells with no net motion of the stimulus in either direction. *Bottom*: During the induction phase, after prolonged viewing of a pattern in constant motion to the left, shift-left interneurons are fatigued (stippled cells). Given a stationary test pattern after induction, normal tremor-related shifts of the pattern to the right are not offset by corresponding shifts to the left, resulting in an illusory streaming of the stimulus to the right. Streaming decays as the illusory image traverses autaptic cells beyond the field of direct stimulus excitation.

Illusory Motion as a Resultant of Orthogonal Induction

In the standard waterfall illusion, the direction of illusory motion is always opposite to the motion of the inducing stimulus; it is rotated 180 degrees from the angular direction of the inducing motion. Thus, motion to the left (180 degrees, where up is 90 degrees and down is 270 degrees) produces the illusory aftereffect of a pattern moving to the right (0 degree).

Suppose one were exposed to two inducing stimuli moving alternately in orthogonal directions. What would be the direction of the illusion? On the basis of the retinoid explanation proposed, we should be able to predict the perceived motion of the stationary test pattern. The directions of the orthogonal induction patterns should evoke discharge in retinoid shift control cells at a frequency proportional to the vectorial component of stimulus motion for each of the four shift control directions. The relative degree of fatigue for each set of directionally oriented interneurons will then correspond to the

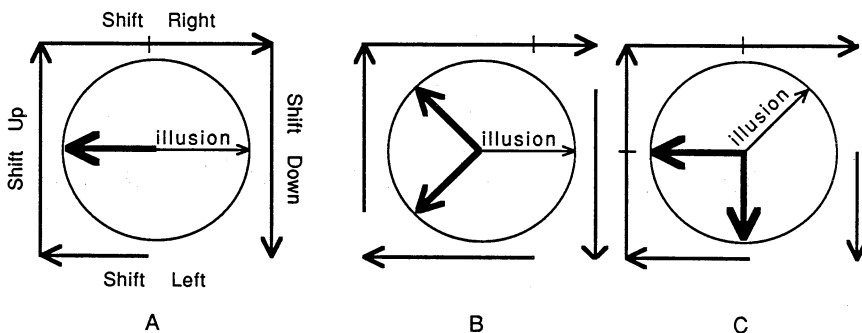


Figure 14.9

Retinoid states after simple and orthogonal induction according to theoretical model. Relative length of each of the outer four directional shift lines indicates relative efficacy of interneurons serving translation in the corresponding direction. Distance from short tick mark to arrow head indicates net advantage in translation efficacy for the direction shown. Within circles, thick arrows indicate direction of inducing motion and thin arrows the expected direction of illusory motion. *A*: Direction of illusion following simple induction. *B*: Direction of illusion following orthogonal induction of 135 degrees and 225 degrees. *C*: Direction of illusion following orthogonal induction of 180 degrees and 270 degrees.

relative frequency of discharge in the shift control cell that drives that set of interneurons. Thus, perceived motion of the stationary test pattern should stream in a direction that is the resultant of the relative efficacy of each set of interneurons considered as a vectorial component for retinoid translation. Figure 14.9 illustrates how the directions of inducing stimuli jointly determine the direction of translation on the retinoid and therefore the direction of illusory aftermotion.

In an experiment reported by Riggs and Day (1980), a visual after-effect of motion was induced by bars moving alternately at two different angles in the frontal plane. During the 10 minutes of induction time, the angle of motion was alternated every 10 seconds. When the angles of inducing motion were 135 degrees and 225 degrees, the direction of illusory aftermotion was 0 degree. When inducing motion was presented at 180 degrees and 270 degrees, the direction of illusory aftermotion was 45 degrees. These results are exactly as predicted by the retinoid model (see examples B and C in figure 14.9).

The Mueller-Lyer Illusion

One of the classical illusions of visual extent is the Mueller-Lyer illusion (Mueller-Lyer 1889, Coren 1986). If two line segments of equal length are each attached at both ends to short angles, one line to the vertices of angles pointing out and the other to the vertices of

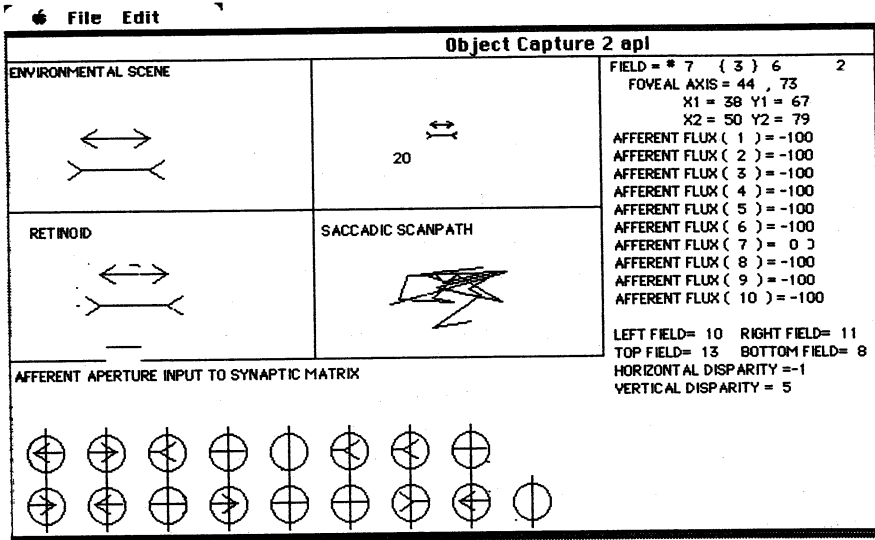


Figure 14.10
Simulation of Mueller-Lyer illusion. Stimulus parts captured on successive saccades are displayed in the bottom frame. Circles indicate afferent aperture. Vertical lines bisecting circles represent vertical plane along normal foveal axis.

angles pointing in (top left frame of figure 14.10), the line with the angles pointing out is perceived as shorter than the other. This illusion can be explained as a normal consequence of the mechanisms for parsing aspects of the visual environment, marking location, and representing distance between objects.

In a comparison of the length of the two lines in the Mueller-Lyer paradigm, the length of each line from one terminus to the other must be registered and the two linear extents compared. Chapter 7 shows how this is done by excursions of the self-locus that are registered in neuronal comparator circuits. The illusion arises because of errors in locating the ends of each line, which are systematically caused by the shapes of the stimuli used in the illusion. The direction of the appended angles biases the centroid location of the terminating line segment. When the end angles point out, the terminating centroids of the attached line are closer together than they are when the end angles point in. This results in a mismatch of the two self-locus traces that are made when the lengths of the lines are compared.

Figure 14.10 shows the result of a simulation of the Mueller-Lyer

illusion. Standard stimuli (top left frame) were presented to the model visual system. Twenty saccades were allowed (middle frame). The bottom frame shows the centroids of stimulus segments that were captured over the course of 20 saccades. For the figure with the angles pointing out, extreme right and left centroids were closer to the center of the figure than they should be. For the figure with the angles pointing in, extreme right and left centroids were either at the terminus of the horizontal line or somewhat beyond it (farther from the center of the figure than they should be). The direction of the results is consistent over a variety of trials using different afferent apertures and initial error tolerances, although the magnitude of centroid error is reduced with smaller apertures and tolerances. However, the trade-off in such cases is that for a small improvement in the accuracy of registering the ends of the horizontal lines, there is a great increase in perceptual effort. Many more saccades must be made and a more extended sequence of retinoid hunting must be tolerated to achieve a modest reduction in the extent of the illusion.

Ambiguous Shapes

Certain well-recognized shapes evoke unstable perceptions. At one moment the stimulus is clearly recognized and classified as a particular kind of object, while at the next moment it is just as clearly seen as a completely different kind of object. I propose that the ambiguity in the perception of such shapes is systematically related to the part of the shape that falls on the normal foveal axis at the moment of visual capture. Given the hypothesized visual mechanisms for finding a centroid and aligning it on the normal foveal axis, if an extended shape has multiple centroids and if the constituent parts within the afferent aperture for the different centroids are also distinctive parts of different objects or different exemplars of an object, then the object that is "seen" at any moment will depend on the particular centroid of the extended shape that has been aligned to the normal foveal axis at that moment. With stimulus patterns having multistable centroids, small saccadic shifts will normally result in the capture of different centroids when a change in gaze is of sufficient magnitude.

The Necker Cube

A common example of an ambiguous figure, in this case one that evokes the perception of bistable perspective, is the Necker cube (Necker 1832). The stimulus is a 2-D projection of a cube with no hidden edges (top left frame of figure 14.11). As the cube is viewed, it is typically seen to "jump" unpredictably between two perspec-

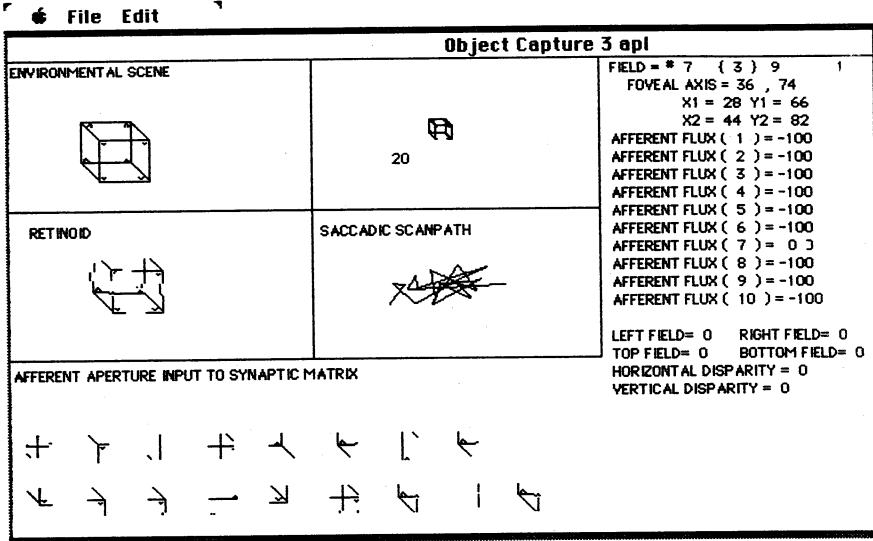


Figure 14.11
Simulation of response to Necker cube. Stimulus parts captured on successive saccades are displayed in the bottom frame.

tives. One is a cube, seen from above front, with the lower square as its front plane and the upper square as its back plane. The other perspective is a cube, seen from below front, with the lower square as its back plane and the upper square as its front plane. Confronted with the same stimulus, will the model visual system perform in a way that would constrain perception to one or the other of the two geometric interpretations, and will there be natural alterations in the captured patterns that can account for the shifting perspective?

Figure 14.11 shows the result of a simulation with the Necker cube as the visual stimulus (upper left frame). In order to facilitate interpretation of the results, the two squares that form the front or back faces of the cube are marked at their corners. The lower square has a small V at each corner. The upper square has a small inverted V at each corner. Twenty saccades were allowed. The center frame displays the scanpaths of all saccades. The bottom frame displays the parts of the stimulus that were successively captured by the visual system during the 20 saccades. A partial reconstruction of an image of the whole cube on the scene assembly retinoid is displayed in the left center frame.

Each captured pattern that represents a corner of the cube uniquely determines one of the possible visual perspectives on an opaque cube

(the exterior form of a solid cubic object). Whenever a corner is captured within the afferent aperture and projected to the synaptic matrix, any image of a remembered (previously learned) cubic object that is evoked in the imaging matrix must correspond in perspective to the angular properties of the current effective stimulus (the captured corner). When the eye moves, the current effective stimulus can change, and a different corner can be captured. Thus, the two different perspectives on the cube can each be evoked at different times even though the whole stimulus pattern is static.

In this simulation run, there were 17 captures of parts of the cube over the course of 20 saccades. We can see from the bottom frame of figure 14.11 that the model first "perceived" a cube from above (first corner captured, top row). At the fifth captured part, the perspective changed to a cube seen from below (second corner captured, top row). This perspective was maintained until the ninth capture (bottom row) and reversed again at the fifteenth capture. Thus, the perceptual experience of a "spontaneously" shifting perspective in the Necker cube paradigm can be explained as a natural consequence of visual mechanisms embodied in the proposed brain model.

Figure Reversal

In the case of the Necker cube, the part of the figure that had been captured (centered within the afferent aperture) was the effective stimulus that determined the perspective from which the figure was viewed. Other kinds of figures will evoke the recognition of entirely different objects depending on which of their parts are captured as effective stimuli. A well-known example is displayed in the top left frame of figure 14.12. When this drawing is viewed, there is typically a spontaneous alternation in the perception of the figure. One moment it is recognized as a vase and at another moment as two faces. The explanation for such figure reversals is that perception is constrained by the effective stimulus within the afferent aperture. Since the parts of the figure around different centroids are proper parts of different well-known objects and because small saccades will often lead to the alignment of different centroids on the normal foveal axis for projection to the detection matrix, the particular object that is perceived will vary from time to time within a common visual field.

The simulation shown in figure 14.12 displays the responses of the model visual system in the vase-face paradigm. The pattern in the upper left frame is the stimulus presented to the system. Saccades are displayed in the center frame. The bottom frame displays the parts of the stimulus that were successively captured by the visual

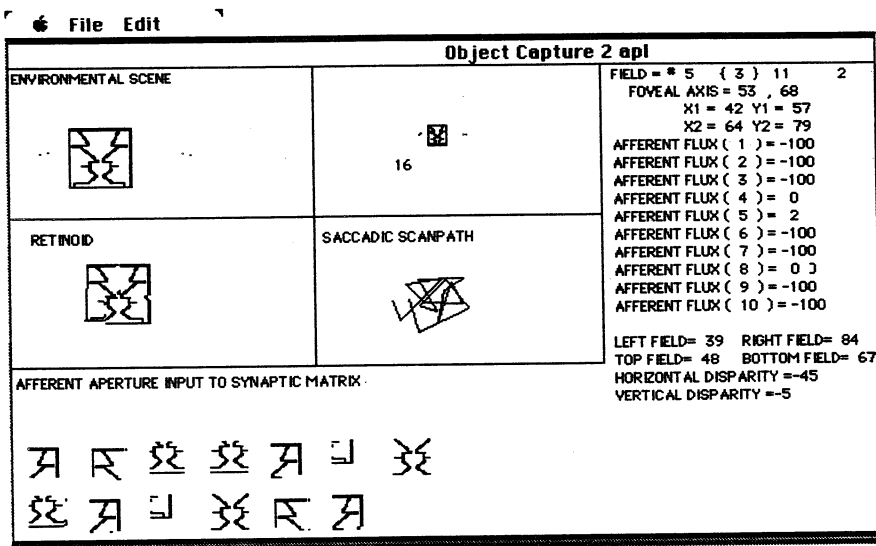


Figure 14.12

Simulation of response to ambiguous figure: Vase-face reversals. Stimulus parts captured on successive saccades are displayed in the bottom frame.

system during the simulation. A reconstruction of an image of the stimulus is displayed in the left center frame.

The sequence of the stimulus parts that were captured (effective stimuli) and projected to the synaptic matrix is consistent with the perception of either of two different objects appearing at differing times within the same visual field. For example, the first pattern that was captured (bottom frame, figure 14.12) corresponds to the brow, eye, and nose in the profile of a face looking to the left. Next is the same profile facing right. This is followed by two successive patterns corresponding to a large part of a vase, and so forth. Although the whole figure is ambiguous, its component parts, captured at different centroids, are unambiguous and clearly associated with different objects. Thus our perception of the object that is "depicted" shifts back and forth between a vase and two faces.