Chapter 4

Modeling the World, Locating the Self, and Selective Attention: The Retinoid System

When we look at the world in front of us, we see a stable, coherent arrangement of objects and environmental features in a spatially extended layout. And when we recall a previously experienced scene, we visualize it in an appropriately integral fashion. Yet it is a physical fact that on any fixation, our window of sharp foveal vision allows us to register clearly only 2 to 5 degrees (at the most) of an extended frontal scene. Nevertheless, by a series of saccadic eye movements, we are able to construct in our brain a properly laid-out representation of the scene as observed from a particular viewpoint. Given that saccadic eye movements present a viewer with a series of scattered glimpses of a spatially extended visual environment, where all sharply defined visual impressions are superposed on the fovea, how can the visual system disentangle the overlaid fovea-centered images and construct an integrated veridical representation of the environment in an egocentric spatial frame?

This fundamental problem led me to hypothesize the existence and detail the neuronal structure of a dynamic postretinal buffer, which I call a retinoid, that in a layered system can register and appropriately combine disparate foveal stimuli into a proper unified representation of a larger real-world scene (Trehub 1977). In the process of modeling the retinoid, it became clear that, together with some relatively simple accessory mechanisms, it can perform many other important cognitive functions:

- Parsing objects in complex visual environments.
- Constructing visual representations of objects and scenes (veridical and/or hypothetical).
- Performing geometric and relational analysis of veridical and hypothetical objects or scenes.
- Locating a representation of oneself within a represented environment.
- Representing the paths of moving objects as well as self-excursion paths in complex environments.

- Effecting selective shifts of focal attention.
- Performing spatial translation of binocular images in stereopsis.

This module, like the retina, registers information in visual space and projects afferents to higher visual centers. It can organize successive retinocentric visual patterns into a unified egocentric or allocentric representation of object space. It serves as a visual scratch pad with spatially organized information stored as short-term memory. The mechanism of storage is assumed to be a retinotopically organized array of excitatory autaptic neurons (Shepherd 1979, van der Loos and Glaser 1972). In the retinoid, an autaptic cell that receives a transitory suprathreshold stimulus will continue to fire for some period of time if it is properly biased by another source of subthreshold excitatory input (arousal). Thus, a sheet of autaptic neurons can represent in its sustained discharge pattern any transitory organized input for as long as diffuse priming excitation is sustained (up to the limit of cell fatigue). If the priming background input is removed or sufficiently reduced, the captured pattern on the retinoid rapidly decays.

Figure 4.1 shows a retinoid composed of an array of autaptic neurons connected by a balanced grid structure of excitatory and inhibitory interneurons. Axon collaterals of shift control cells are in excitatory synapse with selected groups of these interneurons. Any momentary suprathreshold input from an afferent visual array to its homologous autaptic retinoid cells will evoke sustained firing of the retinoid targets if there is a sufficient level of diffuse tonic bias. Thus, any retinal stimulation induces a comparable retinoid pattern of spatially organized discharge. At the same time, each active autaptic neuron induces a subthreshold, priming excitatory postsynaptic potential (EPSP) in each of the eight contiguous interneurons capable of eliciting excitatory and inhibitory potentials (IPSP) in their targeted autaptic cells. A primed interneuron that receives a sufficient increment of excitation from one of the shift control cells will fire and send spike input to its target cell. The retinoid behaves according to the following implicit rules.

1. If an autaptic cell that is not discharging (off) receives sufficient EPSP from an interneuron, it will fire (turn on).
2. If an autaptic cell that is on receives IPSP from an interneuron, it will turn off unless it receives simultaneously EPSP from another interneuron, in which case it will remain on.
3. If diffuse excitatory bias to the retinoid falls below a critical level, all cells in the retinoid turn off.

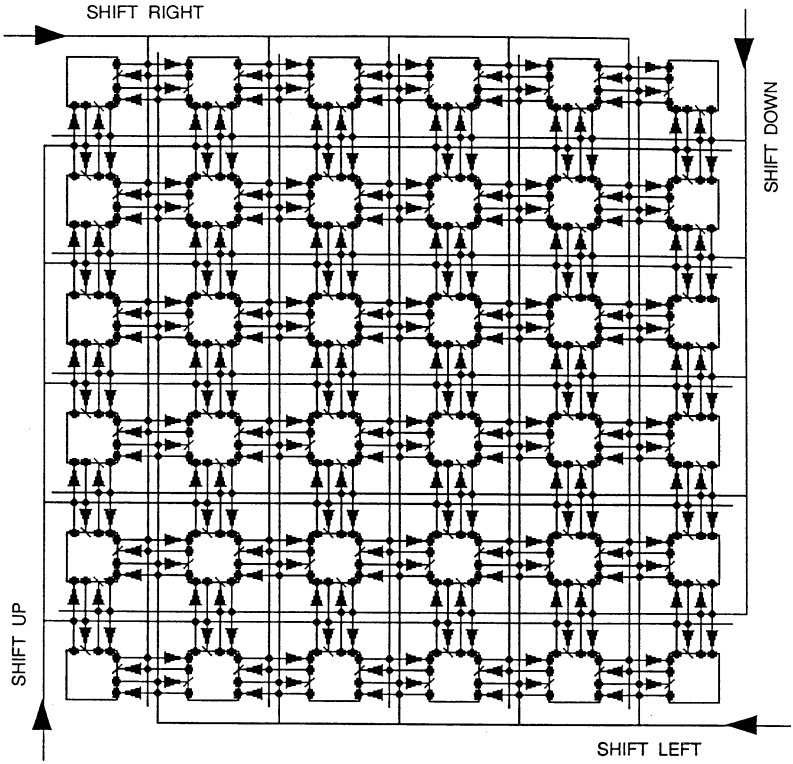


Figure 4.1
Translation retinoid. Squares represent autaptic cells serving short-term memory. Small, filled triangles between autaptic cells represent excitatory and inhibitory interneurons. Shift control cells designated by direction of effect.

Imagine that the sight of an object has evoked its pattern of retinotopic excitation on a normally biased retinoid. The visual pattern will be captured in short-term memory as a spatial analog on the retinoid; this captured pattern can be spatially translated in any direction by appropriate pulse (discharge spikes) from the shift control command cells. For example, a spike train from the shift-right line will transfer standing activity (via interneurons) from any active autaptic cell to the adjacent autaptic cell on its right and erase (via interneurons) the activity in the previously active donor (on the left) unless the donor is also receiving transfered excitation from its left-adjacent autaptic cell. Thus, sustained command discharges from the shift-right control cell will move the entire retinoid pattern to the right in successive increments of a single autaptic cell. Similarly, commands to shift left,

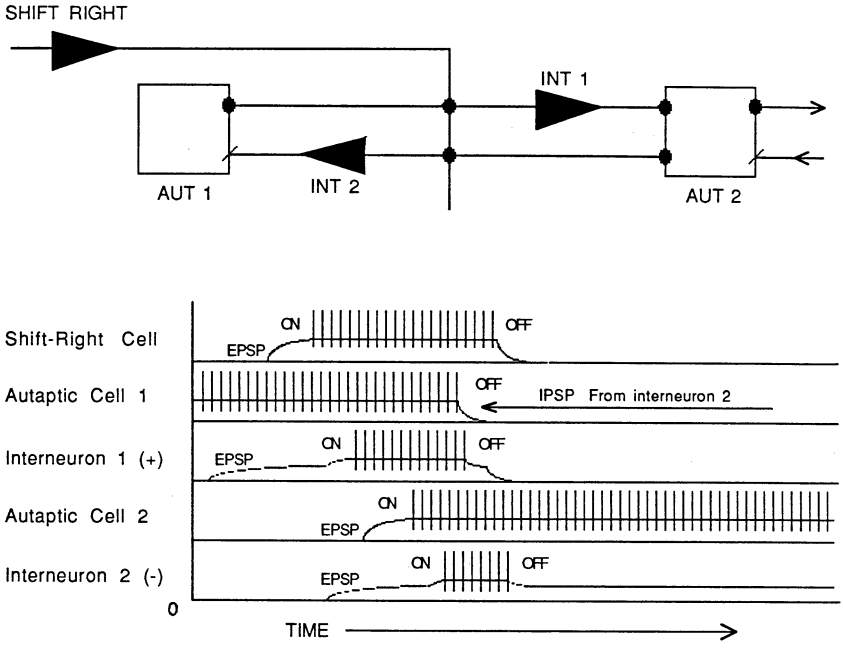


Figure 4.2
Top: Schematic showing two autaptic cells coupled by an excitatory (INT 1) and an inhibitory (INT 2) interneuron. With this arrangement, a pulse from the shift-right cell can transfer autaptic cell activity from left (AUT 1) to right (AUT 2) and inhibit the initial activity of AUT 1. *Bottom*: Illustration of the temporal course of EPSP integration, decay, and spike discharge in each of the five neurons depicted at top.

up, or down will move the captured pattern in the appropriate direction.

The schematic diagram of a local circuit within a retinoid shown at the top of figure 4.2 illustrates the synaptic connections involving a shift-right control cell at the junction between two autaptic cells and the two interneurons required to shift excitation on a retinoid from one autaptic unit to its autaptic neighbor on the right. The bottom part of the figure illustrates the sequence of EPSP and spike activity in each of the five neurons included in the circuit shown at the top as a stimulus is shifted from one autaptic cell to another. Temporally staggered directional pulses, such as right and up or left and down, will move the pattern in oblique directions, with the angle of translation depending on the relative frequencies of the component directional pulses. The higher the frequency is of the discharge spikes

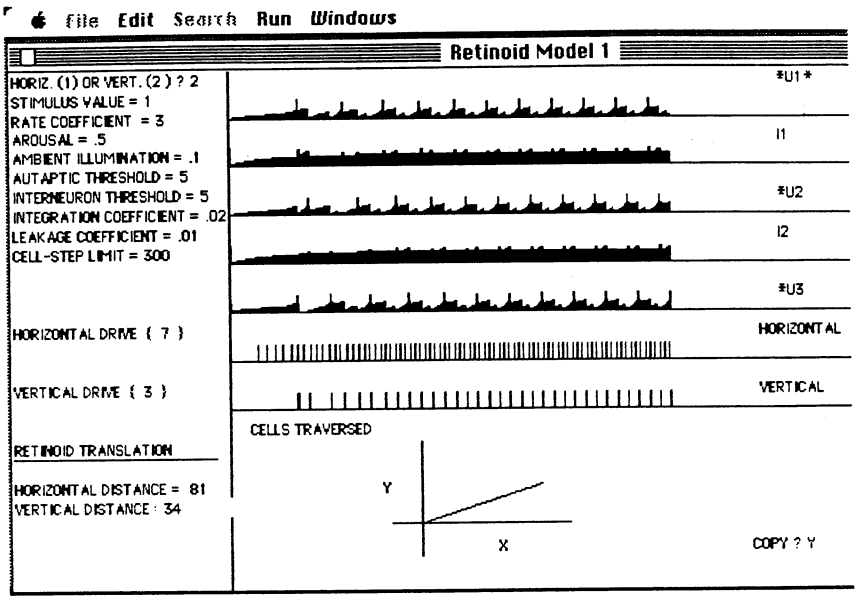


Figure 4.3
Screen printout of retinoid simulation. Lines labeled U show autaptic cell activity. Lines labeled I show interneuron activity. Relative shift-right rate = 7; relative shift-up rate = 3. Successive instances of autaptic cell transfer from left to right are shown on the line labeled HORIZONTAL. Successive instances of upward transfer are shown on the line labeled VERTICAL. In this simulation, initial retinoid activity was translated to a locus represented by a shift of 81 autaptic cells to the right and 34 cells in the vertical direction (shown on the X,Y plot at the bottom of the figure).

within the shift-control pulses, the more rapid will the retinoid representation move. The longer the pulse train is sustained, the greater will be the distance through which the representation is moved. Appropriate control pulse sequences of shift right/left and shift up/down can move a captured stimulus pattern to any position within retinoid space.

Figure 4.3 shows the results of a computer simulation in which a stimulus is translated to a position in the upper right quadrant of retinoid space by input from horizontal and vertical shift control cells. The final position (represented by the terminus of the oblique line within the X and Y axes at the bottom of the figure) is a function of the number of pulses on the shift control lines and the spike frequency within the directionally selective pulses.

Retinoid theory assumes that automatic spatial translation of retinoid patterns can occur when shift control neurons are driven by eye

and/or head movements and by the discharge of cells in lower-level visual mechanisms that detect the motion of objects in the field of view. Cells of the latter type are selectively tuned to both the direction and velocity of motion in the visual field (Barlow, Hill, and Levick 1964; Benevento and Miller 1981). (It should be noted in passing that the translation of images across retinoids is probably also induced by signals from the vestibular apparatus, but detailed consideration of this topic is beyond my scope.)

In the case of pattern translation induced by eye or head motion, the direction in which a representation is shifted on a retinoid sheet is determined by the direction of gaze; the extent to which it is shifted is proportional to the visual angle between the current fixation and the egocentric reference projection, which I call the normal foveal axis. This is the retinoid's coordinate of origin for the egocentric frame. The normal foveal axis corresponds to the line of sight of the fovea when the eyes are straight ahead, the head unturned, and the shoulders square with the upright body.

The Retinoid System

A single retinoid cannot serve as a neuronal substrate for constructing and maintaining in short-term memory a coherent and reasonably accurate representation of a spatially extended scene. Suppose that looking to our right we see a tree and that this pattern is captured on the retinoid and shifted in accordance with the angle of gaze to its appropriate egocentric coordinates. Now if we look leftward and see a barn, its image would also be registered and shifted appropriately on the retinoid. However, a serious problem is encountered at this point: the very same shift-left pulses that translate the image of the barn to its proper retinoid location would also move the image of the tree from its former correct position to an incorrect one (more to the left than it should be). The solution to this problem lies in a layered system of intercommunicating retinoid registers (figure 4.4). With several retinoids in homologous projection, selectively coupled and decoupled, an extremely powerful mechanism is provided that can both assemble a coherent representation of visual space and also perform a variety of other integrative and analytic tasks (Trehub 1977).

In this chapter, we will examine the properties of a retinoid system capable of capturing and restructuring representations of object layouts in a three-dimensional visual environment. But before we deal with operations in 3-D space, it is necessary to describe how a module

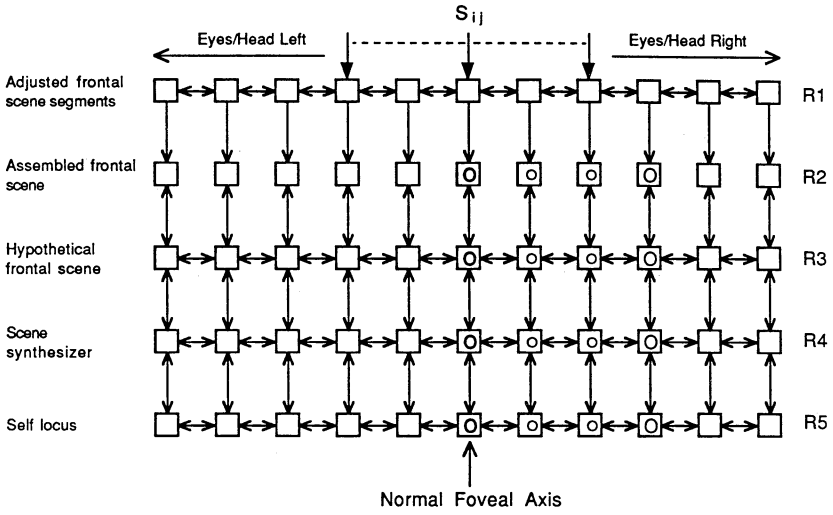


Figure 4.4
Stacked retinoid registers labeled R1–R5. Imagine a dimension of each retinoid array projecting orthogonally to the plane of the page in replicated autaptic cells. Arrows indicate the cells to which excitation can be transferred within the array. Self-locus designated by **0**. Heuristic self-locus designated 0. Excursion path designated o.

of layered retinoids processes inputs from the 2-D surface of the eye's retina.

Scene Assembly from a Two-Dimensional Projection

Let us assume a complete frontal scene to be all discriminable objects within an environment subtended by 180 degrees, with the vertex taken as the head in the normal frontal position. Since the receptive field for sharp monocular foveal vision is 2 to 5 degrees at most, it follows that the eye and/or the head must pivot (either in saccadic or smooth fashion) to scan the whole frontal scene. Thus, a sequence of excitation patterns will be evoked on the foveal region of the retina, and these must be represented in the brain in a way that conserves their real spatial relationships on the 2-D frontal plane.

Figure 4.4 shows a module consisting of a number of retinoid layers (R1–R5). The autaptic cells of each layer are in homologous and reciprocal projection (with the exception of an absence of R2 to R1 input) to the corresponding cells of the neighboring retinoids. The first layer (R1) receives retinotopic input (S_{ij}) from the foveal and near parafoveal region of the retina. Retinoid R1 is a translation retinoid in which the output of its shift control neurons is modulated by eye and/or

head position. It is assumed that the direction and extent of pattern translation on this layer is directly proportional to the degree of eye and/or head shift from the normal foveal axis. As each scene segment is registered on R1 and appropriately shifted in accordance with eye and head position, it is immediately transferred in its proper relative egocentric location to a second nontranslating retinoid (R2) and then erased on R1 by reset inhibition to all cells in the first layer. With successive fixations, a complete and homologously ordered representation of the frontal scene can be assembled on R2, and this information will provide the larger scene context for particular sensory inputs. For simplicity of illustration, an example of a 180 degree frontal scene assembly has been given; however, it is clear that similar mechanisms can assemble a 360-degree scene. Notice that by transferring the retinal image from a retinoid driven by eye and head movements (R1) to another that is decoupled from eye and head movements (R2), positional errors in object representation due to inappropriate effects of shift-control signals are avoided.

Hypothetical Scenes

The retinoid layer labeled R3 in figure 4.4 can construct representations of hypothetical environmental situations (scenes). It is a translation array that receives veridical information from the assembled frontal scene (in R2) and internally fabricated patterns from the retinoid complex labeled R4, called a scene synthesizer. The scene synthesizer consists of two translation retinoids coupled as a functional unit so that either can serve as a stable storage buffer while the other may be engaged in moving excitation patterns from one place to another in its representational space. The source of these patterns can be inputs from other retinoids, memories recalled and projected from the mosaic array in the synaptic matrix, or tracings created by movements of the heuristic self-locus. The combined output of the retinoids in the scene synthesizer is projected to R3. Thus, whereas R2 in figure 4.4 represents the current veridical scene, R3, by combining and spatially rearranging veridical, remembered, and/or fabricated object representations, can create complex hypothetical frontal scenes. With neuronal mechanisms that will be described later, retinoid scenes can be tested according to individual needs. If the properties of a hypothetical scene are judged better than those of a veridical scene and if the person is able to change the environment so that the veridical scene conforms to the hypothetical, the environment may be changed or reorganized to this end.

Self-Location

In order for one to engage in effective behavior within the local environment and to assess the personal consequences of various external states, one must have some internal representation of the real and hypothetical spatial relationships between oneself and other significant objects in veridical space. The self-locus retinoid, together with the other retinoid registers shown in figure 4.4, serves this purpose.

The self-locus retinoid can be thought of as a standard translation retinoid that has no afferent input but does have the capability of point-to-point output projection to other retinoids. The self-locus register is distinctive in that it maintains a uniquely coded region of autaptic discharge with its central "point" located at the center of retinoid space. This fixed position of self-locus corresponds to the normal foveal axis. Autaptic neurons representing the central self-location (the "home" location of the self, so to speak) are constantly active, and autaptic discharge originating from this source of excitation can be spatially translated to any position on the surface of the retinoid by the usual shift control commands. Such heuristic locations and excursion paths of the self-locus can be projected to other retinoids and can be combined with real and/or hypothetical objects and scenes that may be represented on their surfaces to construct internal maps representing goal regions, obstacles, and direct and indirect paths to a goal.

Selective Attention

An important consequence arises from the ability to move replicated excitation from the source point of self-locus to selected regions of retinoid space: the autaptic cells in those regions of a retinoid that are stimulated by the added local excitation of a self-locus excursion (the heuristic self-locus) are preferentially primed and marked relative to other cells in the retinoid grid. These events provide a biological substrate for selective attention on the following bases: cells in a primed region respond more quickly and vigorously than those in other regions, and the heuristic self-locus marker can serve as a spatial reference so that the simple reverse of a self-locus excursion command can be used directly to back-translate any stimulus pattern found at the heuristic coordinate. The latter operation will shift a representation of the environmental pattern so that the region of interest will fall on the normal foveal axis, the source coordinate of the self-locus (figure 4.5). Once the stimulus is in this position, the pattern can be projected to the mosaic array, where it can be recognized with maximum acuity in the detection matrix. In a multilayered retinoid system, it is possible to have two (or more) simultaneous

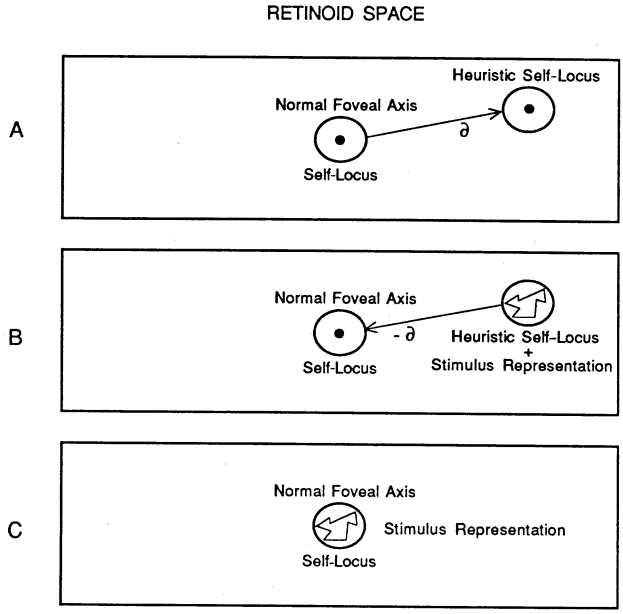


Figure 4.5
Selective attention by excursion of the self-locus. *A*: Priming a region of retinoid space by an excursion of the self-locus that is induced by shift control command ∂. *B*: Stimulus pattern appears in primed region of retinoid. Reversal of the sign of the shift control command ($-\partial$) translates both the heuristic self-locus and the stimulus back to the normal foveal axis. *C*: Stimulus pattern standardized at normal foveal axis where it can be projected to the synaptic matrix to be learned or recognized.

representations of a visual environment—for instance, one with the region of interest at its veridical coordinate with respect to the normal foveal axis and the other with the region of interest translated so that it is centered on the normal foveal axis or both of the layouts together with a representation of a hypothetical scene.

The neuronal process described makes it possible for a person to shift attention covertly over different regions of visual space without corresponding eye movements (Posner 1980; Posner, Snyder, and Davidson 1980; Shulman, Remington, and McClean 1979). Indeed, I suspect that when the tools of neuroscience are sufficiently advanced, it will be found that shifts in the heuristic self-locus normally precede saccades and play an integral role in the programming of saccadic eye movements. Recent findings about the relationship between saccades to visual and auditory targets and the discharge characteristics of single cells in the superior colliculus tend to support this view

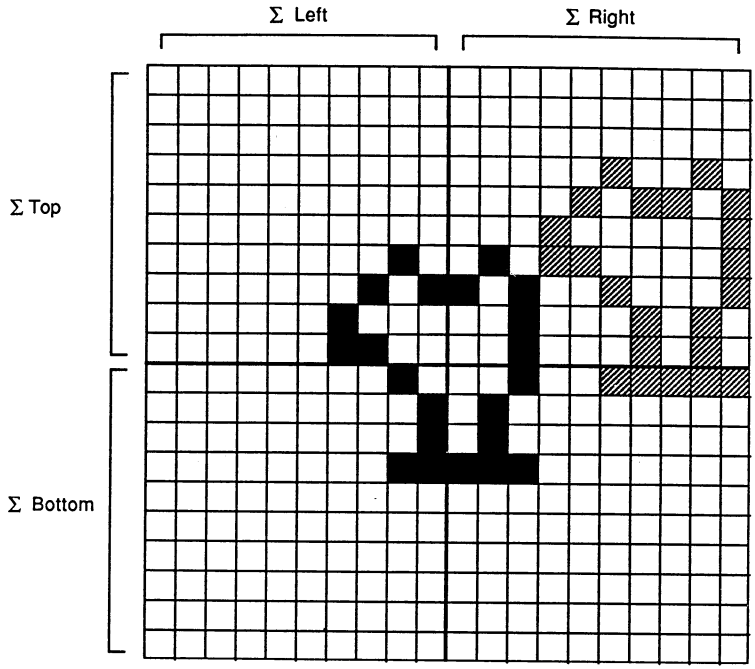
(Mays and Sparks 1980, Sparks and Mays 1983, Sparks and Jay 1987). In a summary of their investigations, Sparks and Jay (1987) conclude that it is the difference between a current eye position and a desired position that determines the direction and extent of the subsequent saccade. The retinoid model assumes that heightened autaptic cell activity, induced by the heuristic self-locus at its resting coordinate, always marks a "desired" eye position, at least in the sense of targeting a localized region of space to be explored. It follows from this formulation that unless eye movements are inhibited, the point of foveal fixation will normally correspond to the resting coordinate of the heuristic self-locus.

Finding and Positioning Pattern Centroids

In normal viewing, the projected image of an object of interest commonly is not centered on the fovea. Saccadic fixations on regions of the visual field do not ensure that the image of an object in a circumscribed region will be centered on the normal foveal axis even if the retinoid representation of that region is translated back to this reference coordinate of the self-locus. For the purpose of efficient learning and subsequent recognition of objects, it is desirable that patterns be positioned in a standard way before they are gated to the synaptic matrix, either to be learned or to be recognized by a best match against a particular profile of synaptic weights in the detection matrix. One way of accomplishing this is to provide a mechanism that ensures that the centroid of an excitation pattern on a retinoid will be shifted automatically so that it falls on the normal foveal axis.

Figure 4.6 illustrates how the standardization of image position is achieved in a retinoid system (Trehub 1986, 1990). Imagine a retinoid structure as if it were organized quadrantally, with each quadrant of its surface receiving retinotopic afferents from its respective retinal quadrant. On a given fixation, if the excitation of a captured pattern is summed independently over each quadrant of the retinoid and if the relative magnitudes of the summed discharges are used to drive its shift control cells, we then have a postretinal neuronal mechanism that can align the centroid of any parafoveal stimulus with the central axis of retinoid space (the normal foveal axis).

In the algorithm shown at the bottom of figure 4.6, each active autaptic cell represents one unit of excitation. The value of the difference in total excitation between the left and right hemifields is compared to a threshold that represents error tolerance (ET), where error is, in effect, determined by the degree of mismatch in excitation. As long as error tolerance is exceeded, a signal is sent to the appropriate



[(Σ Left - Σ Right) > ET] ====> Shift Right

[(Σ Right - Σ Left) > ET] ====> Shift Left

[(Σ Top - Σ Bottom) > ET] ====> Shift Down

[(Σ Bottom - Σ Top) > ET] ====> Shift Up

Figure 4.6
Quadrantally organized retinoid. Intersection of vertical and horizontal axes defines normal foveal axis. Each square represents an autaptic cell. Excitation evoked by a stimulus pattern is summed independently for each of the vertical and horizontal hemifields. The algorithm shown at the bottom of the figure shifts a pattern so that its centroid falls on the normal foveal axis. Diagonally hatched cells show initial retinoid position of a parafoveal stimulus. Solid cells show the stimulus with its centroid on the normal foveal axis after it has been shifted to balance mismatches in quadrantal excitation.

shift control cell to drive the retinoid pattern in a direction that reduces error (to the left if excitation in the right hemifield is greater than in the left and to the right if excitation in the left hemifield is greater). Exactly the same kind of mechanism adjusts the position of a pattern of excitation over the top and bottom fields. Thus, mismatches of excitation across the retinoid hemifields modulate the activity of shift control cells so that excitation is balanced over the retinoid quadrants (within error tolerance) and bring the centroid of an image to the normal foveal axis.

It is easy to see how an algorithm like the one presented can find a pattern centroid and position it on the desired axis. What minimal neuronal mechanism can perform the necessary computation? Figure 4.7 shows a schematic of a neuronal circuit that can do the job. Each active autaptic cell stimulates a paired interneuron (int), which provides excitatory input to a hemifield summation cell—L for summation of all left-field excitation and R for summation of all right-field excitation. Cells L and R are competitively cross-coupled by a pair of inhibitory interneurons so that the level of activity (output frequency) in the dominant summation cell increases as a function of the difference in activity between these two cells (reflecting the disparity in retinoid activity across the hemifields). Cell RC receives as its input and output of the left-field summation cell (L), and it commands the retinoid pattern to shift to the right. Cell LC receives as its input the output of the right-field summation cell (R), and it commands the retinoid pattern to shift to the left. If the activity of L is greater than R, neuron RC (having the steeper integration slope) will fire before LC, and the representation on the retinoid will be shifted a unit to the right. At the same time, the reset cell (marked $-$) will drive the membrane potential of both RC and LC to baseline, and the process will recycle as long as the hemifield mismatch is above error tolerance. The cell labeled "Tolerance" in figure 4.7 sets the level of inhibitory bias on the hemifield summation cells so that the higher the error tolerance is (reflecting the magnitude of inhibitory bias), the greater will be the disparity in retinoid field input to L and R that is required to initiate a pattern shift across the retinoid surface. It is assumed that error tolerance is normally proportional to the size of the retinoid pattern (small objects have low error tolerance; large objects have higher error tolerance). In this example, the mechanism for balancing left and right hemifields was described. The same kind of neuronal structure balances top and bottom hemifields in the process of finding and positioning the centroid of an image.

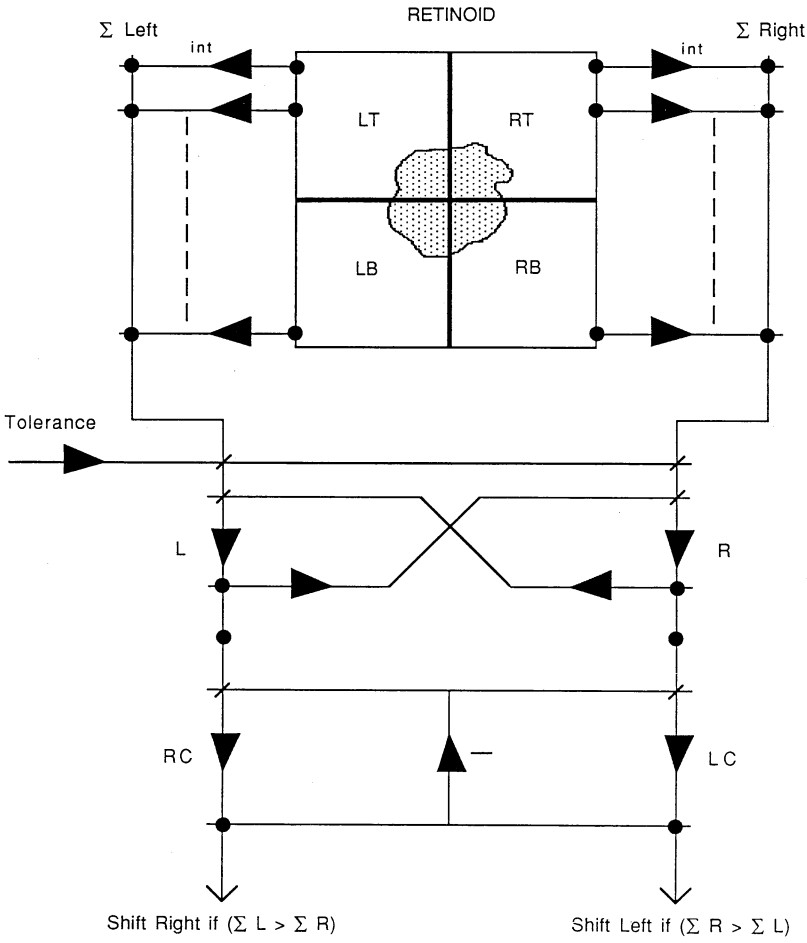


Figure 4.7
Schematic of neuronal mechanism for finding and positioning pattern centroids.

Depth Perception and Stereoscopic Vision

The mechanisms outlined can account for some essential visual-cognitive processes in a 2-D space. The network described here can integrate monocular retinoid outputs in a full binocular system to achieve depth perception and stereopsis in 3-D visual space.

A neuronal structure for stereoscopic vision is illustrated in schematic fashion in figure 4.8, which shows the arrangement of sensory and neuronal elements required for depth perception at one horizontal plane in the binocular visual field—a plane defined along the x-axis (L to R) for the horizontal dimension and along the z-axis in depth. Each small square in the strips designated "left-eye retina" and "right-eye retina" is assumed to contain a uniform 2-D array of retinal receptors and associated ganglion cells. The retinal image at each eye is transmitted retinotopically as a binary-coded edge transform to its associated retinoid, where it is mapped to the egocentric frame and shifted in accordance with the direction and degree of eye deviation from the normal foveal axis. Thus, the position of the visual pattern on each monocular retinoid corresponds to the relative angular position of objects and their features in the frontal field of each eye.

Correlation Clusters

Each diamond-shaped box in the 3-D retinoid shown at the bottom part of figure 4.8 contains a single cluster of retinotopically organized neurons corresponding to the ganglion cells that provide their sensory inputs through the two monocular retinoids. Thus, every cell in each 3-D cluster receives two afferent axons: one relayed from its corresponding ganglion cell in the left-eye retinoid and the other from its corresponding cell in the right-eye retinoid. Each diagonal string of cell clusters in the 3-D retinoid in figure 4.8 represents a line-of-sight array. The principal function of each cluster is to perform a cross-correlation between the micropatterns of inputs arriving from the left eye and the right eye. For this reason, these cell clusters are called correlation clusters (Trehub 1978).

The basic neuronal circuitry of a correlation cluster is shown in figure 4.9. Each of the nine principal cells in the cluster innervates a paired inhibitory neuron, which, in turn, synapses with all other principal cells in the cluster (note that it does not have an inhibitory synapse on its paired principal cell). It is assumed that all inhibitory neurons in this network are of the low-saturation type; they reach a constant peak output at minimal sensory input to their paired presynaptic cell. The magnitude of the correlation between the two micro-

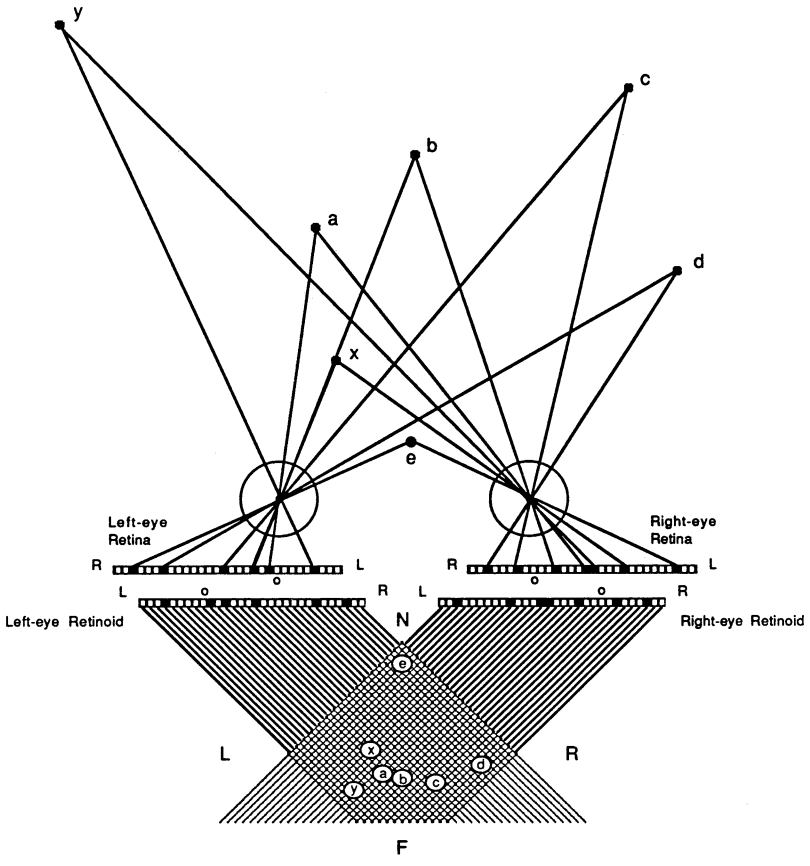


Figure 4.8
Schematic of binocular visual system. Imagine all retinoid elements stacked orthogonally to the plane of the page. Lowercase letters at top of figure represent objects in space with their projections to left-eye and right-eye retinas. Line-of-sight projections from each object terminate on filled squares representing stimulated retinal cells. Left-eye and right-eye retinoids are depicted just below their corresponding retinas. Filled retinoid squares represent stimulated autaptic cells. L and R indicate left and right visual fields. Notice that retinal representations are transformed at each retinoid to correspond with the relative egocentric positions of objects in terms of monocular visual angle. Diamond-shaped cells in the 3-D retinoid represent correlation clusters (A_j) that are innervated by intersecting retinotopic axon projections from the left-eye and right-eye retinoids. N (at top of 3-D retinoid) indicates the near visual field. F (at bottom of 3-D retinoid) indicates the far visual field. Lowercase letters within the 3-D retinoid indicate correlation clusters having maximum evoked activity in response to the corresponding objects in the visual field. The N-F dimension defines the Z-axis. Rows of correlation clusters orthogonal to the Z-axis define Z-planes (disparity planes; depth planes). Notice that relative depth discrimination decreases with object distance.

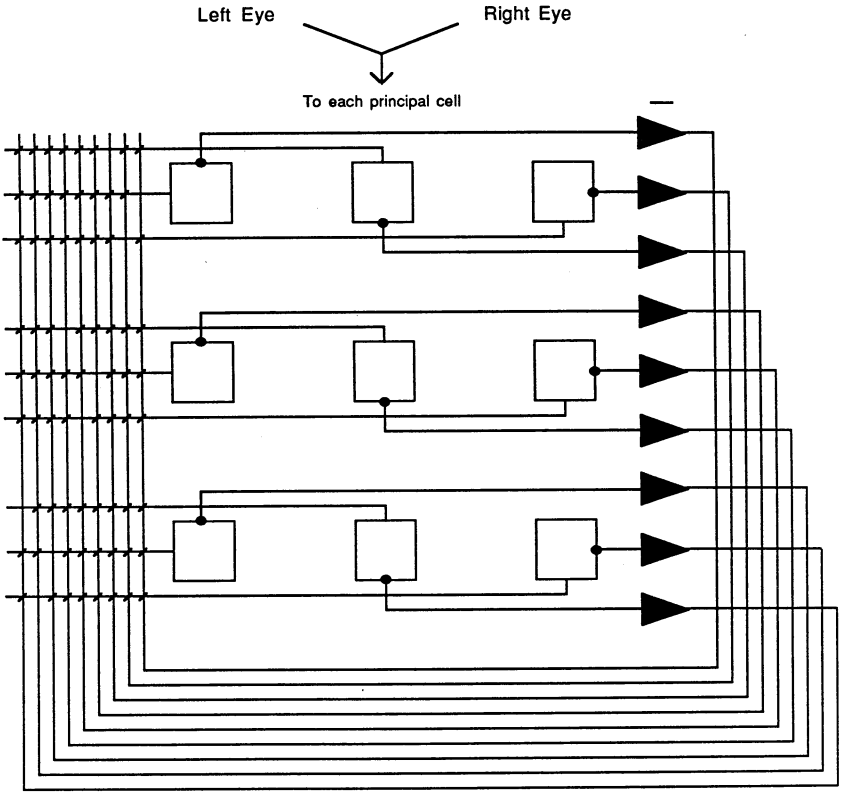


Figure 4.9
Correlation cluster (A_j). Squares represent principal cells of correlation cluster. Small, solid triangles represent inhibitory interneurons (−).

patterns impinging from the left and right retinas on each cluster is given by the total discharge density over all principal cells in the correlation cluster. Assuming that the output of each principal cell increases one unit for each active sensory input, we can write

$$E_j = k + \sum_{i=1}^{n} \dot{\mu}_{ij}\alpha_{ij} - \sum_{i=1}^{n} b(\dot{\mu}_{ij}\beta_{ij}) \tag{4.1}$$

where E_j is the total output of correlation cluster j; k is the endogenous cluster activity constant; $\dot{\mu}_{ij}$ is the ith stimulated principal cell of correlation cluster j; α_{ij} is the number of active sensory inputs on $\dot{\mu}_{ij}$ (1 or 2); β_{ij} is equal to $\sum_{i=1}^{n} \dot{\mu}_{ij} - 1$, the number of active inhibitory inputs on $\dot{\mu}_{ij}$; b is the coefficient of inhibition.

Brightening Neurons

Total output (E_j) from each correlation cluster (A_j) in each line-of-sight cluster string stimulated by the dominant eye converges on an associated neuron I call a brightening cell (λ) (figure 4.10), an excitatory neuron that sends an axon collateral to every principal cell in its paired correlation cluster, provides input for an inhibitory neuron that connects to all A_j in its line-of-sight string, and sends a horizontally and vertically oriented axon collateral to all principal cells along its disparity (Z-plane) row and column in the 3-D retinoid (see figure 4.8).

Given a pattern of binocular input, each brightening cell (λ) integrates E_j from its associated correlation cluster (A_j). Assuming uniform spike thresholds for all λ, the latency of discharge for λ will be a monotonic function of the magnitude of stereoscopic correlation at its associated A_j. Within any line-of-sight string of A_j (with respect to the dominant eye), the first brightening cell to fire inhibits all activity in all correlation clusters except its associated A_j (which receives an increment of feedback excitation from the active λ that is sufficient to null-out the inhibitory input). Thus, only those principal cell groups that correspond to the highest binocular cross-correlation of pattern fragments at each dominant line-of-sight cluster string in the 3-D retinoid will be active. Unless the images of an object seen with both eyes are significantly distorted by normal visual parallax, a single veridical micropattern or a pair of identical micropatterns projecting to the same horizontal retinal plane and within the range of stereoscopic fusion will yield the highest intercorrelation. Since the latter condition (two concurrent identical micropatterns in the range

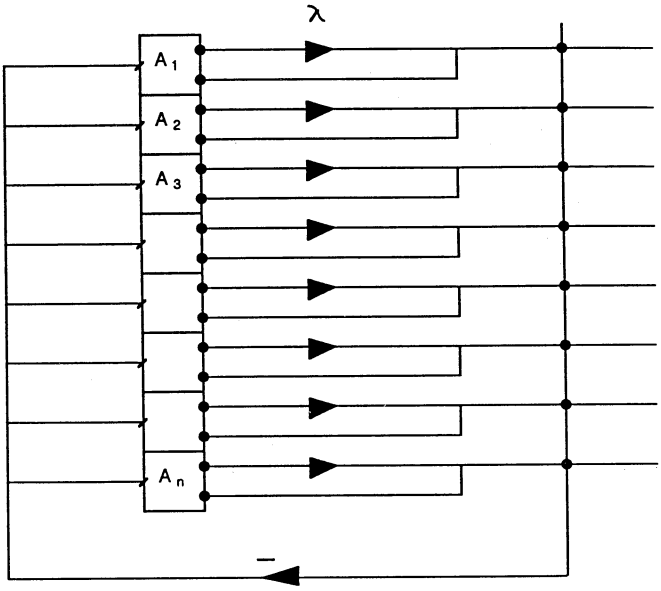


Figure 4.10
Line-of-sight string of correlation clusters (A_j; $j = 1, 2, 3, \ldots, n$) with associated brightening cells (λ) and squelching neuron (−). Each dot represents an excitatory contact with all cells in A_j, and each diagonal slash represents an inhibitory contact with all cells in A_j.

of fusion) is highly unlikely to occur in nature without deliberate contrivance, this putative neuronal mechanism should be effective in squelching false binocular targets.

Julesz (1971) has provided an interesting demonstration of how 3-D shapes can be perceived binocularly in the absence of any monocular shape information. If a 2-D random pattern of dots is presented as a stimulus to one eye and the other eye is simultaneously stimulated with the same pattern having some of its surface slightly offset horizontally, the offset regions appear as discrete shapes at different depths depending on the direction and degree of offset. Monocular viewing of either display reveals nothing but a 2-D dot pattern. The binocular disparity information in the random dot stereogram is sufficient to induce a vivid perception of bounded dot patterns at various depths (Z-planes), and the patterns are seen as belonging to coherent shaped surfaces. The neuronal mechanism for stereopsis described responds similarly. At each Z-plane, the veridical input is conserved in the pattern of stimulus-evoked excitation, which, in turn, is embedded in a flux of low-level excitation projected to all principal cells

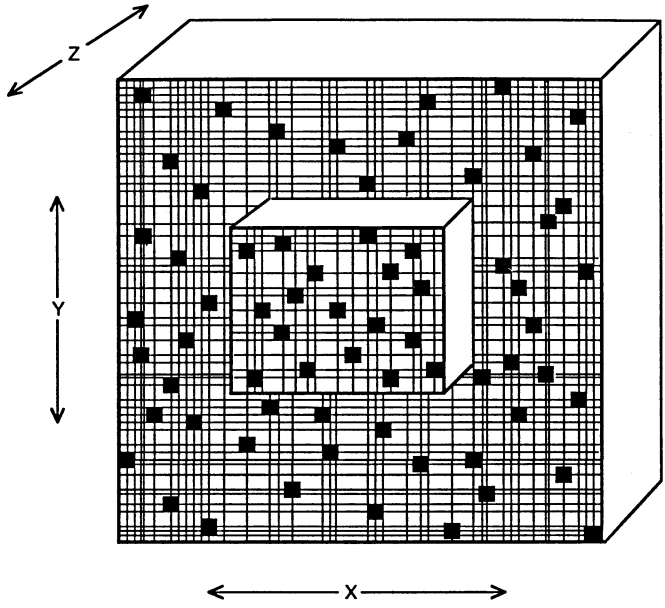


Figure 4.11
Enhancement of planar surfaces (Z-plane brightening in 3-D retinoid). Surface excitation is distributed en passant (Shepherd 1979) to principal cells in each correlation cluster (A_j) via brightening cell (λ) collaterals oriented along the x- and y-axis on Z-planes defined by the stimulus-evoked discharge of the clusters together with complementary inhibition along line-of-sight cluster strings (see figure 4.8). Nodes of excitation are densely distributed over the planar surfaces at the regions of λ-collateral intersection. Brightening cells are represented by solid squares. Lines projecting from each cell represent axon collaterals.

in the plane by λ axon collaterals (Z-plane brightening; see figure 4.11). Sharp planar boundaries are established by the abrupt inhibition of brightening in the appropriate region of one plane and the evocation of adjacent brightening on a different plane in accordance with the activity of correlation clusters and their coupled λ-cells on the Z-axis.

An Example of Performance

This stereopsis model was simulated, using random dot stereograms, to assess performance (Trehub 1978). Each correlation cluster (A_j) in the 3-D retinoid consisted of nine principal cells arranged in a square array as shown in figure 4.9. Random dot stereograms were generated by computer, with each black dot representing a point of stimulation to be registered on the retina. Dot density was arbitrarily set

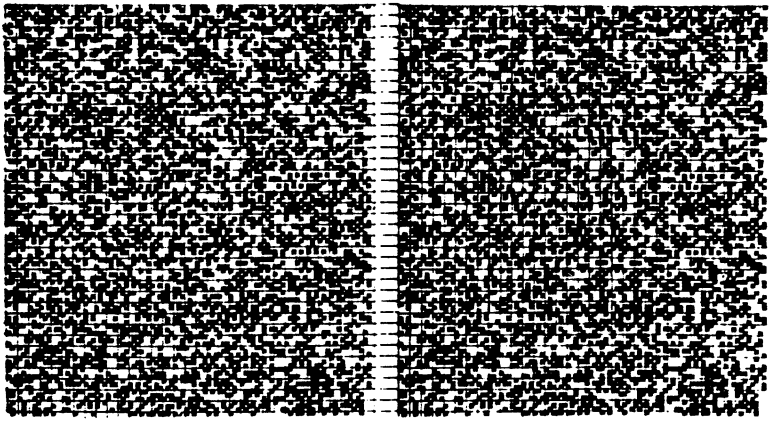


Figure 4.12
Random dot stereograms used as binocular stimuli for simulation test of 3-D retinoid model. Random dot density, 40 percent. Left, pattern A; right, pattern B, identical except for dot displacement. Each A_j cluster is indicated by grid overlay. In pattern B, a large vertically oriented rectangular area of dots has been displaced three A_j units to the left, and within this rectangular area a small square region has been displaced five A_j units to the left with respect to A. Pattern A stimulates left regina. Pattern B stimulates right retina. Source: Trehub 1978. Copyright Academic Press (London) Ltd. Reproduced by permission.

at 40 percent (figure 4.12). Retinal activity was assumed to be appropriately shifted in the left-eye and right-eye retinoids with reference to egocentric spatial coordinates and projected from each monocular retinoid as line-of-sight units of stimulation to the corresponding principal cells in the 3-D retinoid.

The initial output of each correlation cluster (A_j) was determined by equation 4.1, where the endogenous activity constant (k) was arbitrarily set at $k = 10$ and the coefficient of inhibition (b) was set at $b = 0.4$. Excitation on each principal cell from plane-brightening neurons (λ) on each plane over the Z-axis was arbitrarily set at $\lambda = 0.05$ and limited to $\lambda = 0.10$ at cross-over nodes (see figure 4.11). On the basis of the squelching principle and illustrated in figure 4.10, the correlation cluster (A_j) yielding the highest output in each line-of-sight string from the dominant eye inhibited all other cell clusters in its string.

The random dot stimuli presented to each eye are shown in figure 4.12; the result of the simulation test is shown in figure 4.13. The model discriminated the disparity planes (Z-planes) correctly, assigned the proper boundary contours to each of the patterns in depth, and conserved the original surface information.

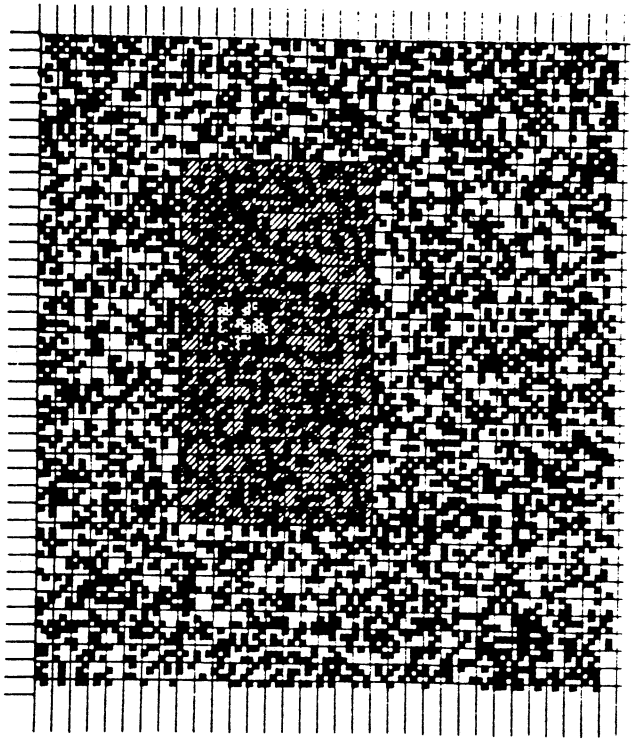


Figure 4.13
Result of simulation test. Large, diagonally hatched rectangular area is "seen" (represented) in midground of 3-D retinoid. Small, stippled square area within hatched rectangle is seen as nearest. Unmarked area is seen as background (behind the other surfaces on Z-axis). Planar coherence is established by contour-demarcated brightening on each Z-plane. Internal planar details are conserved. Source: Ibid. Copyright Academic Press (London) Ltd. Reproduced by permission.

Other Properties of the Retinoid System

The retinoid system and similar neuronal mechanisms based on the dynamic characteristics of autaptic cells can provide a variety of visual-cognitive functions in addition to those considered so far. Two of these properties are described below; other functions, some of which depend on system integration with other circuits, will be introduced in later chapters.

3-D Representation in Monocular Viewing

Beyond the capability of the 3-D retinoid to represent the depth of objects and surfaces on the basis of binocular disparity, there is the fact that depth perception is normally adequate even in the case of monocular viewing. Furthermore, a strong sense of depth can be experienced when we look at 2-D perspective drawings of 3-D objects. From a psychological standpoint, there are a number of cues that are known to contribute to depth perception (relative size, interposition, aerial perspective), but we will consider two biological processes that regulate the 3-D retinoid so that objects in depth are represented on appropriate Z-planes in monocular perception.

The sensory-motor structures that govern ocular convergence and accommodation are assumed to generate corollary signals that provide excitatory bias to the appropriate Z-planes in the 3-D retinoid. For example, in the case of convergence, an inward rotation of the eyes (even with one eye occluded) to achieve fixation of a near object in the central frontal field would selectively bias Z-planes representing near space. The greater the degree of inward rotation is, the closer is the retinoid representation of the object; the less the degree of inward rotation is, the more distant is the representation on the Z-axis. Accommodation influences the 3-D retinoid in a similar fashion. The more convex the lens of the eye is, the nearer would be the Z-planes receiving excitatory bias; the less convex, the more distant would be the representational planes that are biased. Notice that the preferential priming of a particular retinoid plane has an effect comparable to that given by a peak binocular correlation on that plane; in each case, correlation clusters that fall on other planes along a common line-of-sight string are inhibited.

3-D Imagery

Individuals are able to imagine a previously viewed object at different distances in response to an instruction to do so (Kosslyn 1978, 1980). The 3-D retinoid provides a physical explanation for this capability.

Figure 4.14 illustrates how this can be done. An image evoked on

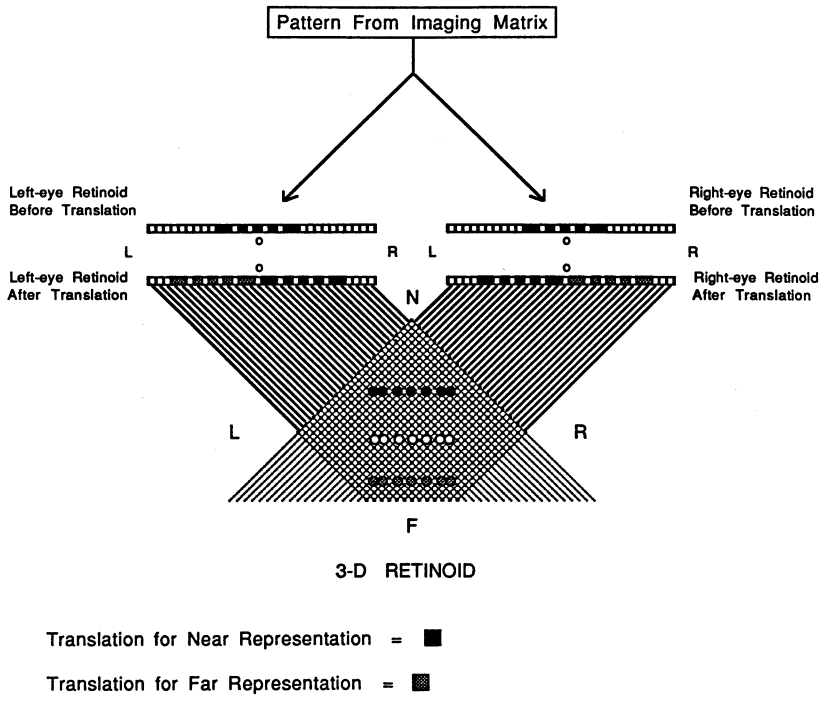


Figure 4.14
Moving imaginary patterns in depth. Input from imaging matrix is captured in the left-eye and right-eye retinoids. Initial position of the pattern within the 3-D retinoid is indicated by unfilled circles. Binocular translation in the medial direction (black squares) moves the image nearer in the 3-D retinoid (black circles). Translation in the temporal direction (stippled squares) moves the image farther away in the 3-D retinoid (stippled circles).

the mosaic cell array by a class cell collateral to the imaging matrix is sent as a retinotopic pattern to both the left-eye and right-eye retinoids. These captured patterns can then be translated in opposite lateral directions by appropriate pulses to the shift control cells governing such retinoid. The result of this operation will be to move the representation of the imagined object along the Z-axis in the 3-D retinoid. For example, if a common pattern is translated to the left on the left-eye retinoid and to the right on the right-eye retinoid, its representation on the 3-D retinoid will shift to a more distant Z-plane. Reversal of the direction of translation in the left-eye and right-eye retinoids will bring the imagined object closer in 3-D representational space (figure 4.14).