Methods in Neuronal Modeling

From Synapses to Networks

edited by Christof Koch and Idan Segev

A Bradford Book

The MIT Press

Cambridge, Massachusetts, and London, England

CHAPTER 11

Simplifying Network Models of Binocular Rivalry and Shape-from-Shading

SIDNEY R. LEHKY and TERRENCE J. SEJNOWSKI

11.1 Introduction

In this chapter we will review two neural network models of visual processing. These simplifying brain models are at a level of organization in the nervous system between the single cell and the cortical column, and at a level of analysis between the computational and the implementational (Marr 1980; Churchland, Koch, and Sejnowski 1989). The goal of these models is to provide explanations for perceptual phenomena based on the representation and processing of visual information in neural populations.

As experimental data accumulate at the level of single neurons, more and more detailed models become possible that mimic more closely the processing of particular circuits. This approach, which might be called realistic brain models, is most useful when the function of the circuit is already known and the knowledge about the circuit is almost complete down to the biophysical level. The model of central pattern generation in chapters 6 and 7 is a good example of this approach. One of the lessons learned from modeling invertebrate circuits is that a wiring diagram is not nearly enough to specify a circuit—specific biophysical membrane properties are also crucial (Selverston 1985). Unfortunately, in most parts of the vertebrate central nervous system we have only a vague notion about the function of circuits, and information about the biophysical level is at best incomplete. Even the patterns of connectivity are uncertain in cerebral cortex.

Another approach to the network level is to start with a function such as a perceptual ability and design simplified neural circuits that can perform the function within the constraints of the state of knowledge. One example of this "simplifying" approach was the Marr and Poggio (1976) cooperative model of binocular depth perception, which demonstrated

(Julesz 1971). These models can be simulated on a digital computer and their properties compared with human performance. However, these models are often too general to compare directly with physiological experiments, and at best they serve as a demonstration of one way that a problem can be solved.

Neither the top-down nor the bottom-up approach is ideal—what is needed is some approach that combines the strengths of both strategies. The value of a model is that it can incorporate both the existing knowledge at the biological level and the performance of the system from psychophysical measurements. The two examples given here illustrate the usefulness of such a combined style of network modeling, taking advantage of both the realistic and simplifying approaches. They are at a level that is beyond that of modeling the details of identified neurons, but not so general that essential features such as response properties of single neurons can no longer be identified.

11.2 Binocular Rivalry

The first model presented here is concerned with interactions between visual processing in the two eyes. Binocular rivalry occurs when incompatible images are presented to the two eyes, such as vertical stripes to one eye and horizontal stripes to the other. In such a situation the visual system is thrown into oscillation, so that first the image from one eye is visible, and then the other, typically for a period of about one second. In general, the entire visual field does not oscillate in unison unless the rival stimuli are sufficiently small (less than 1° in diameter), but rather it breaks up into a constantly changing mosaic of the two images.

In choosing this topic, the desire was to study an aspect of binocular vision that has received less theoretical attention than stereopsis, not only because of its intrinsic interest, but also because its study may provide clues and constraints in constructing a biological model of stereopsis. Rivalry appears to be a problem of intermediate complexity in the sense that the issues can be formulated in terms relevant to the biological concerns of the experimental neurophysiologist as well as the global concerns of the psychophysicist.

11.2.1 Experimental Data

A large body of psychophysical data related to rivalry has accumulated over the past century (O'Shea 1983). The durations of alternating left and right dominance show statistical variation, and the mean durations

depend on the stimuli. The nature of this dependence is unusual. When the stimulus strength (contrast, for example) is increased to one eye, the duration of time for which the opposite eye is dominant decreases. It is therefore possible to independently vary the duration for which each eye is dominant. Also, the oscillations follow a rectangular waveform. This is indicated by the observation that a spot of light flashed to the suppressed eye has its detectability reduced by a constant amount over the entire duration of the suppressed phase, interpreted as showing that the strength of suppression remains constant until being abruptly cut off.

Neurophysiological data on the matter is more sparse. John Allman (unpublished data) has observed in the superficial layers of primary visual cortex in an alert owl monkey some neurons whose neural activity switched on and off in synchrony with behavioral indications that an eye had undergone a transition from suppressed to dominant states. A significant proportion of the cells in the middle temporal area (MT) show similar behavior (Nicos Logothetis, personal communication). This is the only neurophysiological report of oscillations associated with binocular rivalry. Varela and Singer (1987) have recorded from relay cells in the lateral geniculate nucleus (LGN) of anesthetized cats exposed to rival stimuli. They found that strong inhibition occurred when stimuli to the two eyes were unmatched. This inhibition had a latency of hundreds of milliseconds, and was abolished by disruption of corticofugal inputs through ablation of the cortex. Although no oscillations were observed (possibly because of the anesthesia), these data suggest a binocular inhibitory process at an early stage of the visual system whose activity is related to the degree of correlation between images to the two eyes. Together with the results of J. Allman and N. Logothetis, these studies indicate that although rivalry has been chiefly a concern of psychophysics, it is feasible to approach the phenomenon with neurophysiological techniques. This may be one of the simplest experimental paradigms that could link conscious awareness of sensory stimuli with neural activity.

11.2.2 Neural Network Model

By its very nature, binocular rivalry suggests some sort of reciprocal inhibitory linkage between signals from the left and right eyes prior to the site of binocular convergence. Reciprocal inhibition is common at all levels of the nervous system, from peripheral processing in the retina to visual cortex. The simplest neural network implementation of such a system is illustrated in fig. 11.1, which shows the responses of the network under different stimulus conditions, as discussed below (see also

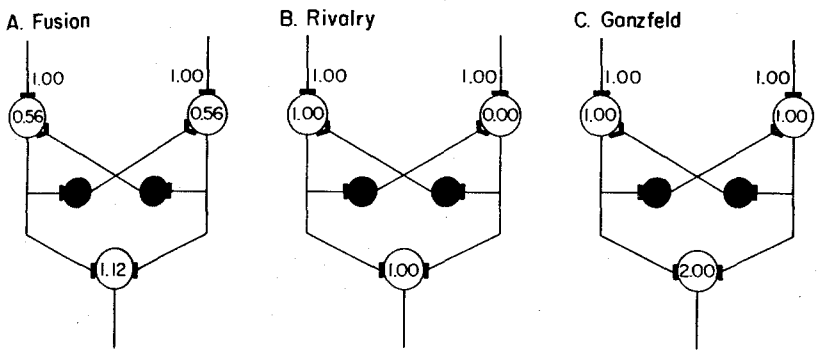

Figure 11.1
Activities in the neural network for three patterns of binocular input. (A) correlated contours (fusion), (B) uncorrelated contours (rivalry), and (C) no contours (*Ganzfeld*). Numbers indicate relative levels of activity when inputs are arbitrarily set to 1.0. Differences in activities are explained by postulating that binocular spatial correlation affects the strength of inhibitory coupling between the left and right sides.

Lehky 1988). The open circles represent excitatory neurons and the filled circles inhibitory neurons. There are two excitatory neurons, one from the left side and one from the right, which converge upon a binocular neuron. These two neurons receive sensory information along inputs from the periphery, indicated by the lines originating from the top of the diagram. Finally, there are two inhibitory neurons, each inhibiting the other side in feedback fashion. This circuit is intended to model binocular processing for only a small, isolated patch of visual field, and is incomplete in the sense that it does not include lateral interactions.

The point of the network is to model in as simple a way as possible the essential neural interactions that may underlie the oscillation of binocular rivalry, without considering many of the biophysical details of actual neurons. Oscillations of course imply a system whose state is changing as a function of time, or a dynamical system. The qualitative behavior of a dynamical system can be found using the mathematical methods of stability theory, as developed in chapter 5. Stability analysis gives the general conditions under which the system will fall into different behavioral states (oscillating or nonoscillating) without providing quantitative information about those states. When a stability analysis is performed on a reciprocal inhibition network, it can be shown that two conditions

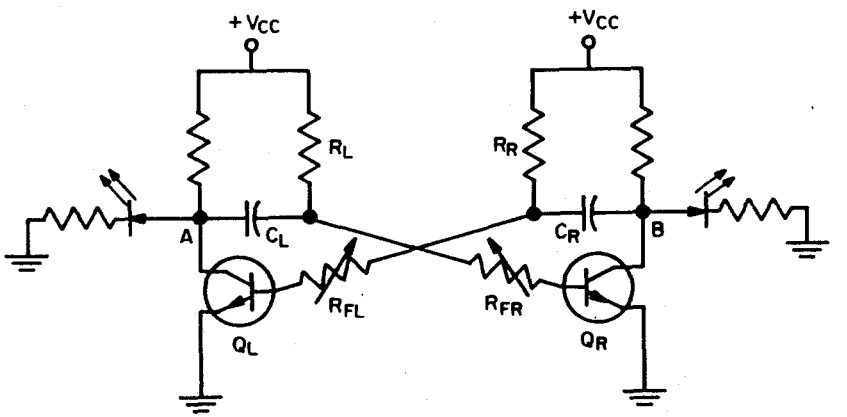

Figure 11.2
Circuit diagram of an astable multivibrator. This circuit is meant to be a physical analog of the neural networks shown in fig. 11.1, in which the two transistors Q_L and Q_R represent the two monocular excitatory units in that figure (open circles), resistors R_{FL} and R_{FR} control strength of inhibitory coupling in the reciprocal feedback inhibition pathway, and capacitors C_L and C_R control the adaptation time constant of the inhibition. The subscripts L and R refer to the left and right sides respectively.

are required for it to go into oscillations (Matsuoka 1984, 1985): (1) the inhibitory coupling must be sufficiently strong, and (2) adaptation of inhibition must also be sufficiently strong.

11.2.3 Analog Electrical Circuit

In order to go beyond a simple qualitative analysis and look at some of the actual properties of the oscillations produced by the network, it is necessary to simulate the behavior of the system. The dynamics of the system are governed by coupled nonlinear differential equations that are not analytically soluble. Therefore one must either solve them numerically on a computer or find an equivalent physical system that is subject to the same equations but can be more easily studied.

Following the second approach, the system that was chosen as an analog to the neural network for rivalry was the electronic circuit shown in fig. 11.2 (analog in the sense that its behavior shared the dynamical features of interest with the neural network). The circuit is called an astable multivibrator and is essentially an oscillating flip-flop, whose

behavior is described in many electronic textbooks and which is easily built.

The analogy between this circuit and the neural network can be seen as we step through the operation of the circuit. The two transistors Q_L and Q_R represent the left and right excitatory neurons. The points labeled A and B in the circuit are equivalent to the outputs of those two neurons, and LEDs were placed at those points to allow visual monitoring of the output voltages at those points. Nothing in the electronic circuit corresponds to the binocular neuron. It would have been straightforward to model a binocular neuron by including circuitry that linearly summed the voltages from points A and B, but that would not have added to our conceptual understanding of what was going on, and would have cluttered and complicated the circuit. The left and right transistors are connected in a manner that may be described as reciprocal inhibition. The "inhibitory" pathway from Q_L to Q_R runs through C_L and R_{FR} to the base of Q_R, and analogously from point B to the base Q_L on the other side. If the voltage at point A is high, that forces the voltage at point B to be close to zero, and vice versa. As the circuit oscillates, the voltages at points A and B alternately go high and low.

The slow charging of the capacitors C_L and C_R along the circuit path connecting the two transistors can be thought of as "adaptation" of inhibition, and this adaptation is necessary for the system to go into oscillation. When the voltage of capacitor C_L on one side gradually charges up to the threshold of transistor Q_r on the other side, the system flips state. At this point the whole process starts over with the other capacitor and transistor, and so on back and forth indefinitely as the system oscillates (the transistors here effectively act as switches; that is to say, their transfer function between "input" base voltage and "output" collector voltage can be approximated by a very steep sigmoid).

Finally, the values of the variable resistors R_{FL} and R_{FR} in the feedback pathway can be thought of as determining the strength of "inhibitory coupling" between the right and left sides. Small resistances correspond to strong inhibitory coupling and large resistances to weak coupling. As we said before, the strength of inhibitory coupling is one of the factors that determine whether a system with reciprocal feedback inhibition will oscillate or not. This can easily be demonstrated in this circuit, for as one increases the values of the feedback resistance (decreases inhibitory coupling) by turning the knob of a potentiometer, a particular point is reached at which the oscillations suddenly stop. Instead of each LED being alternately fully lit or completely dark, both are now on simultaneously, but at some intermediate level of brightness.

In the language of dynamical systems theory this is called a bifurcation point, a point at which a discontinuous change in the behavior of the system (from oscillating to nonoscillating) occurs as one of the system variables (feedback resistance) is continuously changed.

Now we can compare the behavior of this astable multivibrator circuit with that of binocular rivalry. One similarity is that the astable multivibrator produces rectangular oscillations. In addition, the duration of time that the left or right side is "dominant" can be varied independently by changing the value of a parameter (time constant for charging the capacitor) on the opposite side. This is analogous to the behavior of rivalry, if one substitutes "change contrast" for "change capacitor (adaptation) time constant." Although it will not be discussed in detail here, the statistical distribution of the durations of dominance and suppression can also be replicated by the system under consideration here by adding random noise to the model (Lehky 1988). Finally, the electronic circuit passes from an oscillating to nonoscillating state as the value of the feedback resistances (strength of inhibitory coupling) is varied. In the binocular visual system, we know that the behavior goes from oscillating (rivalry) to nonoscillating (fusion) as the correlation of images presented to the two eyes is varied.

This last point brings up an important physiological prediction of the astable multivibrator model, which is that not only is there reciprocal feedback inhibition prior to the site of binocular convergence, but this inhibition involves synapses whose strength is affected by the degree of correlation between the left and right images. High correlation would lead to weak inhibitory coupling across the synapses, and low correlation would lead to strong inhibitory coupling. Furthermore, it has already been mentioned that changing stimulus strength (contrast) in rivalry and changing the adaptation time constant in the astable multivibrator circuit have analogous effects. Therefore, a second prediction would be that the binocular reciprocal inhibition postulated here shows adaptation whose time constant depends on contrast. In other words, adaptation occurs at a faster rate when contrast is increased.

Going back to fig. 11.1, the activities within the neural network as postulated by the model are shown for binocular fusion and rivalry (the values for fusion are in fact taken from an earlier model (Lehky 1983) that considered psychophysical results about how various luminances presented to the two eyes combine to form the perception of binocular brightness). The last diagram in the figure shows some psychophysical data from Bolanowski (1987) that show that when no contours are present (uniform field or *Ganzfeld* conditions) the binocular reciprocal

inhibition disappears, so that luminances presented to the two eyes add up linearly to form the percept of binocular brightness. This is included here in support of a central point of the modeling, that the strength of inhibitory coupling in binocular vision is a function of the spatial patterns presented to the two eyes. The strength of inhibitory coupling for the three conditions can be ordered as follows: uncorrelated contour (rivalry) > correlated contour (fusion) > no contour (*Ganzfeld*).

It is important to point out that the model as presented here is incomplete, since it just says that the degree of correlation between the left and right spatial patterns affects the strength of inhibitory coupling, but does not give any mechanism by which this may occur. This postulated effect of spatial correlation on binocular inhibition is presumably mediated by various lateral interactions not considered in the model.

A functional model such as the one presented here cannot define anatomical location. However, the physiological data of J. Allman and N. Logothetis (personal communication) and of Varela and Singer (1987) suggest that the processes under consideration here are already occurring at an early stage (primary visual cortex and area MT). Varela and Singer further propose that the correlation between the left and right images computed at the cortical level controls, through known feedback connections, inhibitory circuits in the LGN or perigeniculate nucleus that gate the signals passing through the LGN to the cortex. Certainly the precise binocular alignment of layers in the lateral geniculate nucleus points to an important role in binocular vision beyond the inadequate notion of the LGN as the recipient of a rather amorphous set of modulatory influences from the brainstem, or the simple notion of the LGN as a "relay center" (Sherman and Koch 1986; Koch 1987).

In conclusion, although the model presented here was at a much simplified level, it both fits the available data well and suggests new lines of experimental investigation that might not have been otherwise considered.

11.3 Computing Surface Curvature from Shaded Images

In the previous example, intuition was coupled with knowledge of anatomy and physiology of the visual system to arrive at a plausible network model of binocular rivalry. The network is a small but important part of a larger, more complex system. As processing is traced into the visual cortex, it becomes more difficult to find plausible network models based on intuition.

Recently, new "learning" algorithms have been devised as constructive techniques for designing networks that can perform specified transformations between input and output. In this section we describe the application of one of these algorithms, called "error back-propagation," to a problem in visual processing (see Appendix 11.A). Specifically, the algorithm is used to construct a network that can extract certain shape parameters from shading information contained in continuous images of simple surfaces (Lehky and Sejnowski 1988).

Our interest here is in the properties of the receptive fields in such a network, and how they compare to receptive fields actually found in visual cortex. The general finding is that the receptive fields developed by the network are surprisingly similar those found in primary visual cortex. We conclude that neurons that can extract surface curvature can have receptive fields similar to those previously interpreted as bar or edge detectors (Hubel and Wiesel 1962). The receptive field of a sensory neuron is necessary but not sufficient to determine its function within a network. We emphasize the importance as well of the "projective field" of a neuron in determining its function, where the projective field is the output pattern of connections that the neuron makes with other cells.

The approach taken here to the "shape from shading" problem differs fundamentally from that which has been used by researchers in machine vision (Ikeuchi and Horn 1981; Pentland 1984). In that work, explicit rules for extracting the parameter of interest from the image are given in the form of mathematical equations. In this network model, the rules are implicitly contained within the thousands of "synaptic weights" between the units of the network, and do not lend themselves to a more compact description.

It may be useful to compare the machine vision and network models in terms of the three levels of analysis set forth by Marr (1980): the computational level, the algorithmic level, and the hardware or implementation level. They differ most obviously, and perhaps most importantly, in the last level, that of implementation. The reason this may be the most important difference is that the properties and limitations of the hardware influence the choice of algorithm. When constrained to deal with the problem using a network architecture similar to that occurring in the brain, the resulting solution is quite different from what was conceived when the computational problem had been considered in isolation. The machine vision algorithms are local, dealing with the relation between adjacent pixels. Whatever globality they achieve is done

serially, by pasting together a sequence of local analyses. In contrast, the algorithm that the network implements appears to be intrinsically global and parallel, reflecting its architecture. The network seems to handle shape-from-shading as a problem in pattern recognition, looking for particular configuration of light and dark over a fairly large patch of surface (large relative to a pixel). For more difficult problems it may be necessary to have feedback connections, such as those used in the binocular rivalry network, which would give the network a more complex temporal dynamics.

11.3.1 The Task of the Network

The local curvature is an important descriptor of the shape of a surface. The network model described here is intended to extract surface curvature from shaded images. Curvature is defined as the rate of change in the orientation of the surface normal vector as a function of arc length as one moves along a curve lying on the surface. It was selected as the parameter of interest for the network because it is a relatively robust indicator of shape. The magnitude of surface curvature is independent of rotations or translations of the surface (which is not true for shape described in terms of surface normals).

In general, the curvature at a point on a surface depends upon the direction one travels along the surface. The directions for which curvature assumes maximum and minimum values are called the principal directions, and the maximum and minimum curvatures themselves are called the principal curvatures. There is a theorem in differential geometry that states that the two principal curvatures always lie along orthogonal directions on a surface.

Knowing the principal curvatures at a point provides useful information about the local properties of the surface. If both principal curvatures are positive, the surface is convex; if both are negative, it is concave, and if they have opposite signs, the surface is saddle-shaped. If one principal curvature is zero, the surface is cylindrical, and if both are zero the surface is planar. Aside from curvature sign, the magnitudes of the principal curvatures themselves are an indication of the spatial scale of features on the surface. Hence it would be helpful to have a way of estimating principal curvatures and their orientations directly from the shading information in an image. Of course, the shading of an object depends on the direction of illumination. To be interesting and useful, the network should be capable of determining surface curvature independently of illumination direction (i.e., the network should find some aspect of the image related to curvature that remains invariant as illumination direction is changed). Furthermore, the parameters extracted by the network should not depend on the image being precisely aligned at any particular position in the overall receptive field of the network.

To summarize, the task of the network is to find the magnitudes and orientations of the principal curvatures at the center of a simple geometrical surface patch, independently of illumination direction and the position of the surface.

11.3.2 Designing the Network Model

As in the binocular rivalry model, the network to be presented here is meant to correspond to a small circuit receiving input from only a very limited region of the visual field, perhaps a region subserved by a single cortical column. The network therefore determines curvatures for a small patch of a large, complex image. To generate descriptions of images, we believe that the network here would have to be replicated at different spatial locations to cover the entire visual field, and also replicated at different spatial scales, which would all feed into higher-level networks that integrated local curvatures into more general shape descriptions (see Pentland 1986 for discussions of synthesizing descriptions of complex shapes from simple shape primitives).

The network is constructed to extract principal curvatures from a set of simple geometrical surfaces (elliptic paraboloids). The eventual relevance of this to more complex images is contained in the assumption that any surface patch can be locally approximated by such a simple geometrical surface. Since curvature is proportional to the second derivative of surface position, an arbitrary surface having the same curvature as the actual surface forms a second-order approximation to that surface. This is probably sufficiently accurate, given that the psychophysical evidence indicates that the human visual system does not generate very good estimates of surface shape from shading information alone (Mingolla and Todd 1986).

As stated before, the input images to the network were of elliptic paraboloids (parabolic cross section in depth and elliptical cross section in the fronto-parallel plane). This surface was selected because, unlike ellipsoids for example, it has no edges (occluding contours) other than edges caused by the limits of the network receptive field itself. We were interested in studying network responses to shading, without the confounding presence of edges in the images.

Reflection from the surface was assumed to be matte, so that scattering of light from the surface is independent of the location of the viewer. Illumination was also assumed to have a diffuse or Lambertian

component, so that although light came predominantly from a particular direction, there were components from all directions arising from light reflected and scattered about by the general environment of the surface (see Appendix 11.B for details). This diffuse illumination was used in order to eliminate hard shadow edges from the image, again in keeping with our desire to study network responses purely to shading. In any case, illumination in reality almost always has a diffuse component. Specular reflections, which are also common in nature, were not included.

Another concern in constructing the network was ambiguity in the sign of the curvature. Without knowing the direction of illumination, it is impossible to distinguish a positively curved surface from a negatively curved one. The appearance of a convex surface with light coming with a tilt of 30°, for example, is physically indistinguishable from a concave surface illuminated at a tilt of −30°, and a saddle-shaped surface can be found that is also indistinguishable. The inherent ambiguity in curvature sign is well illustrated by the picture of the inside-out mask face in Gregory (1966, p. 127), in which surface curvatures are perceived as the opposite of what they actually are. Because of this ambiguity, a particular image presented to the neural network can correspond to several different surfaces (convex, concave, or saddle-shaped).

To resolve ambiguity in the signs of the two principal curvatures, two assumptions were built into the network. The first assumption placed restrictions on possible directions of illumination. It was assumed that illumination always came from above (light tilt between 0° and 180°). This was sufficient to fix the sign of one of the two principal curvatures. There is in fact evidence that biological systems do make this assumption, coming from observations of the well-known "crater illusion," in which craters come to be seen as mounds when the photograph is turned upside down. The interpretation of the image, whether as a crater or as a mound, is always consistent within the implicit assumption that illumination is coming from above (Ramachandran 1988).

The sign of the second principal curvature in principle could have been disambiguated by making further restrictions on illumination directions (for example, not only does light always come from above, it always comes from the left). However, there did not appear to be any plausible justification for making this further restriction on illumination. Rather, the assumption was built into the network that both principal curvatures have the same sign. That is to say, the assumption was made that the surfaces presented to the network were always convex or concave, but never saddle-shaped. Possibly, this assumption may be relaxed in more complex images where additional information is available from neighboring regions.

11.3.3 Constructing the Network Model

We used the "error back-propagation" learning algorithm as a design technique for constructing a network with the desired characteristics (Rumelhart, Hinton, and Williams 1986; Parker 1986; LeCun 1985; Werbos 1987). The details of the algorithm are given in Appendix 11.A. Essentially, the following was done. The network was presented with many sample images. For each presentation, responses were propagated up through several layers of neural-like units to a layer of output units. The actual responses of the output units were then compared with what the output for that image should have been. Based on this difference, synaptic weights throughout the network were slightly modified to reduce error, starting with synapses at the output layer themselves and then moving back down through the network (hence the name back-propagation). After thousands of image presentations, the initially random synaptic weights organized themselves into a set of receptive fields that provided the correct input/output transfer function. (In all cases, units could assume any level of activity over the range 0.0–1.0.)

It should be made clear at this point that the "back-propagation" algorithm was used purely as a formal technique for constructing a network with a particular set of properties. This is not a model of developmental neurobiology, and no claims are made about the biological significance of the process by which the network was created. The focus of interest here will be on the properties of the mature network.

As input for the network, we generated 2,000 images of elliptic paraboloid surfaces. Each paraboloid differed in the magnitudes of the principal curvatures, the orientation (principal direction) of the minimum curvature (the maximum curvature is always at right angles to this orientation), the slant and tilt of the illumination direction, and the location of the center of the surface within the overall image. Image parameters were all selected with a uniform random distribution. Principal curvatures were generated over the four-octave range 1/2° to 1/32°. The direction of minimum curvature was between 0° and 180°. Light tilt was also between 0° and 180°, and light slant was between 0° and 60°. The center of the paraboloid surface could lie anywhere within a circular disk comprising the central third of the image. These 2,000 images served as the training corpus for the neural network learning algorithm.

The particular network we used had three layers: an input unit layer, an output unit layer, and a hidden unit layer between them (fig. 11.3a).

There are no lateral interactions between units within a layer, nor any feedback connections. Each unit in a layer is connected to every unit in the subsequent layer. The overall organization of the network is seen in fig. 11.4c, which shows the response of the fully developed network to a typical input image. The two hexagonal regions at the bottom represent responses of the 122 input units. Responses of the 27 units in the hidden layer are represented in the 3 x 9 rectangular array above the hexagons. Finally, the responses of the 24 units in the output layer are represented in the 4 x 6 array at the top. In all cases, areas of the black squares are proportional to the activity of a particular unit. These three layers will be more fully described below.

The response properties of the input and output units were predefined for the network, based on what operations we wanted the network to perform, as well as being constrained by biological plausibility. Through the learning algorithm, the network proceeded to develop connections between the input units and the hidden units (the receptive fields of the hidden units), and connections between the hidden units and the output units (the projective fields of the hidden units). These hidden unit response properties essentially act as a mapping, or transform, that converts the inputs to the desired outputs. Before further considering the hidden units and their receptive fields, we describe the properties of the input and output units. The input layer consisted of two hexagonal spatial arrays of units, called the on-center units and off-center units (for an explanation of these cell types see chapter 10; only the spatial and not the temporal aspects of the receptive fields were incorporated into our model). A hexagonal array was chosen over a square array because we were interested in determining stimulus orientations, and a hexagonal array has a greater degree of rotational symmetry than a square one (biologically, there is too much scatter in the positions of retinal ganglion cells to readily classify them as falling into a simple geometrical lattice, whether square or hexagonal (Wässle, Boycott, and Illing 1981). The two input arrays (off-center units and on-center units) were superimposed on each other, so that each point of the image was sampled by both an on-center unit and an off-center unit. Each of these arrays consisted of 61 units, for a total of 122 units in the input layer (61 happens to come out to an even number on a hexagonal array). The receptive field of each input unit was the Laplacian of a two-dimensional Gaussian, or in other words the classic circularly symmetric center-surround receptive field found in the retina and lateral geniculate nucleus. This receptive field organization is illustrated in fig. 11.3b, which also shows that the receptive fields of the input units were extensively overlapped

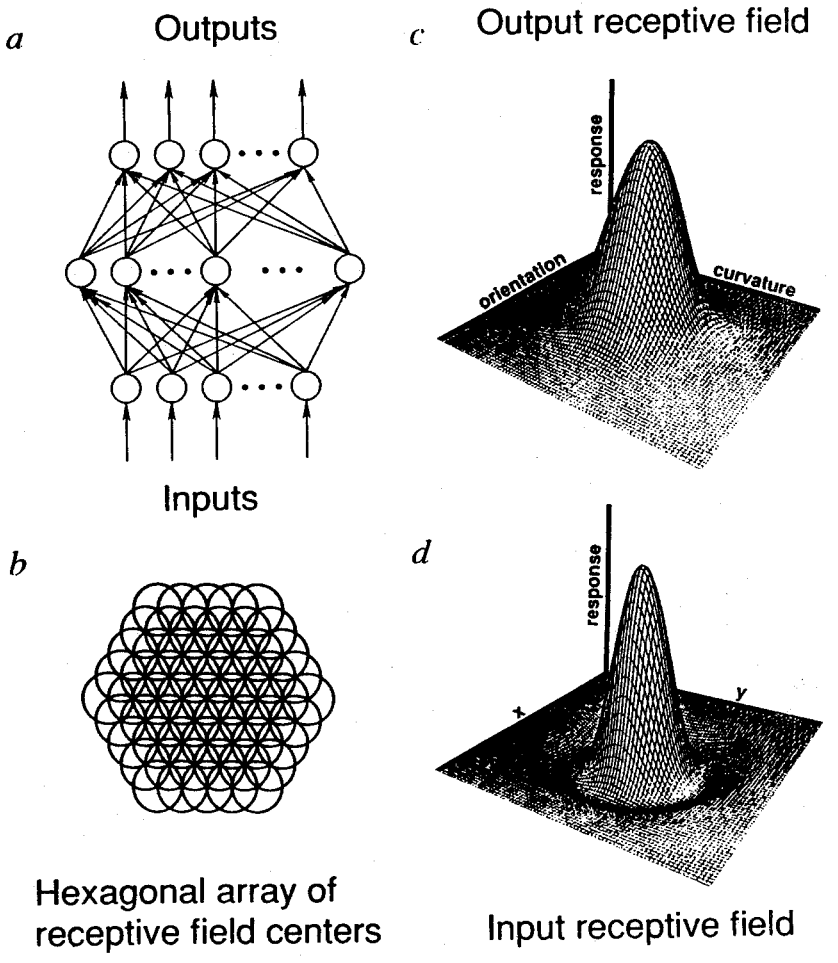

Figure 11.3
(a) Schematic diagram of network used here, having three layers: an input layer (122 units), a middle "hidden" layer (27 units), and an output layer (27 units). (b) Receptive field of an input unit, defined as the Laplacian of a Gaussian. This figure shows an on-center unit. The network also included off-center units, in which excitatory and inhibitory lobes were reversed. (c) Input units were organized into hexagonal arrays. Circles represent receptive field centers. The input image was sampled by two superimposed hexagonal arrays, representing on-center and off-center cells. (d) Output units had two-dimensional tuning curves in a parameter space defined by curvature orientation and curvature magnitude. Each of the 24 output units had peak responses at a different combination of orientation and magnitude.

(fig. 11.3c). The receptive fields of on-center units had excitatory centers and inhibitory surrounds, while off-center units had opposite polarity. Responses of the input units to an image were determined simply by convolving the image with the units' receptive fields. Aside from being biologically motivated, choosing these center-surround input receptive fields (Laplacian of a Gaussian) is advantageous from a computational viewpoint. Specifically, these receptive fields, acting as second-derivative operators, have responses that tend to reduce changes in image appearance arising from changes in illumination direction (Marr and Hildreth 1980).

Moving to the output units, they have responses tuned to both curvature magnitude and curvature orientation, as illustrated in fig. 11.3d. In other words, they have receptive fields tuned to a local region in a two-dimensional magnitude-orientation parameter space (in contrast to the input units, which are tuned to a local region of geometrical space). The equation defining output responses is:

$$R = A(m)\ \ B(o) \tag{11.1}$$

where $A(m)$ is Gaussian as a function of the logarithm of curvature magnitude (lognormal function), and $B(o)$ is a Gaussian function of curvature orientation. This sort of multidimensional response is typical of those found in cells of the cortex, although cells responding specifically to curvature have not been demonstrated. The output units had curvature tuning curves $A(m)$ with peaks of either $1/8°$ or $-1/8°$, depending on whether the unit responded to a convex or a concave surface. Half-width bandwidth of curvature tuning was one octave at $1/e$ height. Orientation tuning curves $B(o)$ had peaks set to 0°, 30°, 60°, 90°, 120°, and 150°, with half-width bandwidth of 30° at $1/e$ height.

However, the problem with having a nonmonotonic, multidimensional response is that the signal from a single unit is degenerate. There are an infinite number of combinations of curvature and orientation that give an identical response. The way to solve this ambiguity is to have the desired value represented in a distributed fashion (distributed coding), by the joint activity of a population of broadly tuned units in which the units have overlapping receptive fields in the relevant parameter space (in this case curvature magnitude and orientation).

The most familiar example of this kind of distributed coding or representation is found in color vision. The responses of any one of the three broadly tuned color receptors is ambiguous, but the relative activities of all three allow one precisely to discriminate a very large number of colors. Note the economy of this form of encoding; it is possible to form fine

discriminations with only a very small number of coarsely tuned units, as opposed to requiring a large number of narrowly tuned, nonoverlapping units (Hinton, McClelland, and Rumelhart 1986; Sejnowski 1988). The output representation of parameters in the model under consideration here will follow the coarse tuning approach. With this brief description of the concept of distributed representations, we can now examine the actual output representation used here, as illustrated in fig. 11.4. Again, the output units are represented by the 4 x 6 array at the top of the figure. The output units have tuning curves that are overlapping with peaks at different curvature orientations as one moves horizontally along a row. For the image used in this example, the curvature orientation was 10°. One can see from the size of the black squares for the output units along a row that the largest responses come from units that have peak responses close to 10°, and responses drop off as orientation tuning moves away from that value. The orientation value specified by any one unit is ambiguous, but the joint activities serve to precisely define orientation.

Moving vertically from top to bottom, the four rows represent different curvatures. The top two rows represent the value of the smaller of the two principal curvatures and the bottom two rows represent the larger principal curvature. Within each of those two pairs of rows, the top row responds if the curvature is positive (convex surface) and the bottom one responds if the curvature is negative (concave surface).

Unfortunately, this manner of representing curvature leaves the curvature magnitude ambiguous (unlike the situation in the orientation domain). For example, if a unit gives a small response, we don't know if it is because curvature is above or below the peak of the tuning curve. This is because there are no sets of output units tuned to different but overlapping ranges of curvatures, as there were units tuned to different but overlapping ranges of orientations (peak values of curvature are at either $1/8°$ or $-1/8°$, which are far enough apart so as not to overlap).

The way to remedy the situation would have been to introduce a greater number of output units, having overlapping tuning curves in the curvature domain. Having units sensitive to different ranges of curvatures is equivalent to making them sensitive to different spatial scales; surfaces with large curvatures have variations in reflected light intensities occurring at fine spatial scales, and small curvatures have variations occurring at broad spatial scales. Therefore, constructing a network whose outputs have overlapping curvature tuning curves would involve sampling the input image at different spatial scales (i.e., repeatedly convolving the input image with center-surround input receptive fields with

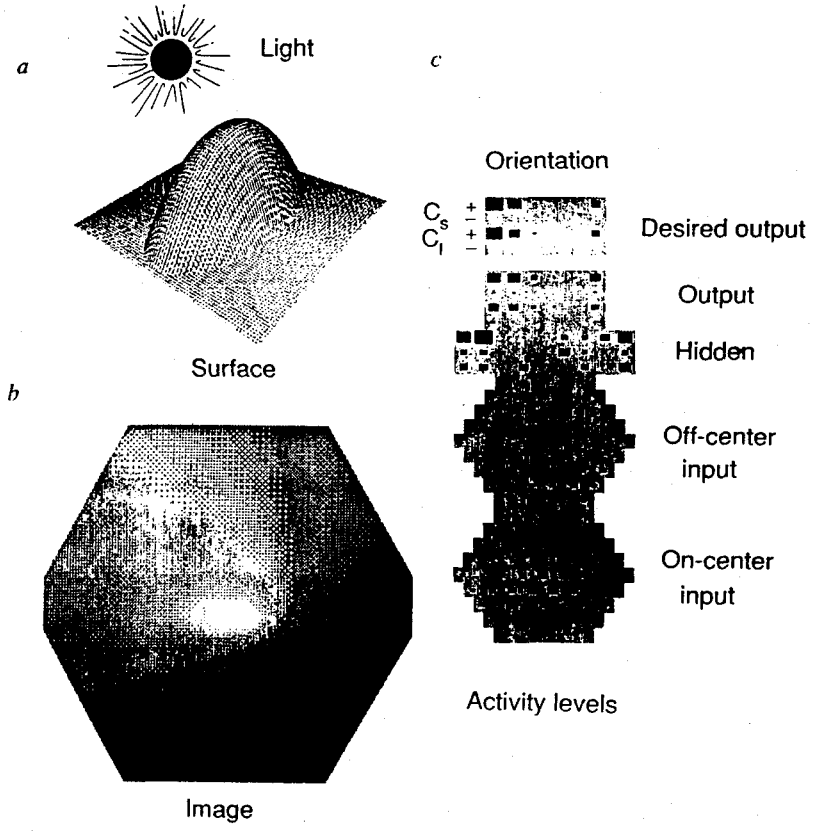

Figure 11.4
Typical input image and the resulting activity levels within a fully trained network. (a) Example of elliptic paraboloid surface (flat base did not fall within input field of the network). (b) Example of input image synthesized from light reflected off the surface of an elliptic paraboloid surface. A set of 2000 such images were used to train the network. (c) Responses of network to the image. Double hexagons show responses of on-center and off-center input units. Area of a black square is proportional to a unit's activity. Resulting activities in the 27 hidden units are shown in the 3x9 array above the hexagons, and activities of the 24 output units are shown in a 4x6 array at the top. This output should be compared with the 4x6 array at the very top (separated from the rest), showing the correct response for the image. The 4x6 array is arranged as follows. The six columns correspond to different peaks in orientation tuning. Rows correspond to different curvature magnitudes.

different diameters, instead of a single diameter as has been done here). However, this has not been implemented here. The question of the integrating multiple spatial scales has been deferred until the neural network is presented with stimuli more complex than the smooth paraboloid surfaces used here.

11.3.4 Properties of the Fully Developed Network

A network containing 122 input units, 27 hidden units, and 24 output units was trained in the manner described earlier. The 2,000 input images were presented to the network one at a time, and after each presentation the connection strengths were changed slightly to make the values of the units on the output layer compare more closely with the desired output values. Around 40,000 trials were required before performance of the network reached a plateau at a median correlation of 0.88 between actual outputs and the correct outputs over the 2,000 input images used to train the network. The connection strengths developed by the network are presented in fig. 11.5, which will be fully described below. Learning curves (correlation as a function of number of trials) for different numbers of hidden units are shown in fig. 11.6, although only results for 27 hidden units will be discussed in detail below. The learning curves show that adding more hidden units improves performance up to a point, but beyond that, additional hidden units contribute nothing. This observation is corroborated by the finding that if there are too many hidden units, the superfluous ones fail to develop strong synaptic connections. Note also that a network with no hidden units (direct connections from the input to the output layers) performs better than a network with three hidden units. There is a bottleneck caused by the hidden layer, which reduces the total number of weights when there are only three hidden units. The network with no hidden units is essentially a perceptron and its performance is surprisingly good; this is probably a consequence of the on-center and off-center preprocessing stage, which is evidently a good input representation. However, the one-layer network fails badly at many images which would pose no difficult for humans. This indicates that some part of the problem is second-order or higher.

Ability of the network to generalize (produce the correct output when presented with patterns it had never seen before) was tested in the following manner. The set of 2,000 images was randomly divided in half. The network was trained on one set of 1,000 images, and then the mature network was tested (without learning) on the other half. The results are given in fig. 11.7, as histograms of all the correlation coefficients between actual and correct output for the 1,000 training images, and also

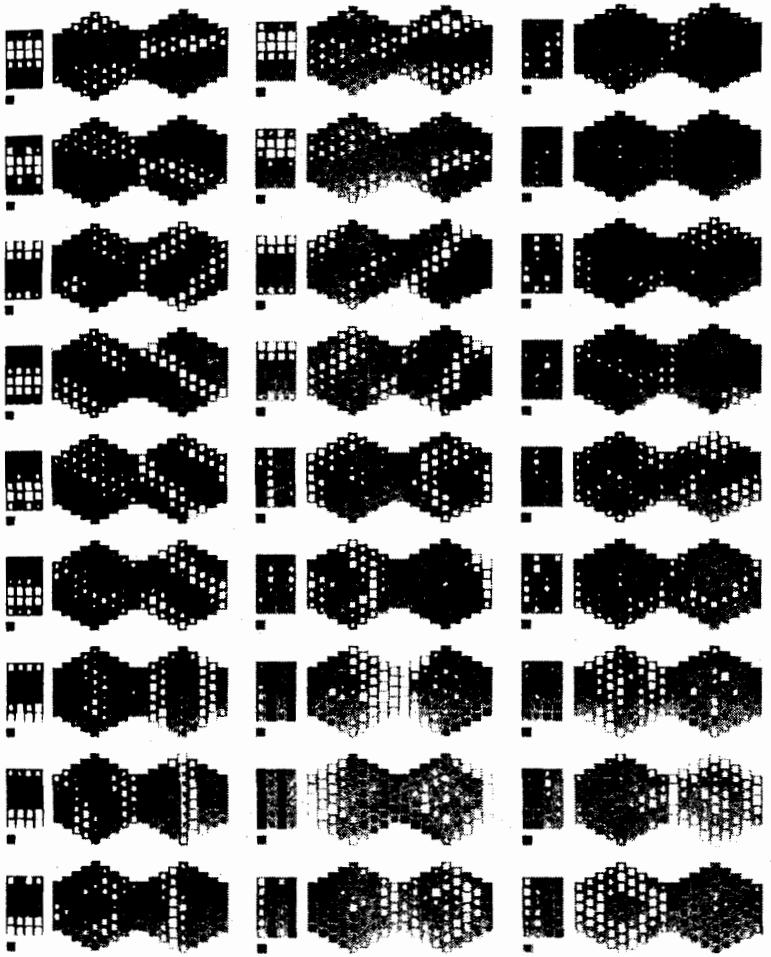

Figure 11.5
Hinton diagram showing connection strengths in the network. Each of the 27 hidden units is represented by one hourglass-shaped icon, showing an input receptive field from on- and off-center units (double hexagons), and an output projective field (4x6 array at the top). Organization of the 4x6 array is the same as described in fig. 11.3. Excitatory weights are white, inhibitory ones are black, and the area of a square is proportional to the connection strength it represented. Isolated square at the upper left of each icon indicates the unit's bias (equivalent to a negative threshold). Black horizontal lines group together units of the same type based on receptive field organization.

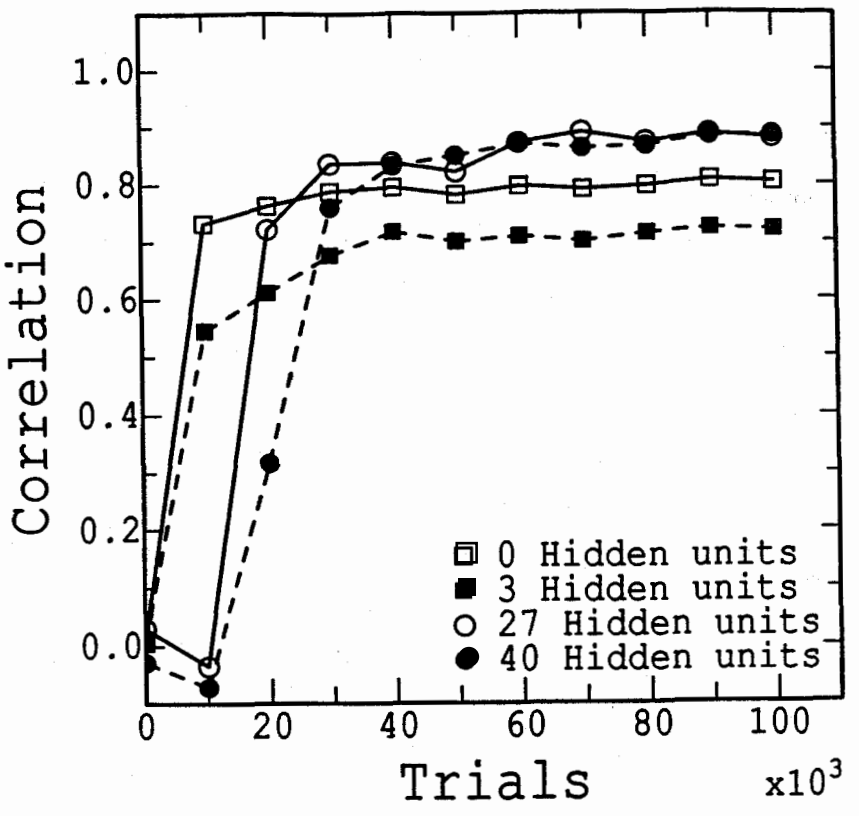

Figure 11.6
Learning curves for the network, showing correlation between actual and correct responses of the output units as a function of the number of learning trials. Learning curves for networks with different numbers of hidden units are shown here.

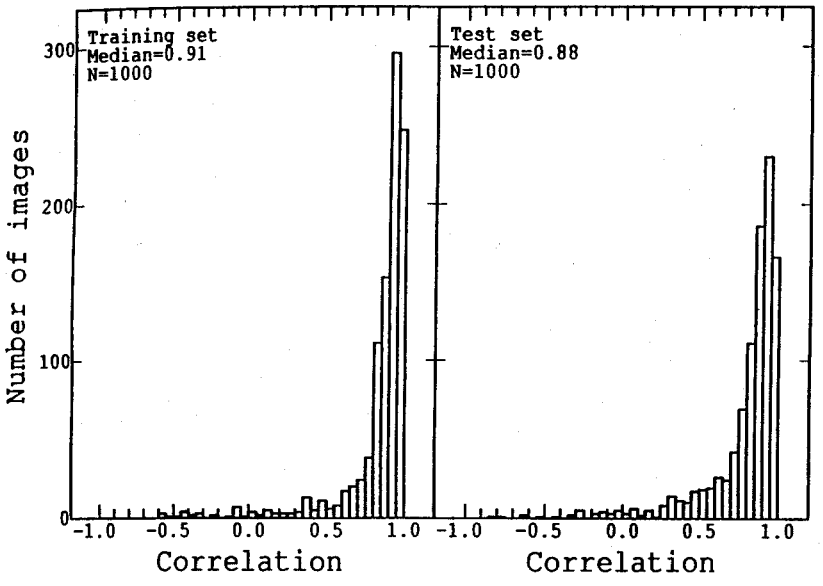

Figure 11.7
Demonstration of network's ability to generalize and give correct responses to images that were not part of its training set. The network was trained on a set of 1,000 images, and then tested on a different set of 1,000 images. (a) Distribution of correlation coefficients between actual and correct outputs for the images of the training set. (b) Distribution of correlation coefficients for images not part of the training set.

the 1,000 test images. These histograms indicate that the network generalizes well to patterns that were not in the training set.

Figure 11.5 is of central importance in understanding the fully developed network. It is a Hinton diagram, showing all the connection strengths between units in the fully developed network. Each of the 27 hidden units is represented by one of the grey hourglass-shaped figures. Connection strengths are represented as black and white squares of varying size. The white squares are excitatory weights, the black squares are inhibitory weights, and the area of the square is proportional to the magnitude of the weight. Within each hourglass figure two sets of connections are shown. First, there are the connection strengths from all input units to that particular hidden unit, and second, connection strengths from that hidden unit to all the output units. The two hexagonal arrays on the bottom of each figure show the connections from

the on-center and off-center input arrays to the hidden unit (hidden unit receptive field), and the 4 x 6 rectangular array at the top are the connections from the hidden unit to units in the output layer (hidden unit projective field). Finally, the value of the bias, or negative threshold, is shown in the isolated square at the upper left corner of each hidden unit figure.

The pattern of excitatory and inhibitory connections in the two hexagonal input arrays can be interpreted as receptive fields of the hidden units. Most of the hidden units appear to be orientation-tuned to a variety of directions. These oriented fields have several excitatory and inhibitory lobes, which may occur in various phases. This is the pattern found in simple cells in cat and monkey visual cortex, which are often fit with Gabor functions (DeValois, DeValois, and Yund 1979; Kulikowski and Bishop 1981; Andrews and Pollen 1979; Wilson and Sherman 1976. See also fig. 10.5; the earlier studies of Hubel and Wiesel (1962, 1965) had focused on the central two or three lobes, which are the most prominent). In addition to units that were clearly orientation-selective, a few units had receptive fields that were more or less circularly symmetric.

Upon examining projective fields, three types become apparent. Type 1 has a vertical pattern of organization to the 4 x 6 array of weights, type 2 has a horizontal organization with alternate rows similar, and type 3 has a horizontal organization with adjacent rows similar. These classes of hidden units appear to provide information to output units about orientation of the principal curvatures (type 1), their signs (convexity/concavity) (type 2), and their relative magnitudes (type 3). A few hidden units were difficult to classify, and several failed to develop large weights.

Overall, then, we distinguish three types of hidden units: those providing orientation information, those providing information about the signs of principal curvatures, and those providing information about relative magnitudes of the two principal curvatures. It should be noted, however, that the receptive fields of some hidden units are somewhat irregular, and that combinations of units might be working together to provide information that is not apparent from examining single units. Incidentally, during construction of the network the three types of units always developed in a particular temporal sequence. Sign units came first, followed by orientation units, and finally magnitude units.

An interesting question is whether the hidden units in this network act as feature detectors or as parameter filters. By a feature detector we mean a unit that responds strongly when presented with an appropriate and specific stimulus and poorly to all other inputs, in essence

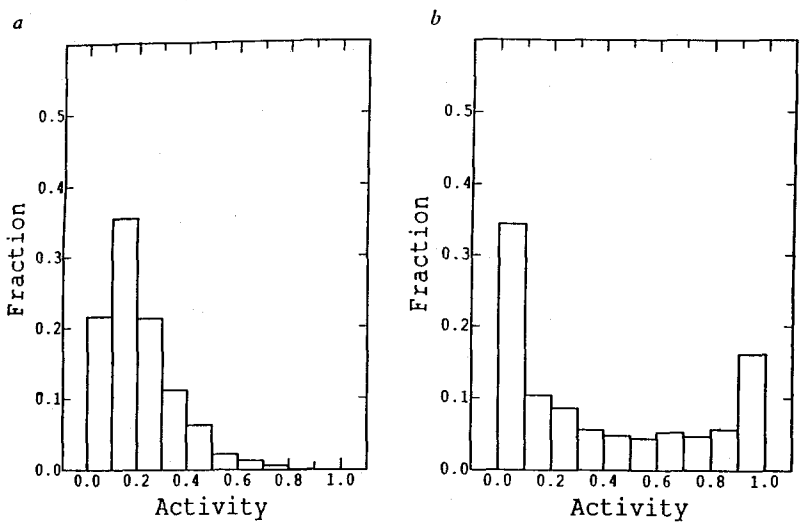

Figure 11.8
Distribution of responses of single units when the network was presented with a set of 2,000 images. (a) Unimodal distribution typical of a type 1 unit, selective for curvature orientation. Type 3 units, selective for the relative magnitudes of the two principal curvatures, also had unimodal distributions. We interpret units with this sort of response as filters indicating the values of particular parameters. (b) Bimodal distribution typical of a type 2 unit, selective for the signs of the principal curvatures (convexity/concavity). Units having this sort of response, which tended to be either fully on or off, are interpreted as feature detectors.

an all-or-nothing response. By a parameter filter we mean a unit that responds with a continuous range of activities when presented with various stimuli, a unit whose responses are defined by a rather broad tuning curve for a particular parameter. To investigate the matter, we looked at the responses of individual hidden units when presented with the corpus of 2,000 images. For each input image, the response of the unit fell between 0.0–1.0. By plotting a histogram of the unit's response levels when presented with many inputs we hoped to classify the unit, with the histogram of a parameter filter unit showing a unimodal distribution having a peak at some intermediate level of activity, and a feature detector unit having a bimodal response histogram with activities concentrated at either very high or very low levels.

We found examples of both kinds of behavior among the hidden units (fig. 11.8). Of the three hidden unit classes, the orientation units (type 1)

and the magnitude units (type 3) had unimodal distributions, and we classified them as parameter filters. In contrast, the curvature sign units (type 2) invariably had bimodal distributions in which they tended to be fully on or fully off. We interpret type 2 as feature detectors that discriminate between surface convexity and concavity.

The results of the learning procedure shown in fig. 11.5 were representative of many training runs. Each started with a different set of random weights. Similar patterns of receptive fields were always found, and the same three types of units could be found by examining the connections to the output units. However, there was variation in the details of the receptive fields, and in the number of units that did not develop any pattern of connections. It appears that only a limited number of hidden units are needed to achieve the maximum performance, which always had an a correlation close to 0.88. The extra hidden units undergo "cell death" (occurring as a consequence of the weight decay term in the learning algorithm; see also Appendix 11.A), since they serve no useful role in the network and can be eliminated without changing the performance.

The term "receptive field" of a unit was used above to refer to the pattern of excitatory and inhibitory connections arriving from the preceding layer of units. However, the true receptive field is something different; it is not a pattern of synaptic connections, but the response of that unit to a stimulus, such as a spot or bar of light, as a function of spatial position. The stimulus response depends not only on the immediate synaptic inputs to a unit, but also on the filtering properties of units in the preceding layers.

To examine these true "receptive fields" we have explored the responses of units in the network to bars of light, in essence conducting "simulated neurophysiology." The bars were varied in position, orientation, width, and length. These bar stimuli were chosen because there is an extensive experimental literature using them, so that responses of real neurons are known.

The responses of hidden units to bars were easily predictable from the pattern of excitatory and inhibitory connections they received from the input units. Tuning curves were measured for all stimulus parameters (bar position, orientation, width, and length). Just by looking at the pattern of connection strengths in the weights diagram (fig. 11.5) it was possible to form a good estimate of what the optimal bar stimulus would be. The ease in understanding hidden unit responses is not surprising, since the only intermediary between the stimulus and the hidden units is the array of simple center-surround receptive fields of the input units.

The situation was quite different when measuring responses of output units to bar stimuli. Finding an optimal stimulus took extensive trial and error. Again, this is not surprising. The response of each output unit is determined by a weighted combination of the inputs from all 27 hidden units (although in practice a smaller number of hidden units tend to predominate). In other words, the response of each output unit is some combination of all the 27 patterns shown in fig. 11.5. It is therefore very difficult to grasp intuitively what the responses to a particular stimulus will be.

Despite the complex organization of receptive fields for output units, it was possible to obtain smooth tuning curves for the various bar parameters. Tunings were generally broader than in hidden units. One feature of some output unit responses was the presence of strong "end-stopped inhibition." Responses dropped off precipitously when bar length was extended beyond a certain point. Also, as the bar was swept across the network in a direction transverse to the optimal orientation, the output units responded over a broader region than the hidden units did (about 1.5–2.0 times as wide). This is as expected, since output units receive convergent input from a number of hidden units.

The responses of units in the hidden and output layers are reminiscent of the behavior of some units that have been found in primary visual cortex (Hubel and Wiesel 1962, 1965). It is possible to think of the hidden units as having simple cell-like properties, and output units as behaving like some types of complex cells. However, in drawing this analogy it should be kept in mind that the relationship between hidden units and output units is strictly hierarchical (in the sense that responses of output units are entirely synthesized from inputs from the preceding hidden units), while the extent to which such a hierarchical relationship holds for simple and complex cells has not yet been settled (Gilbert 1983; Bolz and Gilbert 1986).

A major point to make about the receptive field properties of units in this network, and perhaps the most important lesson to be drawn from this entire modeling study, is the following. Knowledge of the receptive field of a unit does not appear sufficient to deduce the function of that unit within a network. The receptive fields shown in fig. 11.5 can easily be interpreted as bar detectors, or edge detectors (or from another point of view, spatial frequency filters). What they are in fact doing is extracting principal curvatures from shaded surfaces. Yet it seems unlikely that this interpretation would have occurred to someone. The questions this network raises for standard interpretations of receptive fields of real neurons are obvious. While we can determine the receptive

field of a neuron, there is at present no practical method of determining its projective field. Yet it would appear that the projective field is as important as the receptive field in determining the function of a neuron. The network model provides an alternative interpretation of the observed properties of cortical cells, namely that they can be used to compute shape from shading rather than, or in addition to, detecting edges. The information contained in the shaded portions of objects in images can partially activate the simple and complex cells in visual cortex, and these responses can be used to extract curvature parameters from the image. It might prove interesting to test the response properties of cortical cells with curved surfaces similar to those used here.

Although the response properties of the processing units in our model were similar in some respects to the properties of neurons in visual cortex, the way that these response properties arise could be quite different. For example, projections from the lateral geniculate are all excitatory, but we have allowed feedforward connections that are both excitatory and inhibitory. Also, the oriented responses of cells in visual cortex may arise in part from inhibitory interneurons, as discussed in detail in chapter 10 (see also Sillito 1985). However, our network is not meant to be a literal model of actual cortical circuitry, which is certainly much more complex, but rather a model of the representation in visual cortex of information about surfaces in the visual field. The same representations could be constructed differently in different systems, as indeed the orientation tuning of neurons in different species might also have different origins though they might serve the same function.

Nonetheless, we believe that our model can evolve toward a more detailed account of real cortical circuitry as more information is found about detailed patterns of connectivity between different cell types in cortex. As a first step towards creating a more realistic network, we constructed one that constrained all the connections from the on-center and off-center cells in the input layer to be excitatory, as found for principal cells in the lateral geniculate nucleus. The average performance of the network was nearly identical to the previous network, and same response properties were found for the hidden units as before. Evidently, the off-center cells were able to substitute for the on-center inhibition, and vice versa.

As was mentioned earlier, the network model handles shape from shading as a problem in pattern recognition, looking for particular configurations of light and dark over an extended region. In this respect the network resembles the model of Koenderink and Van Doorn (1980) (which they described but never implemented) and a more recent model by

Pentland (1988) that uses the sums of filters much like our hidden units. The information about the curvature parameters of a particular patch of image was contained in the distributed pattern of activity in hidden unit layer. Whatever precision the network loses by adopting such a delocalized analysis may be compensated by increased robustness.

Humans are only able to extract an approximate estimate of the curvature from local shading information, and it will be interesting to compare the accuracy of the network constructed here with that of humans under controlled psychophysical conditions (Mingolla and Todd 1986; Ramachandran 1988). It is also clear that humans use other cues to estimate curvatures, such as the outline of bounding contours. This suggests that the network presented here should be considered only a small part of a much more complex system that uses multiple cues.

11.4 Conclusions

It is clear that the present generation of neural models cannot begin to reflect the complexity of the real nervous system. From anatomy and physiology we know that the visual system is a tangled web of multiple inputs, feedback loops, and lateral interaction, and moreover that each unit within that web is a complex entity in itself, with the various nonlinear temporal and spatial integrative aspects of the dendritic tree being one part of that complexity. It is from a confrontation with this complexity, from a desire to see some pattern to it, that one is led to attempt the extraction of essential features and incorporate them into a simple model. This leads to the most difficult part in constructing a model, which is to decide what is an "essential feature" and what is "simple." These are ultimately matters of intuition and judgment, although of course the choices are related to the types of questions being asked (whether they concern psychological phenomena or biophysical problems).

The proposed network model of rivalry is intended as a functional description rather than a realistic model, and the "equivalent circuit" in the model may eventually translate into a much more complex network in the real nervous system. For example, the individual neurons in the model are likely to represent populations of real neurons that have some functional properties in common. Another limitation of the model is the difficulty in precisely specifying where in the system the neurons should be found. However, the model indicates places to look and predictions for what might be found there. The model suggests that it is worthwhile

to investigate sites of reciprocal inhibition prior to the site of binocular convergence. This suggestion depends on both the psychophysical data and knowledge of mechanisms that are known to exist in the nervous system in the areas of interest. All of this provides experimentalists some justification to embark on a line of investigation that they might not otherwise have thought interesting.

An electrical circuit analog was used to study the nonlinear dynamics of the underlying neural model of binocular rivalry. It might be objected that the electrical circuit, which is made of transistors and resistors, is not relevant to the biological problem, because neurons have properties that are different at the biophysical level. Many of the qualitative features of the electrical analog circuit are insensitive to the detailed biophysical properties of the nonlinear summing elements; when such properties are important, then provision must be made for incorporating these properties into the electrical analog. In the case of the simple circuit for binocular rivalry, we can verify by simulation that a neural model with plausible biophysical properties displays the same qualitative properties as the electrical circuit. The value of the electrical circuit model is that it can be physically built and studied under real-time conditions, which is usually not possible with digital simulations. With analog VLSI technology (Mead 1989) it is possible to build many thousands of such circuits and to study their properties under conditions that would be difficult or impossible otherwise.

In the proposed network for extracting curvature parameters from shaded images, the receptive field properties needed at the intermediate level of hidden units were similar to the properties of simple cells in primary visual cortex. These properties were not determined by the intuition of the model builder, but by the learning algorithm. The modeler specified the function that the network was required to perform only by giving examples of inputs and the desired outputs. Once trained, the network model was able to compute curvature parameters for new images nearly as well as for the ones it was trained on. The learning algorithm that was used to construct the model need not have a counterpart in the nervous system for the resulting model to have validity; the process is not intended as a model of development but simply as a technique for generating hypotheses for what might be found in the nervous system.

It might be instructive to compare our approach with that taken in chapter 10, where a realistic model of the retina, LGN and layer IV of visual cortex were simulated. The questions that motivated that model were structural ones such as the projection patterns of LGN cells in cortex and the origin of orientation selectivity. The question of what

function the oriented simple cells had was not raised—it was enough for the model to mimic the actual responses of simple cells under a variety of conditions. In contrast, the questions that motivated our model related to the properties of the images and the information that could be extracted by the network. In our model of binocular rivalry, we were led to the hypothesis that feedback connections to the LGN are able to gate incoming sensory information. In our cortical model of shape-from-shading, we found units in the hidden layer of our network that were similar to simple cells; our primary concern was in interpreting the function of these cells and not in their genesis. It was only after examining the outputs of these cells—their projective fields—that we were able to understand their function in the network. In the realistic simulation of oriented cortical neurons presented in chapter 10, the cells had no output projections, and hence could not have any function. It is clear that these two models are complementary in their strengths and weaknesses. Both types of models are needed if we are to understand the visual cortex at all levels.

Neural network modeling is still at an early stage of development, but it is already clear that new principles are emerging concerning the representation of information in neural populations, and transformations that are possible with these coding schemes. For example, Georgopoulos, Schwartz, and Kettner (1986) have shown that in motor cortex, information about the intended direction of arm movement is distributed in populations of neurons that are broadly tuned to the direction. Lee et al. (1988) have presented evidence for the coarsely coded representation of eye movements in the deeper layers of the superior colliculus. Zipser and Andersen (1987) have applied the same approach used here to the problem of transforming from retina-based coordinates to head-centered coordinates. They report that the properties of the hidden units in their model are similar in essential respects to those of some neurons in parietal cortex. The success of their model and ours depended to a large extent on incorporating what was known about the single-unit properties and the style of representation found in cerebral cortex into the models. Learning algorithms provide a new technique for drawing out the implications of these assumptions and exploring some of the principles of distributed processing in sensory and motor systems. Ultimately, evaluation of this claim rests upon the ability of the models to provide results useful in organizing data and suggesting new experiments.

This work was supported by a Presidential Young Investigator Award to Terrence J. Sejnowski and grants from the National Science Foundation, Sloan Foundation, General Electric Corporation, Allied Corporation Foundation, Richard Lounsbery Foundation, Seaver Institute, Air Force Office of Scientific Research, and Office of Naval Research. The address of Sidney Lehky is Department of Biophysics, Johns Hopkins University, Baltimore, Maryland 21218. The address of Terrence Sejnowski is The Salk Institute, P. O. Box 85800, San Diego, California 92138. Electronic mail should be addressed to terry%sdbio2@ucsd.edu.