can be found in [14], while Van der Lugt [20] discusses these issues in the context of optical data processing.


1. Barrow, H. G., and Tenenbaum, J. M., "Computational Vision," *Proceedings IEEE*, vol. 69, no. 5, May 1981, pp. 572–595.
2. Bauer, M. E., "Technological Basis and Applications of Remote Sensing of the Earth's Resources," *IEEE Transactions on Geoscience Electronics*, vol. GE-14, no. 1, January 1976, pp. 3–9.
3. Biberman, L. M., and Nudelman, S. (eds.), "Photoelectric Imaging Devices," Plenum, New York, 1971.
4. Bruss, A. R., "Some Properties of Discontinuities in the Image Irradiance Equation," Massachusetts Institute of Technology, Artificial Intelligence Laboratory, A.I. Memo no. 517, April 1979.
5. Chien, R. T., and Snyder, W. E., "Hardware for Visual Image Processing," *IEEE Transactions on Circuits and Systems*, vol. CAS-22, no. 6, June 1975, pp. 541–551.
6. Clark, W. A., "From Electron Mobility to Logical Structure: A View of Integrated Circuits," *Computing Surveys*, vol. 12, no. 3, September 1980, pp. 325–356.
7. Danielsson, P.-E., and Levialdi, S., "Computer Architectures for Pictorial Information Systems," *Computer*, vol. 14, no. 11, November 1981, pp. 53–67.
8. Haralick, R. M., "Glossary and Index to Remotely Sensed Images Pattern Recognition Concepts," *Pattern Recognition*, vol. 5, no. 4, 1973, pp. 391–403.
9. Horn, B. K. P., "SEQUINS and QUILLS—Representations for Surface Topography," Massachusetts Institute of Technology, Artificial Intelligence Laboratory, A.I. Memo no. 536, May 1979.
10. Ikeuchi, K., and Horn, B. K. P., An Application of the Photometric Stereo Method, Massachusetts Institute of Technology, Artificial Intelligence Laboratory, A.I. Memo no. 539, August 1979.
11. Ikeuchi, K., "Numerical Shape from Shading and Occluding Contours in a Single View," Massachusetts Institute of Technology, Artificial Intelligence Laboratory, A.I. Memo no. 566, February 1980.
12. Ikeuchi, K., "Shape from Regular Patterns (An Example of Constraint Propagation in Vision)," Massachusetts Institute of Technology, Artificial Intelligence Laboratory, A.I. Memo no. 567, March 1980.
13. Nagy, G., "Digital Image Processing Activities in Remote Sensing for Earth Resources," *Proceedings IEEE*, vol. 60, no. 10, October 1972, pp. 1177–1200.
14. Nawrath, R., and Serra, J., "Quantitative Image Analysis: Applications Using Sequential Transformations," *Microscopica Acta*, vol. 82, no. 2, September 1979, pp. 113–128.
15. Palmieri, G., "Image Devices for Pattern Recognition," *Pattern Recognition*, vol. 3, no. 2, July 1971, pp. 157–168.
16. Proceedings Society of Photo-Optical Instrumentation Engineers, vol. 48, "Acquisition and Analysis of Pictorial Data," *The Modern Science of Imagery Conference*, San Diego, Aug. 19–20, 1974.
17. Smith, D. A., "Using Enhanced Spherical Images for Object Representation," Massachusetts Institute of Technology, Artificial Intelligence Laboratory, A.I. Memo no. 530, May 1979.
18. Tanimoto, S. L., "Advances in Software Engineering and Their Relations to Pattern Recognition and Image Processing," *Proceedings 5th International Conference Pattern Recognition*, Miami Beach, Dec. 1–4, 1980, pp. 734–741.
19. Turin, G. L., "An Introduction to Digital Matched Filters," *Proceedings IEEE*, vol. 64, no. 7, July 1976, pp. 1092–1112.
20. Van der Lugt, A., "A Review of Optical Data Processing Techniques," *Optica Acta*, vol. 15, no. 1, 1968, pp. 1–33.
21. Young, I. T., Balasubramanian, Dunbar, D. L., Peverini, R. L., and Bishop, R. P., "SSAM: Solid-State Automated Microscope," *IEEE Transactions on Biomedical Engineering*, vol. BME-29, no. 2, February 1982, pp. 70–82.


Levine, "Vision in man and machine," McGraw Hill 1985

CHAPTER

# THREE

# BIOLOGICAL VISION SYSTEMS

## 3.1 INTRODUCTION

Humans and animals are capable of accomplishing a wide range of perceptual tasks, most of which we understand very little about from the point of view of how the brain functions. We have noted that the visual system may be considered as a mechanism that converts input light patterns into perceptions, which are usually reported verbally or by a motor action. Most of our basic knowledge about this activity relates to the early stages of the processing, as will become apparent from the descriptions in this chapter and those following. Moreover, even the early stages are not fully understood.

If we model the visual system as a living optical transduction device followed by a computer, we must appreciate that even though the transducer is physically in the eye, some portion of the image processing may also take place there. In fact, the less advanced the living form, the greater the fraction of its computational power that is located in the eye. Physically it is convenient to distinguish three stages, as shown in Figure 3.1. None of these is simple in design, although the first stage of optical processing is the least complex, mostly because the smallest amount of neural activity is associated with it. The second, retinal stage, providing the sensory transduction as well as some cellular processing, is also located in the eye. The last stage in this sequence is essentially a rubric for the myriad of connections and many complex levels of processing. These begin at the retinal level and continue in an analysis hierarchy (see Figure 1.10) toward and within the brain, the so-called meat machine or wetware.

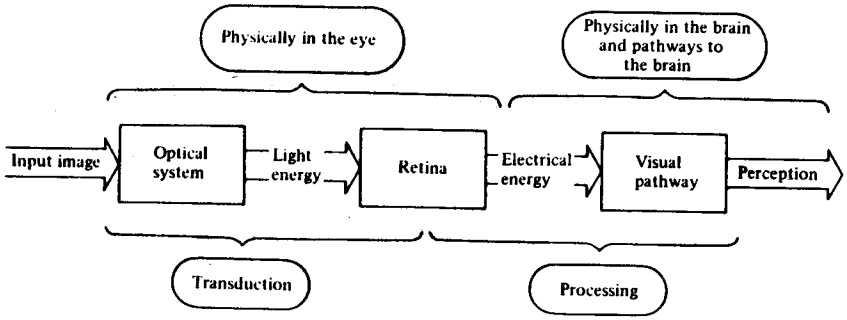



**Figure 3.1** The human visual system viewed as consisting of a camera and a computer, providing for transduction and processing, respectively.

Section 3.2 will describe the optical system, the "front-end" or "camera" aspects of the human visual system. At this point light energy is input and is controlled somewhat, but the major function of the system is to provide a focused image at the retina. Therefore, we see that energy conversion takes place at this interface. The retina, described in Section 3.3, serves the twofold roles of transduction and processing. First the signal is converted to a frequency-coded format by a photochemical process and then the resulting electrical signals are further analyzed. The amount of high-level retinal processing is inversely related to the intelligence and evolutionary complexity of the animal. This is true only in general terms, since the phylogeny of the vertebrate eye is unknown.

Since the eyes of the various vertebrates are not greatly different, it has been convenient for experimental purposes to examine and study such animals as cats and monkeys rather than humans. This, of course, can only be carried out to a certain point, because the visual pathway, which will be described in Section 3.4, finally outputs as a human action. However, animals have a limited facility for explaining their behavior. This serves at present as a real stumbling block to achieving further knowledge about the higher processes of the human visual apparatus. A new experimental technique must be found before any great strides can be taken to increase our comprehension of the brain.

Notwithstanding the above arguments, we do know how the neurons in the brain function as individual computational units, even if we are at a loss to explain how they cooperate and compete to provide a given perceptual experience. Section 3.5 will explain in very simplified terms a general model of cellular behavior which seems to be consistent with our present knowledge.

## 3.2 THE OPTICAL SYSTEM

In Figure 3.1, the ocular optical system of a human is seen to produce a transformation of the light energy of the visual input stimulus impinging on the



eye to an output which is similarly a high-energy signal. A horizontal cross section of the human eye is shown in Figure 3.2. The input light pattern enters the cornea and then passes in sequence through the anterior chamber, the pupil opening of the iris, the lens, and the vitreous humor before impinging on the layer of photoreceptors which constitutes the retina at the rear. The latter is responsible for the actual transduction from light energy into electrical energy in the form of a train of frequency-modulated pulses and will be discussed in detail in Section 3.4. A system block diagram showing the main stages of the optical system is shown in Figure 3.3.

The human eyeball, which is about 25 mm in diameter, is flexible, so that even this first step of optical processing results in image distortion. The stages

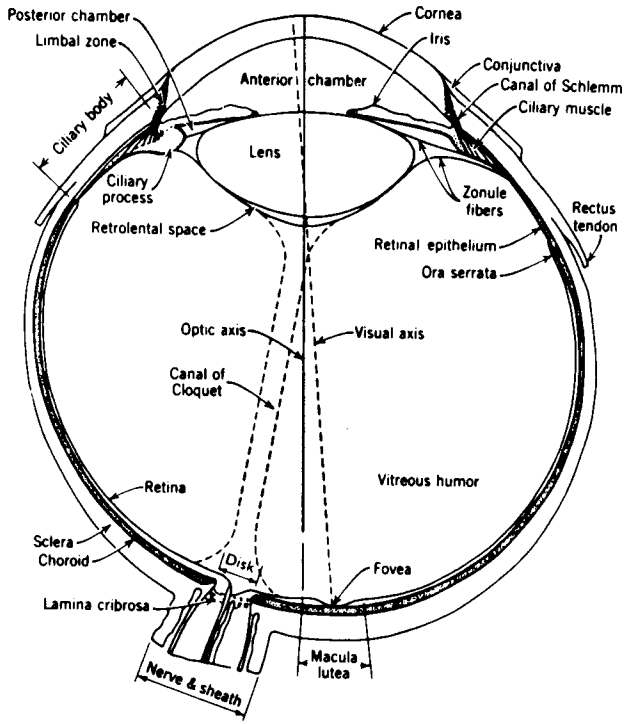



**Figure 3.2** A horizontal cross section of the human eye viewed from above and showing the different stages of the optical system. The retina appears at the rear and is well protected from environmental disturbances. [*From W. R. Uttal, "The Psychobiology of Sensory Coding," Harper & Row, New York, 1973, p. 102, in J. L. Brown, "The Structure of the Visual System," in C. H. Graham (ed.), "Vision and Visual Perception," Wiley, New York, 1965, pp. 39–59; after G. L. Walls, "The Vertebrate Eye," Cranbrook Institute of Science, Bloomfield Hills, Mich., 1942, as modified from M. Salzmann, "The Anatomy and Physiology of the Human Eyeball in the Normal State," University of Chicago Press, Chicago, 1912.*]

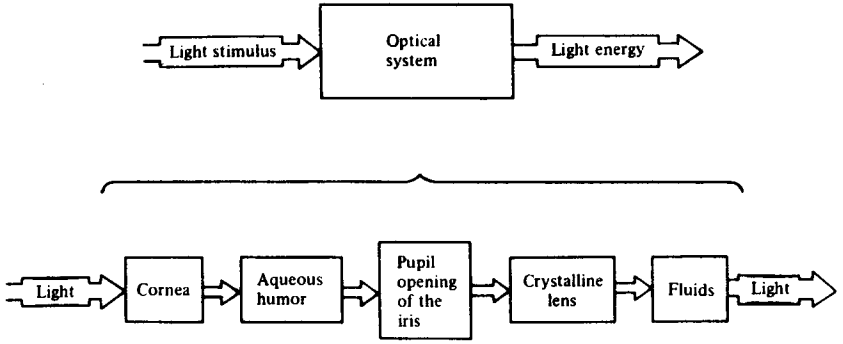



**Figure 3.3** A block diagram of the light pathway of the human eye as sketched in Figure 3.2.

shown in Figure 3.3 which the light path must traverse contain many impurities which obstruct and nonlinearly transform the signal. Evidently no correction is made for the ensuing chromatic or spherical aberrations. Remarkably, only about 50 percent of the light energy entering the cornea in fact even arrives at the retina. The optical system therefore projects a recognizable but definitely imperfect image on the receptor cells.

This method of image formation in the vertebrate eye has not necessarily been replicated by the evolutionary process in other animals. Figure 3.4 shows several interesting cases of invertebrate photoreceptor organs. For example, the limpet (Figure 3.4*a*) and the nautilus (Figure 3.4*b*) do not even possess focusing lenses and the images falling on their retinas are controlled in a very rudimentary fashion. The photoreceptors of the limpet are protected by a secretion and are located within a so-called visual pit. The latter has the effect of reducing the amount of input ambient light, thereby increasing the contrast and enhancing the effectiveness of the eye in detecting enemy shadows. Image formation in the nautilus, a spiral-shelled mollusk related to the octopus and the squid, resembles that in the simple pinhole camera, except that the eye is continuously being washed by seawater. This type of arrangement has the advantage of always keeping the image in focus, with the concomitant disadvantage of seriously reducing the amount of light falling on the retina. Other invertebrates possess more sophisticated optical systems containing lenses. In the case of the scorpion (Figure 3.4*c*), an aggressive, nocturnal, eight-legged animal, the optical system is relatively large and located externally, while in the slow-moving snail (Figure 3.4*d*) the light must first pass through retinal ephithelium, the retina, and a liquid secretion before reaching the internal lens. On the other hand, the squid (Figure 3.4*e*) has highly developed eyes which contain a lens capable of forming an image. If two eyes are necessary for the survival of the normal vertebrate (Figure 3.4*f*), then four eyes must be twice as useful. Figure 3.5 shows such a fish, the *Anableps microlepis*, which is capable of simultaneously monitoring both aquatic and aerial activity. When it is

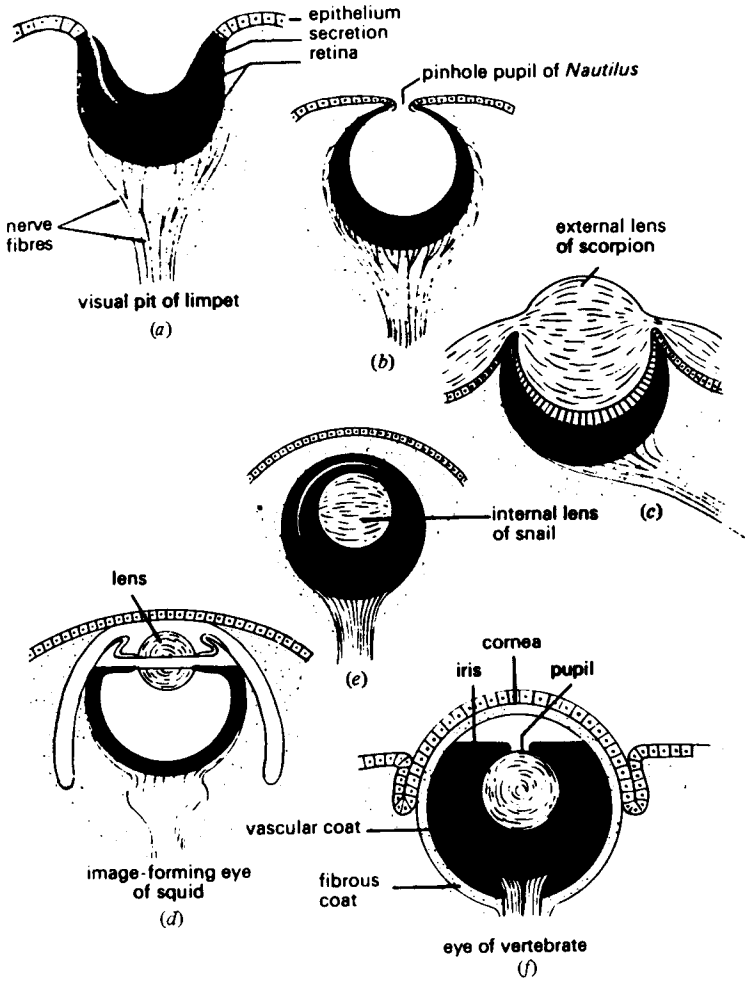
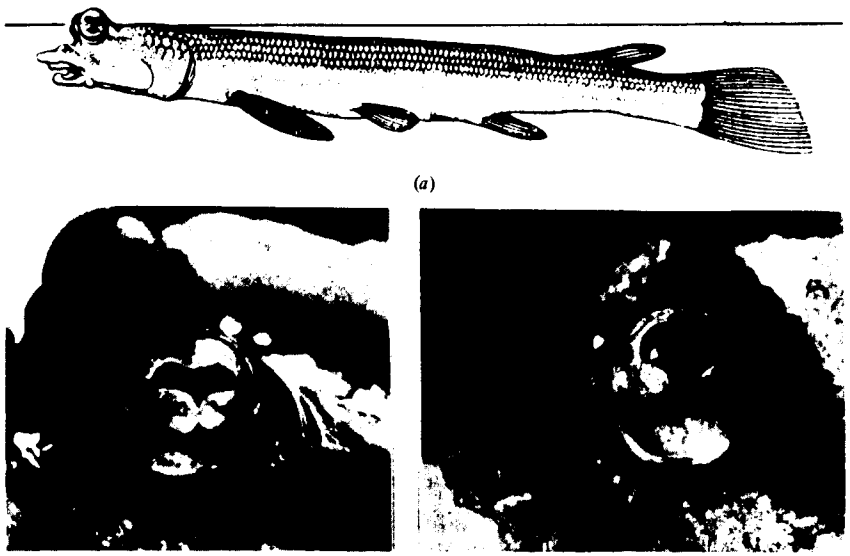



**Figure 3.4** A comparison of several interesting eyes of invertebrates (*a*–*e*) with those of a vertebrate (*f*). How and why these eyes developed as they did is not known. (*From R. L. Gregory, "Eye and Brain the Psychology of Seeing," World University Library, McGraw-Hill, New York, 1966, p. 24.*)

submerged, a horizontal flap is retracted to effectively provide only two eyes. It possesses a strong muscle capable of moving its lens to the proper position for maximum visual acuity depending on whether it is attacking an insect on the water surface or other prey underwater.

Another interesting eye is the compound eye of the invertebrate arthropod, which possesses thousands of lens facets. Each lens facet contains one

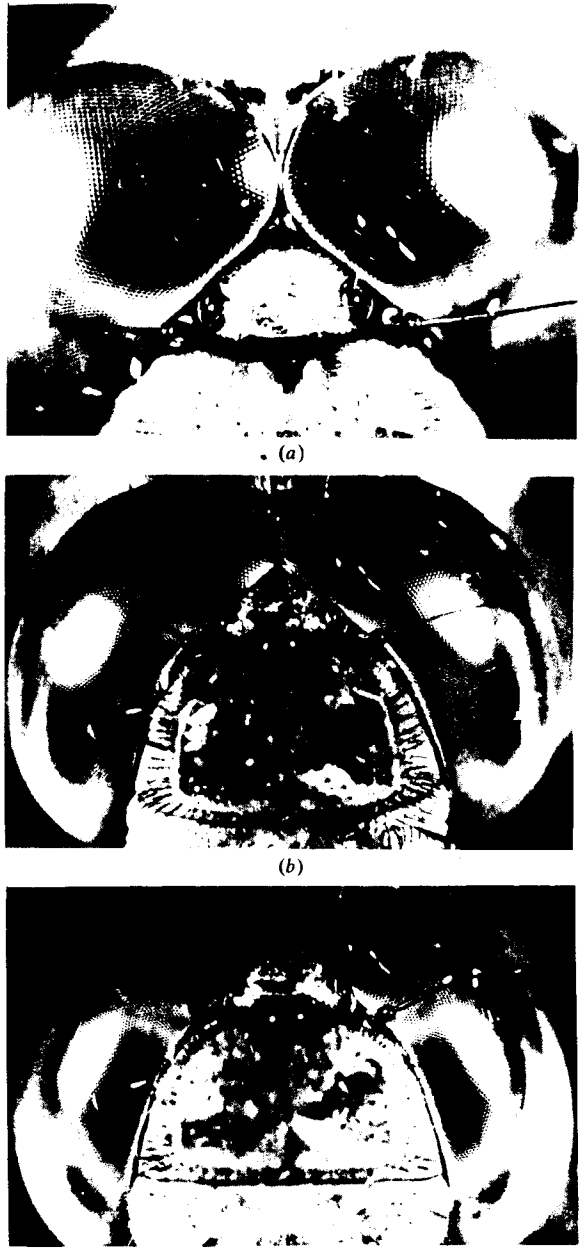


(*a*)

(*b*)

**Figure 3.5** The *Anableps microlepis* can control its visual system so that it effectively has four eyes. The light-adapted condition shown in (*a*) allows it to simultaneously view prey both above and below the waterline. In (*b*) the pupillary aperture is one complete unit, a situation which occurs when the fish is submerged and the pupil is dark-adapted. (*From H. O. Schwassman and L. Kruger, Experimental Analysis of the Visual System of the 4-Eyed Fish Anableps microlepis, Vision Research, vol. 5, 1975, pp. 269–281.*)

receptor element and is stimulated by the light directly impinging upon it. An example of an insect is shown in Figure 3.6. This mosaic arrangement would seem to yield a one-to-one mapping between the viewed light patterns and the retinal electrical signal. For the most part this is correct, except that there is some interaction at the electrical-signal-processing level. Later on we shall discuss some interesting experiments with the compound eye of the horseshoe crab, which demonstrate how this lateral interaction tends to accentuate borders between light and dark areas. Instead of a dense receptor mosaic, suppose that an animal has a lens system but only one photoreceptor capable of generating an electrical signal. This is the case for the arthropod *Copilia* (Figure 3.7), in which mechanical scanning of the receptor (recall that the *Anableps microlepis* moves the lens) is seemingly used to ensure that the anterior lens focuses on this transduction element. Finally, consider the rattlesnake, which has two sources of imagery, one in the visible and the other in the infrared range. There is evidence that these are integrated to give the snake a composite view of its environment [15].



(*a*)

(*b*)

(*c*)

**Figure 3.6** (Caption on page 66.)

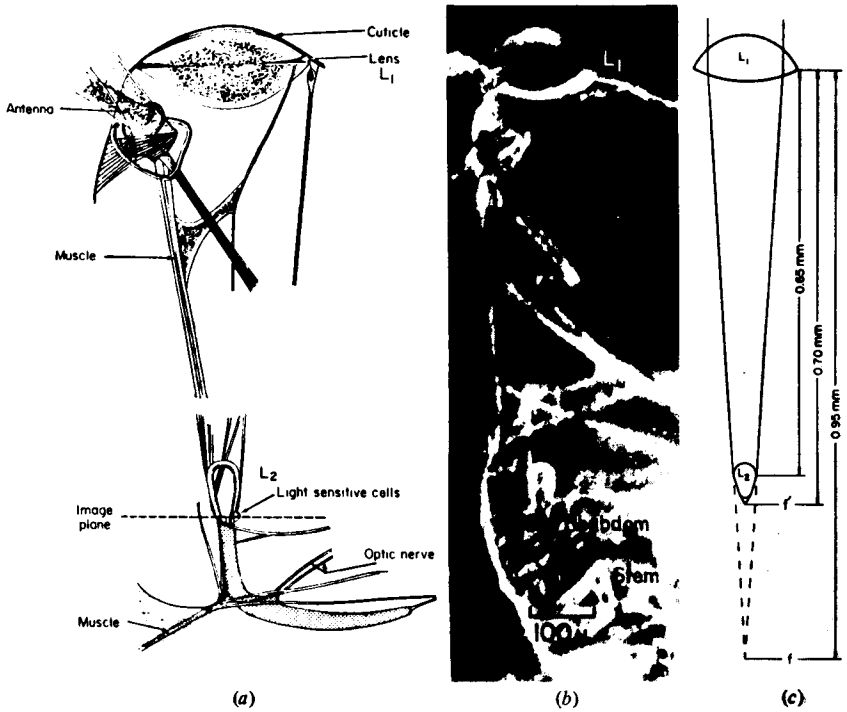



**Figure 3.7** *Copilia quadrata.* (*a*) The scanning eye of the copepod, *Copilia quadrata*, which possesses a fixed anterior lens ($L_1$) as well as a smaller, movable posterior lens ($L_2$) attached to a single photoreceptor within its body. This interior lens and receptor assembly continuously moves across the image plane of the anterior lens, possibly in analogy to the mechanical scanning microdensitometer discussed in Chapter 2. (*b*) Dark-field photomicrograph of the eye. (*c*) The optical system showing the positions of the corneal anterior lens ($L_1$) and the crystalline-cone posterior lens ($L_2$); *f*: focal point of the corneal (anterior lens); *f'*: focal point of the total optical system. [*From J. J. Wolken, "Comparative Structure of Invertebrate Photoreceptor," in H. Davson (ed.), "The Eye," vol. 6, "Comparative Physiology," Academic, New York, 1974, pp. 111–154, after H. Grenacher, "Untersuchungen über das Sehorgan der Arthropoden, insbesondere der Spinnen, Insekten and Crustaceen," p. 145, Vanderhoeck and Ruprecht, Göttingen, Germany, 1879.*]

**Figure 3.6** Photographs of the head of the dragonfly *Orthetrum*, showing two foveas, one pointing directly ahead and the other forward and upward, for catching prey in flight. (*a*) Camera at 35° to the longitudinal axis of the animal, forward and upward. (*b*) Camera at 15° to the longitudinal axis, forward and upward, with minimum size of pseudopupil between the two foveas. (*c*) Camera on the horizontal axis, looking straight forward. The black area of facets is the pseudopupil. The larger the pseudopupil, the greater the density of the visual units (ommatidia) looking in that direction. The white dirt on the eye serves to mark the facets. (*Photographs provided by G. A. Horridge, Australian National University, Canberra.*)





Let us now consider each of the different stages in Figure 3.3. The "cornea" (which is covered with a film of tears) is the front surface of the eye which bends the light to form the image. It is transparent and supported by an opaque layer of fibrous membrane called the "sclera" (see Figure 3.2), part of which is seen as the white of the eye. Although the air-cornea-aqueous humor (contained in the anterior chamber) pathway is responsible for approximately two-thirds of the optical power of the eye (42 diopters compared with 57 to 62 total), it is not a very good optical instrument.

After the light emerges from the aqueous humor, it passes through the pupil, which is a diaphragm or opening in the center of the iris [11]. The characteristic pigmentation of the iris is what gives us the color of our eyes. Curiously, women usually possess larger pupils than men and blue-eyed persons have larger pupils than those with brown eyes. This circular hole is similar to an aperture stop for a lens in a photographic camera [18]. Thus the iris behaves in a certain sense just like a servomechanism by contracting and expanding the size of the pupil. In this fashion it controls the amount of light which passes on to the next stage, which is the crystalline lens. Figure 3.8*a* is a schematic frontal view of the human eye. Variations in the diameter of the diaphragm are achieved by the contraction of two kinds of smooth muscle fibers, the sphincter and the dilator pupillae. The sphincter is responsible for constriction and runs parallel to the circular iris; the dilator activates expansion and functions radially. In cooperation these muscles can achieve a 16-fold

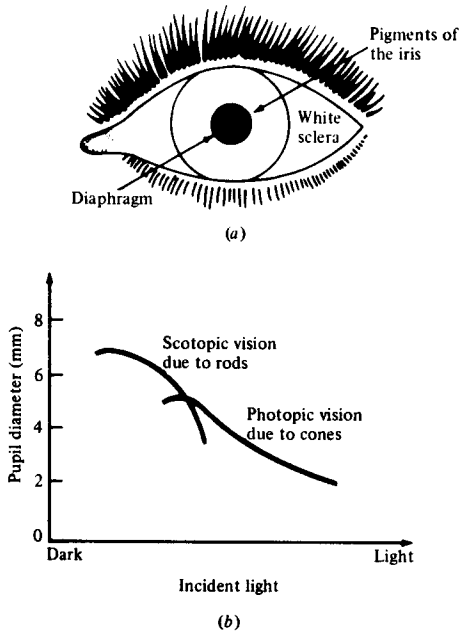



**Figure 3.8** The two modes of operation of the pupil. Note that the incident light is only one of many factors that control the diameter of the pupil at any time. (*a*) Schematic frontal view of the human eye. (*b*) Sketch of the variation of pupil size as a function of incident light.

change in aperture area by varying the diameter from about 2 mm to a maximum of 8 mm. Control of the size of this area is such that it is normally kept as small as possible in order to maximize the focus. The time constant associated with this action is about 10 s for the full diameter range. The activation of the control is extremely complex and is due to numerous factors but mainly to the amount of incident light and the emotional state of the viewer. Indeed, a recent survey has listed 23 different sources of variation, including sexual preference, political attitude, fatigue, semantic stimuli, and signal wavelength [32]. The dynamic characteristics of this pupillary light reflex system also appear to be related to certain neurological disorders. Readers who are interested in models of this dynamic control system should consult [21, 22, 24, 25, 33]. Pupil size as a function of the amount of incident light is sketched in Figure 3.8*b*.

Varying the size of the pupil of the eye serves three objectives [8]. The first effect is the light reflex function discussed above, which controls the amount of light that enters the eye and therefore impinges on the retina. The second, known as the near response, constricts the size of the pupil in order to control the depth of focus of near objects. The third, which is particularly important under bright light conditions, is reduction of the pupil size in order to reduce image aberrations. It appears that only the first two factors can be quantitatively controlled by external inputs.

After the pupil, the incident light passes through the crystalline lens, which is responsible for about one-third of the total optical power of the eye. It provides accommodation to near and far vision by changing its geometry, a function which is achieved in a camera by moving the lens. This second lens in the light pathway is made up of nonrigid laminae, much like an onion, and control is achieved by the action of the ciliary muscles, which vary the laminar thickness and shape. The control response time is approximately 0.4 s. Thus this lens guarantees that the image is in focus at the retinal plane, where an inverted image is provided.

Finally, we have the fluid, or vitreous humor as it is called, which is gelatinous and is essentially the means by which the shape of the eye is maintained. Light passes through the fluid to the retina, which is responsible for the electrooptical conversion between the incident light patterns and the resulting first stage of electrical activity in the nerve cells.

We may view the optical system in Figure 3.3 as an input/output transformation of a three-dimensional space whose range is restricted by the physical capacity of the human eye. We observe this space through a two-dimensional "window." The input light patterns result in an output falling on the retina which originates from the three-dimensional scene viewed by the eye and which contains objects to be recognized. The height of the resulting inverted image in this retinal plane is related to the size of the object subtended (see Figure 3.9). Let $S$ be the size of the object at a distance $d$ from the eye and $P$ the size of the projection on the retina. Then if we assume that the focal length of the eye is about 17 mm,



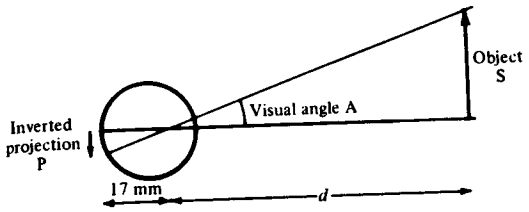



**Figure 3.9** The retinal projection of an object viewed by the human eye.

$$P = \frac{17S}{d} \quad \text{mm} \tag{3.1}$$

and the visual angle $A$ is given by

$$A = \tan^{-1}\left(\frac{S}{d}\right) \quad \text{degrees} \tag{3.2}$$

For example we can compute that a thumbnail (= 1.5 cm) at arm's length (= 60 cm) will subtend a visual angle of about 1.5°. The most sensitive part of the retina, the "fovea," subtends an angle of only about 2°.

The breadth of the scene is governed by the limit of the eyes' peripheral vision. Experiments have shown that retinal stimulation occurs for bright point

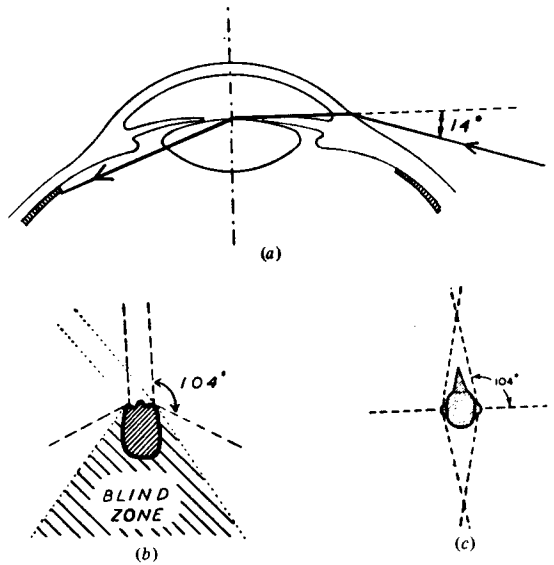



**Figure 3.10** Peripheral vision. (*a*) A schematic showing the limits of the peripheral vision of man. Light originating from behind the head will project onto the retina. (*b*) The blind zone of man when physical movement of the eye is taken into account. (*c*) The peripheral vision of a bird is much more extensive than that of man because of the placement of its two eyes. (*From M. H. Pirenne, "Vision and the Eye," 2d ed., Associated Book Publishers, London, 1967, p. 20.*)

sources of light that extend to about 104° from the optical axis, as shown in Figure 3.10*a*. We note that even objects lying behind the viewer are able to affect the retinal projection. Taking into account the ability of the eyeball to move horizontally as well, Figure 3.10*b* shows what portion of the circle centered at the human is a blind zone. Although this small angle is indeed impressive, it does not compare with the peripheral vision of birds, which have their eyes mounted on the side of the head, as shown in Figure 3.10*c*. No doubt this developed as a result of the distinctive requirements of life in the bird's habitat.

We see, therefore, that the geometry of the ocular optic system limits to a certain extent the ability of the eye to view the three-dimensional world around it. There are two additional important limitations which arise in the dioptric system. First, optical aberrations due to diffraction tend to limit the spatial frequency response, acting as a low-pass filter to image patterns at the input. We shall consider this concept in greater detail in later chapters but for the moment it is simplest to visualize spatial frequency as the number of alternating black and white bars of equal width per unit length of a given visual stimulus. The higher the frequency, the greater the number of oscillations; above a certain frequency the ocular system attenuates the signal reaching the

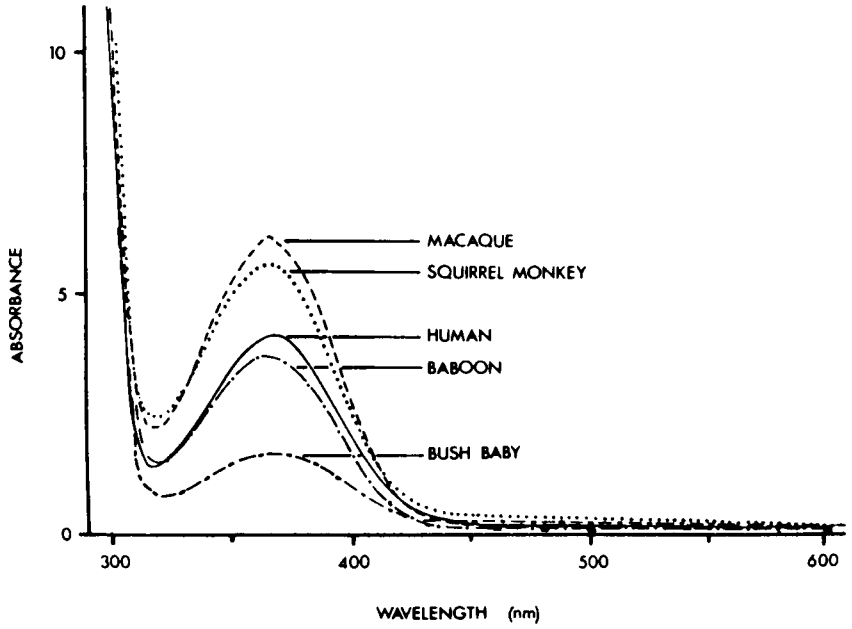



**Figure 3.11** The absorbance spectra of various primate lenses. (*Adapted from G. F. Cooper and J. G. Robson, "The Yellow Colour of the Lens of Man and Other Primates," Journal of Physiology, vol. 203, 1969, pp. 411–417.*)



retina much in the same way as does any optical instrument. A mathematical treatment of this phenomenon is given in [17, pp. 265–272] where the characteristics of the retinal image with respect to these optical factors is detailed. The second aberration is chromatic in nature in that the refractive indices of the various ocular media are dependent on the wavelengths constituting the input signal. For example, Figure 3.11 shows the absorbance spectra of the lenses of various primates. It is interesting that these spectra also change considerably with age (see [17, p. 246]).

In this section we have considered as a black-box transformation the optical preprocessing that takes place in the human eye. The input is the three-dimensional visual space and the output a corresponding inverted two-dimensional retinal projection. We note that even this potentially simple operation is quite complicated when it occurs in nature, as opposed to being man-made, as was the case for the computer vision systems discussed in Chapter 2. Generally the farther we are in the human vision system from the stimulus input, the more involved and unknown are the operations.

## 3.3 ELECTROOPTICAL RECEPTORS

A striking feature of the primate retina is that it consists of essentially five distinct elements. These are situated at the rear of the eyeball and are structured in vertical layers, the whole assembly having approximately the thickness of a sheet of paper. These processing elements, shown schematically in Figure 3.12, are: (1) rod and cone photoreceptors; (2) horizontal cells; (3) bipolar cells; (4) amacrine cells; and (5) ganglion cells. An actual cross section of a human retina is shown in Figure 3.13. The input light must first pass through the relatively transparent optic nerve fibers, the blood vessels, and the different layers of cells before reaching the only existing visual transducers, which are situated at the outermost extent of the outer plexiform layer. This obviously attenuates the arriving signal, a curious inversion by nature! Furthermore the rod-and-cone transducers are so arranged that their light-sensitive surfaces actually point away from the incoming image. The output electrical signals from the retina (see Figure 3.1) are transmitted by the axons of the ganglion cells, which together form the optic nerve. Electrooptical transduction is accomplished by the rods and cones and these will be discussed in this section. The next four layers of cells are actually involved in image signal processing and their anatomical arrangement will be covered in the next section. We note that these would normally be considered as a functional extension of the brain, or the "human vision computer."

Since about the early 1960s and especially during the 1970s, the electron microscope has contributed a significant amount of knowledge regarding the ultramicroanatomy of the rods and cones as well as the specific synaptic interconnections in the retina. These are extremely complex, and considerable uncertainty still remains regarding many facets of the retinal structure. Around

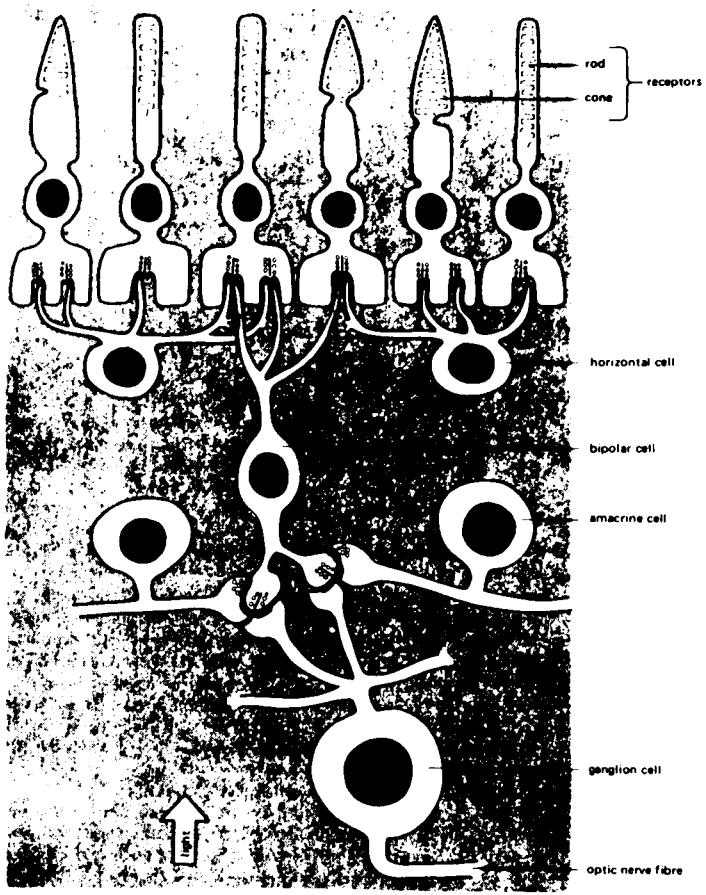



**Figure 3.12** A highly schematic diagram of the human retina. [*From C. Blakemore, The Baffled Brain, in R. L. Gregory and E. H. Gombrich (eds.), "Illusions in Nature and Art," Duckworth, London, 1973, pp. 9–48.*]

the turn of this century important contributions to our knowledge of the vertebrate retina were made by Cajal [5, 6] (see [17] pp. 770–904 for a translation of [5]), who suggested that a commonalty of structure existed for all vertebrates. More recent discussions can be found in [3, 4, 7, 10, 14].

We can see from Figure 3.2 that the retina upon which our visual world is projected is concave and surrounds nearly 200° of the eye. One small portion, where the optic nerve leaves the eye, is a blind spot at which no photoreceptors exist at all and which is therefore insensitive to light. In the human eye any light which fails to be absorbed by the receptors is then absorbed by the retinal epithelium and choroid layer, which tend to minimize the effect of stray light.

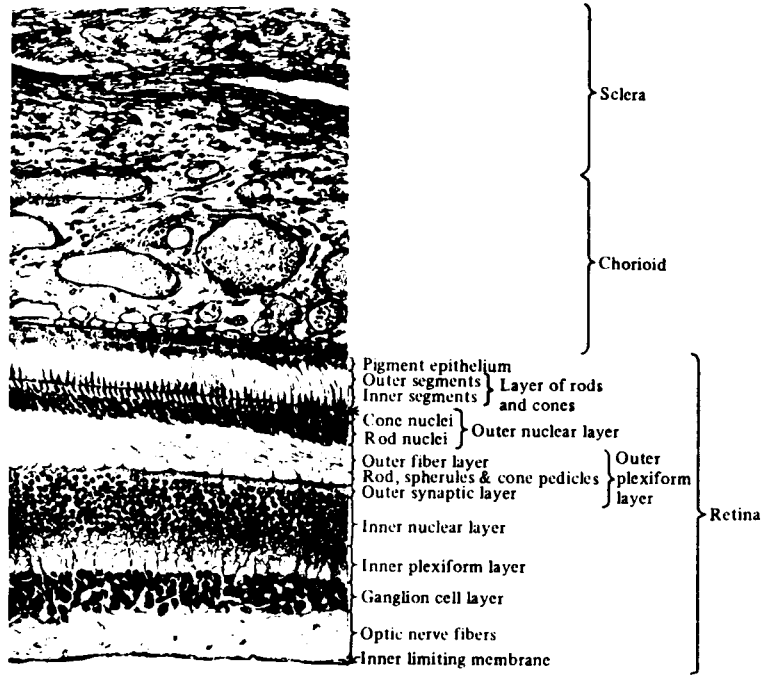



**Figure 3.13** Cross section through the human retina in the region of the central area at a medium magnification. The light entering the eye from the outside through the pupillary aperture—from below in the figure—passes through all retinal layers until it reaches the bacillary layer of the rods and cones, where it elicits coded electrical signals. These, in turn, pass in the opposite direction, from the rods and cones to the ganglion cells, along whose fibers they are transmitted to the brain. (*From S. L. Polyak, "The Retina," University of Chicago Press, Chicago, 1941.*)

There are also large differences between the sensitivity of the retina at its center and periphery because the distribution of the transducing elements, the rods and cones, is not uniform.

The "fovea" of the retina (see Figure 3.2) defines the visual axis of the eye and is responsible for highly detailed and exact vision. There are no blood vessels covering this area to interfere with the impinging image, and the correspondence of the cones with the next levels of cells is one-to-one. Nevertheless, the fovea is quite small, consisting of about a 1.5-mm-diameter depression (corresponding to 5.2° of the visual angle) in the retina situated near the optic axis. Figure 3.14 shows a cross section of the human retina at the fovea. Its center contains only cone receptors and no rods, and the cones here are much longer and thinner than those on the periphery. This rod free area, responsible for central vision, is about 0.3 mm in diameter, corresponding to only 1° of visual angle and only 0.5 percent of the total extent of photoreceptor coverage. This turns out to be twice the visual angle subtended by the sun or

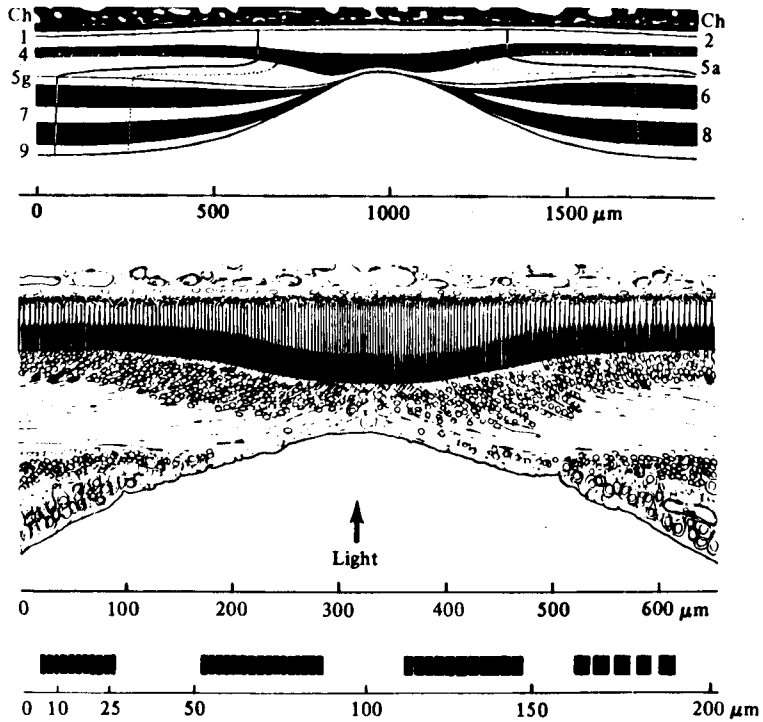



**Figure 3.14** The central fovea of the adult human eye, whose diameter is approximately 1.5 mm (5.2°). The upper sketch shows, semidiagrammatically, changes in the relative thickness and position of the retinal layers. It also shows the relationships of the photoreceptor layer (2) and the deeper layers (4 to 9) caused by the latters' displacement owing to the formation of the fovea. The three black layers indicate the outer nuclear (4), inner nuclear (6), and ganglion cell (8) bodies. White dots in layer (4) represent the rod nuclei; Ch refers to the choroid membrane. The broken lines encompass the rodless territory and the portion of the foveal pit functionally related to it. The solid lines mark the region (2) within which the inner and outer segments of the cones are observed to be very thin and long. The middle drawing represents the foveal center filled with its thin, elongated cones. The most centrally located rods correspond with the most central rod nuclei in the outer nuclear layer (4). Note the practical disappearance of the remaining inner layers in the foveal center. The lower sketch represents samples from four localities showing relative size and number of cones (inner segments), beginning from the left: center of the fovea, slope of the same, edge of the same, and periphery of the central area. Upper sketch reproduced at 80×, middle at 250×, lower at 700× magnification. (*Adapted from S. L. Polyak, "The Retina," University of Chicago Press, Chicago, 1941.*)

the moon. Lateral vision is governed by the loosely defined concentric regions surrounding the fovea centralis, which depend on the density of the cones in the region. The "parafovea" is defined as being 2.5 mm in diameter (8.6° of the visual angle) and this area already contains more rods than cones. Figure 3.15*a* shows the density of rods and cones in the parafovea. The next concentric region, the "perifovea," is defined by the annulus having an inner diameter of



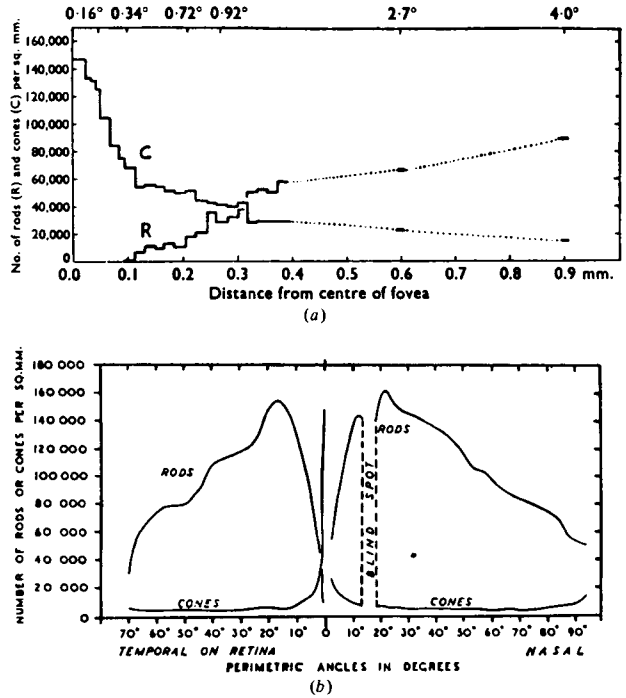



**Figure 3.15** The spatial distribution of rods and cones. (*a*) The distribution of rods and cones centered around the visual axis, which is at the center of the fovea. (*b*) The complete density distribution for rods and cones in the human retina. (*From M. H. Pirenne, "Vision and the Eye," 2d ed., Associated Book Publishers, London, 1967.*)

2.5 mm and an outer diameter of about 5.5 mm (19° of the visual angle). Then, beyond this region is the "peripheral retina," constituting about 97.25 percent of the retinal concave surface and consisting largely of rods. Figure 3.15*b* is an overview of the spatial distribution of rods and cones for the complete human retina. We note that there are about an order of magnitude more rods than cones in the human retina: about $120 \times 10^6$ rods compared with $6.5 \times 10^6$ cones in each eye. We are not really aware of this distribution of photoreceptors and the ensuing loss of acuity away from the fovea because the latter is usually centered on the image we are observing.

As we observed in Chapter 1, it is most common in picture processing by machine to specify a rectangular tesselation of the image plane. Thus, each pixel is taken to have a square shape, although triangular and hexagonal pixel shapes have also been suggested. The idea is based on the observation that retinal receptors are distributed in a hexagonal array of cones with the intermediate space filled by the smaller rods. This tesselation is optimal in the sense that each element has a maximum number of equidistant neighbors. Such

a scheme turns out to be impractical for implementation on a general-purpose computer. Another observation is that the sampling frequency in the retina is greater near the fovea and falls off towards the periphery. Such a so-called foveated array has been represented mathematically by using logarithmic spiral grids [34]. A short review of data structures for computer images is presented in [31].

This unique distribution of photoreceptors is responsible for the "duplicity theory" of the retina, in which two kinds of vision are distinguished in man. Of course their underlying processes constantly interact; for example, we are aware of their functioning when we initially enter a darkened cinema. First consider photopic vision, which describes the activity of the cones. These are responsible for day vision and the exacting discriminative power of the retina, providing a high degree of acuity. Color processing is also an important function of these transducing units. The second aspect, scotopic vision, is provided by the rods. It is concerned with night viewing and therefore tends to integrate the input light in order to allow for increased sensitivity under these generally more trying conditions. The rods are thought to be achromatic. Thus, the fine mosaic of the cones in the fovea adapts quickly to bright light and color, while the rods are relatively slow in response, are coarsely distributed, and only adapt to the shades of gray of dim light. For both rods and cones, the coded output signal associated with a particular input light pattern is a sequence of frequency-modulated pulses. The actual process of transduction is still not completely understood.

Figure 3.16 is a drawing of a rod and cone from a human eye. Although vertebrate receptors vary in size and shape, their basic organization is quite similar. The outer part contains a photosensitive material, while the inner part forms the contact with other cells. For the retina to be capable of detecting the incoming light patterns, it must contain a photosensitive light-absorbing material. This so-called visual pigment is different for rods and cones.

The outer segment of the rod contains a substance referred to as "rhodopsin" or "visual purple," which has been known for over a century and has been studied extensively. It appears pink in color and is bleached or made white by light. There also exists a close correspondence between the spectral response curves of human rhodopsin, which can be readily extracted and studied in vitro, and human rods. It is therefore quite reasonable to attribute the behavior of the latter largely to the properties of the former. Figure 3.17 shows an example of the spectral sensitivity curve for human rhodopsin. Essentially, therefore, the rods behave like bandpass filters, basically operating within the visible spectrum between approximately 400 and 700 nm, as shown in Figure 2.2.

In distinction to the rods, the cones possess three different kinds of photopigments, each of which responds differently to a stimulus wavelength, as shown in Figure 3.17. Cone pigments are difficult to extract, and have been studied mainly in situ by reflection densitometry [19]. From these experiments we know that the spectral absorption functions for different species are similar in general shape. We shall see in Chapter 7 that these three different kinds of



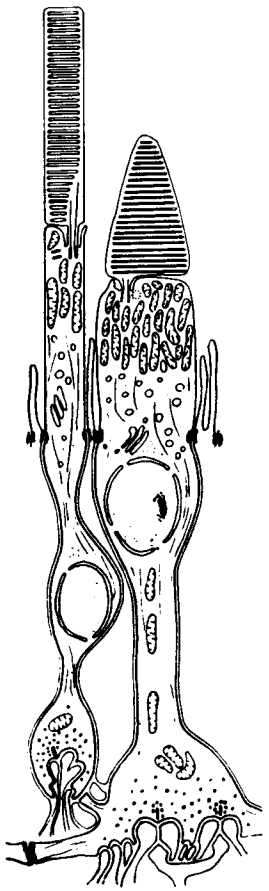


**Figure 3.16** Human rod (left) and cone (right). (*From L. Missotten, "The Ultrastructure of the Human Retina," Editions Arscia Uitgaven, N.V., Brussels, 1965.*)

cones with their different bandpass characteristics are responsible for three independent information channels, which then characterize color vision. It is not clear whether the cones in primates interact, although it is possible that they do at the ganglion cell level and most probably in the lateral geniculate nucleus (see Figure 3.18). There is evidence that neighboring cones in the retina of the turtle do interact [2], so that some questioning of the premise of independence for humans may be warranted. In the case of vertebrates, it appears that this electrical coupling tends to reduce photoreceptor noise in the presence of low light levels. It is also interesting that spatial visual acuity is not necessarily degraded and photoreceptor communication might even enhance image processing at high light levels [12].

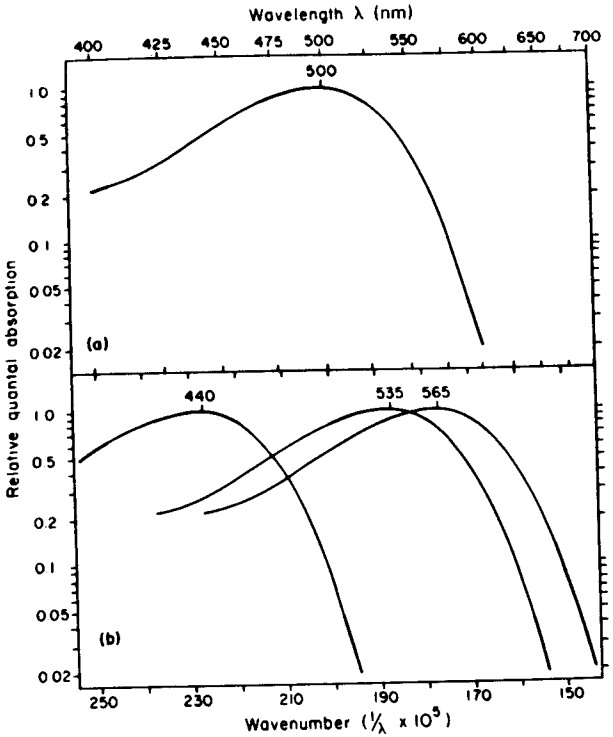



**Figure 3.17** A representative set of absorption spectra for rhodopsin, the rod photopigment (*a*), and the cone photopigments (*b*). [*From I. Abramov and J. Gordon, "Vision," in E. C. Carterette and M. P. Friedman (eds.), "Handbook of Perception: Biology of Perceptual Systems," vol. 3, Academic, New York, 1974, pp. 327–406.*]

In one particular location of the retina there is some additional spectral processing. Observe in Figure 3.2 that a region of about 6° to 10° centered on the fovea has been labeled "macula lutea." This is a yellow screening pigment, which has only been found in primates and which tends to filter out light in the blue-violet portion of the spectrum. This filtering effect, in addition to the filtering action at short wavelengths of the lens in the optical system (see Figure 3.11) and the absence in the central ($\frac{1}{2}^{\circ}$) region of the fovea of cones with pigments sensitive to blue, significantly attenuates the contribution to human vision at these high frequencies.

Summarizing this section, we may say that the rod and cone photoreceptors in the retina are the transducing elements of the human visual system. They transform the focused image on the retina into an electrical energy signal according to the principles to be described in Section 3.4.



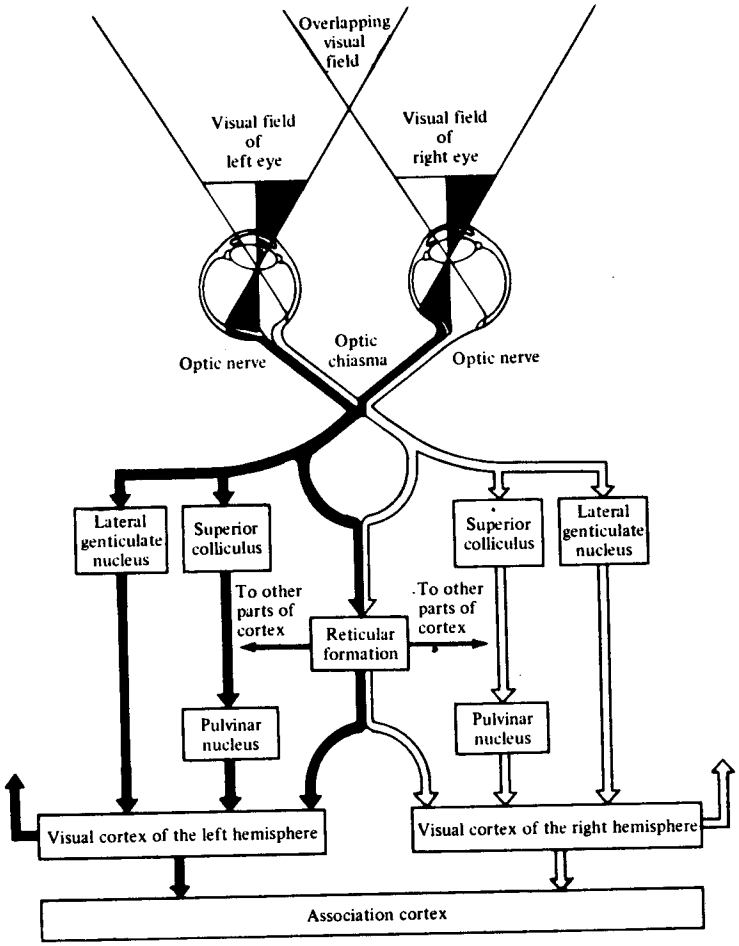



**Figure 3.18** Block diagram showing the major known interconnections in the visual pathway.

## 3.4 THE VISUAL PATHWAY

In Section 3.3 we referred to a stack of five layers of cells in the retina: the photoreceptors responsible for image transduction and the bipolar, horizontal, amacrine, and ganglion cells, which are already involved in some low-level processing. This section will attempt to delineate the major visual pathways involving these cells, as well as their connections. In doing so it will become clear that visual processing in living organisms is constituted as a hierarchy of analyses. It goes without saying that this will be a necessarily incomplete

presentation, first, because the topic is very complex and would require a comprehensive treatment to do it justice and second, because our knowledge at this point of many of the details of even the major interconnections in the brain is severely limited. The literature is often conflicting in nature and is based on a relatively small sample of cells.

Figure 3.18 is a simplified diagram of the visual pathways of the primate visual system. We note that the axons of the retinal ganglion cells form the optic nerve, and that these fibers of each eye are divided into two groups according to which part of the retina they project from. The axons meet just before they reach the brain at the optic chiasma, after which a large proportion of them pass on to the lateral geniculate nucleus. The latter is arranged in distinct horizontal layers and its function is not well understood. We see in the diagram that a number of fibers also project to the superior colliculus and the reticular formation. From the lateral geniculate nucleus, where some processing of the image information does occur, the fibers project to one or the other of the two hemispheres of the occipital cortex. Signals from the reticular formation go to other parts of the cortex and are thought to be responsible for sensory-motor control [30]. Not much is known about the detailed functional organization of the brain, which contains many distinct horizontal layers of cells of various sizes and shapes. However in Chapter 8 we shall discuss some very interesting experiments which suggest the existence of a data organization which could support form recognition. Nevertheless the organization of the brain, or even of the sections of the brain responsible for vision, is nearly a complete unknown from the point of view of information processing. In the following chapters we shall present what evidence exists for image processing in the different stages of the visual pathway.

Let us now examine in more detail, enough to comprehend the complexity of the organization, the elements of the anatomy of the visual pathway. Figure 3.19 shows in a more or less realistic fashion the layers of cells in the retina just before they exit from the eye. This figure should be compared with the more schematic version shown in Figure 3.12, which is greatly simplified and represents our knowledge until about the early 1960s. Indeed the situation is even more complicated than in Figure 3.19, since many subcategories of cells have since been discovered, although their exact nature and function are often not clear. Basically two kinds of connections exist, those that are vertical and carry information from the photoreceptor to the brain and those that are horizontal and provide for lateral interaction among cells. The following observations about these connections are pertinent to experiments made on human and monkey retinas.

After the light striking the eye passes through the different strata of cells and membranes and is transduced by the layer of rods and cones, all subsequent stages of processing are electrical in nature. At the first level are the horizontal cells, probably of two kinds, one large and the other small. These cells interconnect and mediate between the rods and cones by means of the outer synaptic layer of the nuclei of these receptors. The connections are simpler



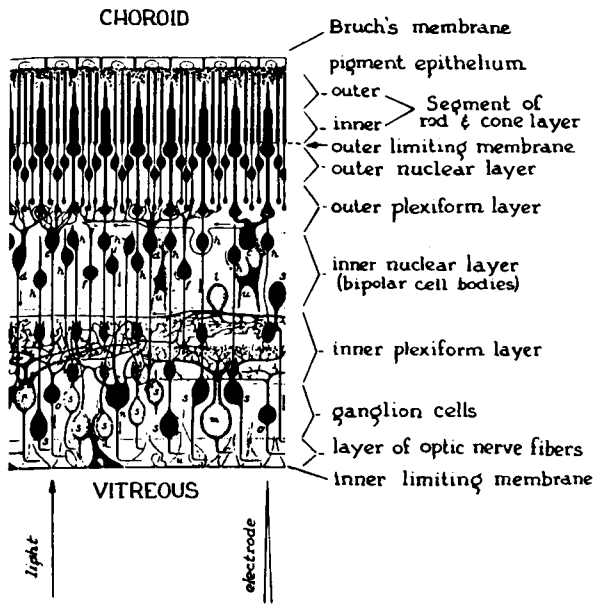



**Figure 3.19** A sketched representation of an actual retinal cross section. Figure 3.12 is an abstract version of this diagram and Figure 3.13 is a view at a lower magnification. Here it is about 400. [*From T. C. Ruch, "Vision and the Retina," in T. C. Ruch and H. D. Patton (eds.), "Physiology and Biophysics," vol. 1, "The Brain and Neural Function," 20th ed., 1979, Saunders, Philadelphia, pp. 461–513, from S. L. Polyak, "The Retina," University of Chicago Press, Chicago, 1941.*]

than those that appear at the input to the next level of cells, the bipolars, thereby implying a lesser degree of sophistication in the processing. There are four kinds of bipolars: rod bipolar, invaginating midget bipolar, flat midget bipolar, and flat bipolar. The rods synapse (connect) with the rod bipolars, while the cones synapse with the others. Each receptor projects to at least one bipolar cell. In the fovea, usually one cone is connected to one bipolar, which is then, in turn, usually linked to only one ganglion cell. Outside the central region of the fovea, many rods or cones communicate with a single bipolar, to the point, for example, that beyond 20° from the visual axis hundreds of cones converge on an individual bipolar. Obviously some process of spatial integration related to visual resolution is occurring, with inputs being provided by both photoreceptors and horizontal cells. It has been hypothesized on the basis of experiments with catfish that the horizontal cells produce an integrated signal, which is used to filter out the low spatial frequency components in the image [23]. The bipolar output signal paths or, equivalently, the axons of these cells synapse with amacrine and ganglion cells. The role of the amacrine cells is to horizontally interconnect and modify the signals at the junction of bipolar and ganglion cells. They may be involved in such visual processes as inhibition,

a phenomenon dealing with contrast enhancement that will be discussed in Chapter 6.

The ganglion cells are connected to bipolar and amacrine cells and also synapse with each other. Although the fine details of this anatomy are not well known, in later chapters we shall describe some very interesting experimental results dealing with image processing by ganglion cells in animals. Apparently, there are three kinds of ganglion cells, referred to as W-, X-, and Y-type cells [26]. They are thought to serve different functional roles. At the output from the retina the axons of these cells form the optic nerve, a bundle about the thickness of a pencil, which contains only about 1 million fibers. This is an extremely small number considering the tens of millions of rods and cones in the retina and the fact that this represents the entirety of the visual input to the brain. As noted previously (see Figures 3.2 and 3.15*b*), in the location where this optic nerve leaves the visual field, about 16° nasally from the optic axis, there is a blind spot (scotoma) with horizontal and vertical dimensions of about 5° to 6° and 7° to 8°, respectively. As can be seen from Figure 3.18, the input from a particular optic tract to a particular brain hemisphere contains only half of the visual field. Because of the action of the visual optical system, the left half of each visual field projects onto the right half of the retina; for the left eye this is on the so-called nasal retina and for the right eye on the temporal retina. These come together at the optical chiasma, where they collect, to proceed to the right hemisphere of the brain. Similarly, the right part of the visual field projects to the left hemisphere. Since the visual fields of the two hemispheres overlap in front of the viewer, each hemisphere possesses data about this common projection. These binocular views are combined by the brain to provide us with the ability of stereoscopic depth perception [9, 30]. At the periphery, however, there is no overlap and only monocular vision is possible.

In man about 20 to 30 percent of the fibers in the optic nerve connect to the superior colliculus, while in lower vertebrates and birds most nerve fibers actually terminate here. Figure 3.20 shows a comparison of the visual pathway for two different species. For example, the frog does not have any cortex and its visual pathway ceases in the so-called optic tectum. Nevertheless the frog is capable of some elementary but interesting image processing operations, which we shall discuss in later chapters. Some animals, such as squirrels, are somewhat intermediate in that they do possess a small visual cortex. Thus we may generalize by pointing out the two categories, one having a relatively large visual cortex such as man, monkey, or cat; the other a rudimentary or nonexistent cortex such as frog, rabbit, and squirrel. In the first case it appears that most processes of vision are carried out in the brain, while in the second case the retinal ganglion cells are capable of some rather important vision computations.

In man it appears that the superior colliculus is responsible for controlling eye movements [20, 35]. Upon leaving this area, most of the fibers then connect to the pulvinar nuclei of the thalamus, after which they pass on to the occipital cortex, as shown in Figure 3.18. The pulvinar nuclei mediate the pupillary



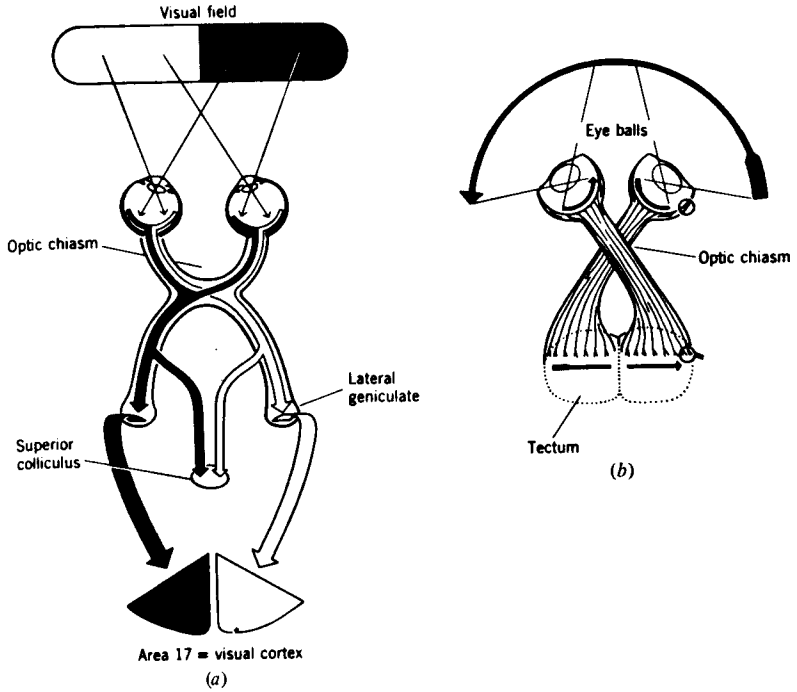



**Figure 3.20** A comparative view of the visual systems of a human and a frog. In the human visual system (*a*) we see the predominance of messages going via the lateral geniculate nucleus of the thalamus up to the cortex with relatively little going to the superior colliculus, while in the frog (*b*) it is the connections to the tectum (which is the amphibian analog of the mammalian superior colliculus) that predominate. Note also the splitting of the two halves of the visual field, so noticeable at the human optic chiasma yet absent in the frog. (*Adapted from M. A. Arbib, "The Metaphorical Brain," Wiley, New York, 1972.*)

reflexes discussed in Section 3.3, thereby providing some feedback from the higher to the lower vision processes. The thalamus, which also contains the lateral geniculate body, is a portion of the brain where all the sensory signals except those for olfaction congregate before passing on. We could imagine some degree of low-level interaction at this point which would integrate the various environmental data that the body senses.

The lateral geniculate body of the thalamus is a major pathway for man in that a significant fraction of the optic fibers connect here, as do fibers from other parts of the central nervous system. It is a laminated structure containing six distinct layers of cells. At present no evidence exists to imply that any major visual analysis occurs at this point, although there is strong evidence that the cells here are implicated in color vision (see Chapter 7). Signals from more than one layer of the lateral geniculate body project primarily to the visual cortex. Another area whose function is unclear and which receives input from the optic

tract is the reticular formation. Signals from here pass on to other parts of the cortex.

In 1893 Ramon y Cajal was responsible for publishing the first comprehensive study of the anatomy of the human brain [6]. About 30 years earlier Golgi had established its organization, using a technique which allowed for the selective staining of small groups of neurons. By late in the nineteenth century it was known that the cortex is divided into different areas, each one responsible for a different function. The occipital area of the brain, where the sensory paths terminate, is called the "cerebral cortex." It contains about 70 percent of all the neurons in the human central nervous system, which attests to its importance. For example, the so-called area 17 of this occipital cortical lobe is referred to as the "primary visual" or "striate" cortex, the former term describing its function and the latter its appearance. As might be expected, the fovea of each retina occupies a disproportionately large projection on the visual cortex. It is probably also the case that inputs from both retinas first converge here on a single neuron cell. It is generally assumed that this physical location is concerned with the perception of light, color, and form, although the complexity of the analysis is completely uncertain. Surprisingly, the partial destruction of this area does leave man with some pattern recognition abilities intact. Experiments to relate visual stimuli with cortical events in vertebrates are difficult to perform and probably limited in maximum scope. After all, a monkey is quite incapable of accurately verbalizing its perceptions, a major link for man between the neurophysiology and psychology of vision.

The visual cortex projects back to the lateral geniculate body, whence, as we saw, it receives input, thereby providing top-down feedback in the hierarchy. Also, it is only the first stage of visual processing and thus is involved in further interconnections to the important association areas nearby, which in man represent the major part of the cortex. It is interesting to note that the visual, auditory and somatic areas of the cortex are together responsible for about one-quarter of its total size.

What physical characteristics does this important part of the human body, the cerebral cortex, have? It is a folded plate about 2 mm thick, which fits just inside the skull. The total area in man if unfolded would only be about $\frac{1}{7}$ m$^2$. It is densely packed with neurons, containing about $10^5$ neurons per square millimeter of surface, whereas the cortex as a whole consists of a network of about $10^{10}$ neurons. These cells are not randomly located but are arranged in layers, which alternate regions densely packed with cells with those that are sparsely populated. Within a plane in a particular region, a large degree of apparent uniformity may be observed. Processing in the cortex appears to be very local in the lateral direction, providing for the effect of a columnar structure radiating from the surface. This aspect will be discussed in Chapter 8, where the consequences for image processing will also be given.

Information from area 17, the primary visual cortex, projects to the adjacent prestriate cortex, areas 18 and 19, together referred to as the "association cortex." From here, the visual pathway projects to the "inferotemporal cortex," largely corresponding to areas 20 and 21. Areas 17, 18, and 19 are generally considered as the "visual cortex" and are sometimes referred to as visual areas I, II, and III, respectively. They seem to be mainly concerned with relatively low-level processing roles. However, lesions in the association areas definitely affect visual pattern discrimination capabilities. Thus, these areas serve the important function of associating memory with visual input patterns. In contrast, the inferotemporal cortex is significant to visual discrimination learning and is probably involved in higher-level visual processing. Perception obviously requires that all these centers function in concert.

The visual pathway is a massively interconnected and complex "vision computer" about which there is obviously much to discover. In this section we have briefly discussed from a macroscopic point of view the various subprocesses and their connections. Taking the computer analogy one step further, we might say that vision in the human is achieved by the interaction of many processors. The low level of processing is the one we know the most about, largely because of experimentation with animals. But what about the microscopic units, the individual "integrated circuit" computational elements which make up these processors? These are the neuron cells, which are discussed in the next section.

## 3.5 SIGNAL CODING AND PROCESSING

It has been mentioned previously that the stimulus signals are coded into sequences of frequency-modulated pulses. In fact the neuron, which is the elemental anatomic unit of the nervous system, operates by processing these pulse trains. It is well-known that frequency modulation offers a higher degree of noise rejection and stability than amplitude modulation, and it is interesting that the basic human computational circuit functions according to this design [27]. This section will describe in an introductory fashion the mechanisms involved.

Figure 3.21 is a schematic diagram of a single nerve cell, showing the "soma" or cell body surrounded by a thin plasma membrane, which is filled with cytoplasm and contains a nucleus. Such cells are usually roughly about 30 $\mu$m in diameter. Neural networks are configured by means of the "dendrites," which are the inputs to the cell, and the "axons," which constitute the single output. There are many inputs, typically perhaps 2000 for small cells to 16,000 for large cells, and this collection of dendrites for a particular cell is called a "dendritic tree." Note that some of the inputs are excitatory in that they promote the cell firing while others are inhibitory and retard it. The dendrites are perhaps 200 to 300 $\mu$m long and thus a cell may receive input from locations a considerable distance from it. On output, information flows from the cell via its axon termination to another by chemical means. This is achieved by the so-called chemical transmitter substance, which diffuses across a synaptic gap, as shown in Figure 3.21.

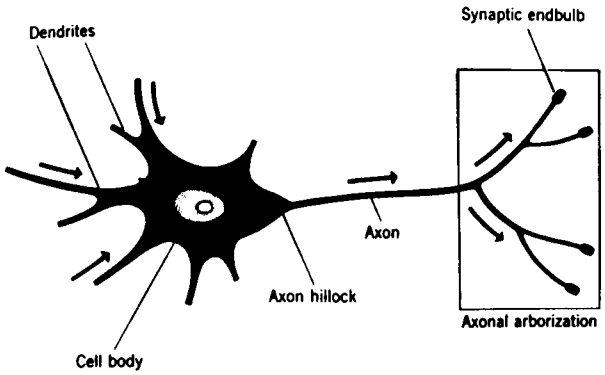



**Figure 3.21** Schematic view of a neuron. Activity from receptors and other neurons modifies membrane potentials on the dendrites and cell body. The effects of these changes converge upon the axon hillock whence—for appropriate spatiotemporal patterns of incoming activity—a pulse of membrane potential will be propagated along the axon, branching out into the axonal arborization to activate the synaptic end bulbs, which modify the membrane potential of other neurons or of muscle fibers in turn. In this way, axons from many cells serve as input connections to the dendrites of a particular cell. The output from the above cell appears on its axon. (*From M. A. Arbib, "The Metaphorical Brain," Wiley, New York, 1972.*)

In this section, we will describe how this chemical process results in a coded pulse train. The range of lengths for axons can be even larger than for dendrites, from about 50 $\mu$m to possibly even several meters. A single axon may be involved with perhaps hundreds of synaptic contacts. As we have seen in the previous sections, many cells are usually packed together in layers and the connections must be rather intricate. These nerve cells are embedded in a supporting protective network of glial cells, whose function is not yet completely understood. Figure 3.22 is a schematic diagram of the basic computational unit, while Figure 3.23 shows in a more or less realistic fashion how three neurons might be interconnected via their dendritic trees.

In addition to these basic elements, the human body contains specialized cells which are dedicated to sensing the environment and therefore act as energy transformation units. The rods and cones discussed previously are one such set of sensing elements, which convert light energy at the input to membrane potential changes at their output. Other sensory transduction systems are auditory (hearing), somesthetic (touch), olfactory (smell), and gustatory (taste). The

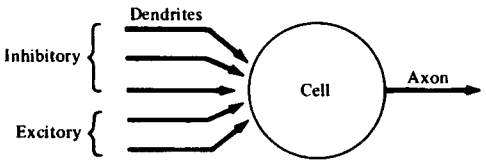



**Figure 3.22** An input/output representation for a typical neuron.



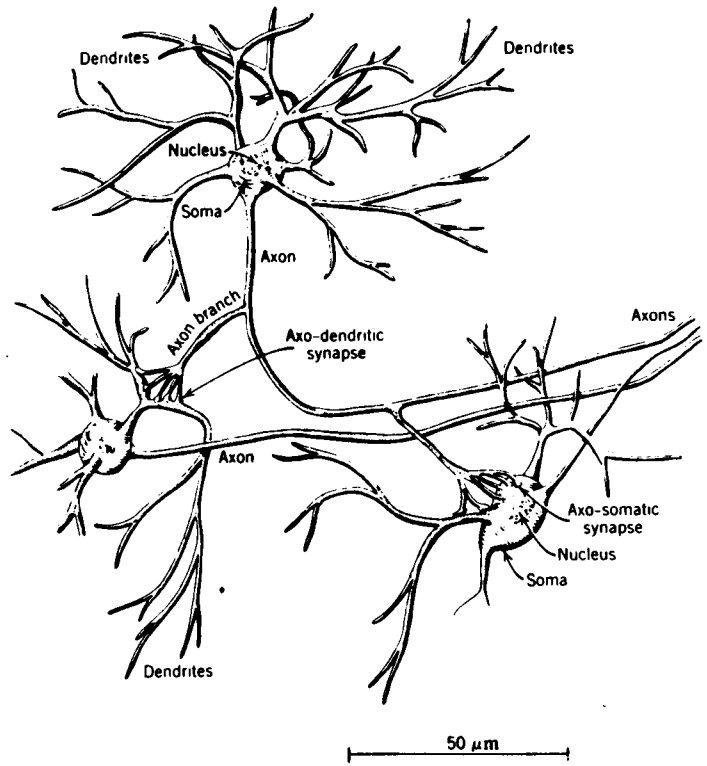
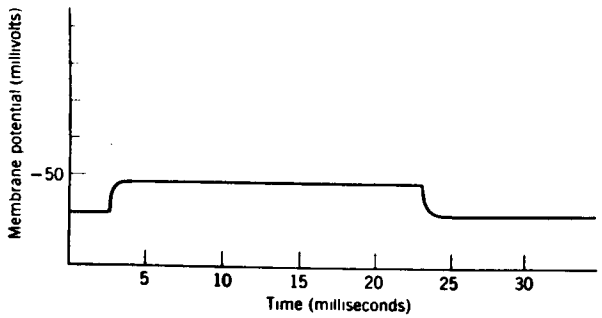


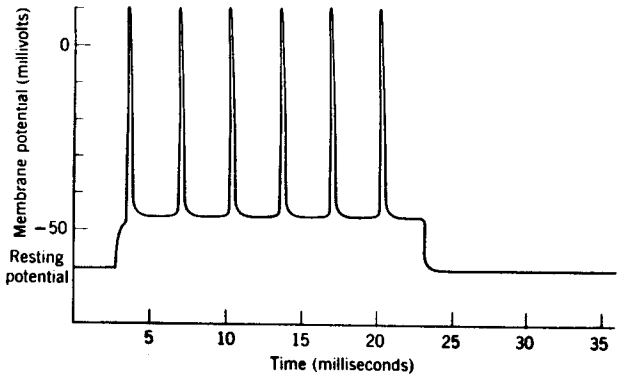

**Figure 3.23** A semischematic rendition of the interaction of three neurons. (*From C. F. Stevens, "Neurophysiology: A Primer," Wiley, New York, 1966, p. 2.*)

transducer units of the visual system are extremely sensitive and require only very small amounts of light to be activated; perhaps one photon is sufficient for the receptors of the human eye. Detailed knowledge of the photochemistry of the eye is available; however, the mechanism by which the transduced electrical signal is created is unfortunately not yet understood.

What follows is a simplification and generalization of how the process of frequency coding is achieved. The explanation is presented in stages, beginning with the neuron axon, which, as we have seen, is the output. Consider an electrophysiological experiment in which a microelectrode is inserted into the interior of an initial segment of axon near the cell body and a step voltage stimulus is applied. The resulting membrane potential or inside-outside voltage of the axon is then measured with a probe. It will be seen that a maintained voltage stimulus produces a frequency-coded pulse train whose frequency of oscillation is proportional to the stimulus voltage.




(*a*)




(*b*)

**Figure 3.24** When the amplitude of the depolarizing rectangular pulse stimulus is below threshold, the axon response is passive (*a*). If it goes above threshold, a sequence of spike action potentials superimposed on the generator potential is communicated along the neuron's axon (*b*). (*From C. F. Stevens, "Neurophysiology: A Primer," Wiley, New York, 1966, pp. 21–22.*)



"depolarization" and "hyperpolarization," as positive and negative excursions from the resting potential of the neuron, taken to be about −60 mV. If we apply a negative rectangular pulse to the axon, we observe a passive response consisting of a hyperpolarization; the resulting negative pulse is merely a slightly distorted version of the input. Similarly, for a positive pulse going below a certain threshold value we obtain a passive depolarization (see Figure 3.24*a*). On the other hand, if the amplitude exceeds the threshold, the probe records an active response, referred to as an "action potential" and shown in Figure 3.25. This voltage transient is initiated by the input going above the threshold and does not depend in any way on the pulse width. Therefore exceeding the stimulus threshold results in an action potential whose shape is substantially independent of the input and which travels along the axon at about 1 m/s, in a range of 0.1 to 10 m/s, the higher values attributed to the larger fibers. Of course distortions do occur over the length of this transmission line, but we will neglect these here.

Two aspects of the action potential are directly related to the frequency-coding property of the input: the "latency" and the "refractory period." The latency, or effective rise time, of the action potential is defined as the time between the application of the stimulus and the peak of the resulting output. This response time decreases exponentially as the stimulus intensity increases. The shape of the curve is similar for different axons but the time scale would tend to vary from cell to cell. The second aspect, the refractory period, is the minimum time between two successive stimuli which will evoke two consecutive action potential responses. This period might be measured in milliseconds. It turns out that the threshold for the second stimulus to fire the

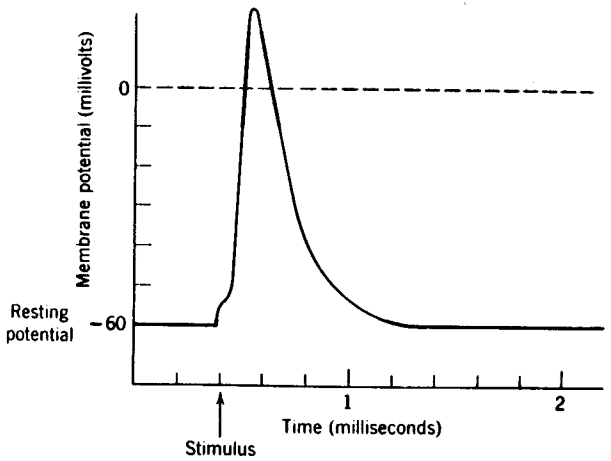



**Figure 3.25** An idealized version of an action potential. (*From C. F. Stevens, "Neurophysiology: A Primer," Wiley, New York, 1966, p. 14.*)

neuron is dependent on the refractory period. There is a dead-zone period, the so-called absolute refractory period, before which it is impossible to have another output pulse. After the dead zone has expired, the input amplitude threshold for the second pulse decreases exponentially as the refractory period increases. Thus, depending on the amplitude of the stimulus input, the threshold will decrease to a point where the neuron will be able to fire again. It again goes without saying that the characteristic of this relative refractory period is similar in shape for different axons and differs only in scale. From the above discussions, we see that if a constant voltage above threshold is applied, the latency and refractory period will both control the frequency of the output pulses. For example, a strong stimulus will yield a smaller refractory period and a faster rise time, thereby resulting in a higher frequency. An example of such an axonal response is shown in Figure 3.24*b*. By experimenting with a particular neuron, we may obtain a curve similar to that shown in Figure 3.26, a representative relationship between stimulus intensity and nerve impulse frequency. The time scales will differ for different axons. A typical output might have a 10 mV depolarization above threshold, resulting in 10 to 500 pulses per second.

It is interesting that some neurons do not behave in the manner described when subjected to a maintained input. In fact the axons of these cells accommodate to the unchanging input and the threshold remains high, so that no further impulses can occur. In order for this threshold to change and thereby fire the neurons, the input must either be increased or decreased. Thus we observe that this axon responds only to differential changes in the input. This type of cell is referred to as a "phasic neuron," in distinction to the previously described so-called tonic neuron.

Having described the existence of a process which converts axonal voltage potentials to a frequency-modulated pulse train, we shall examine the next stages in the process, all of which together permit cellular signal processing to

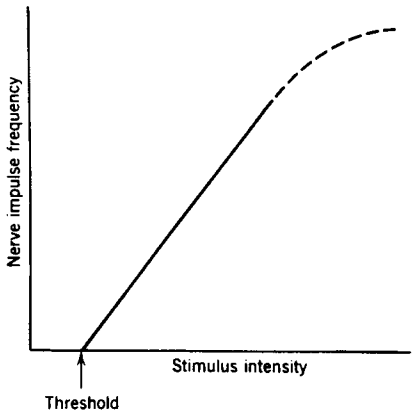



**Figure 3.26** The relationship between the nerve impulse frequency and the stimulus intensity. (*From C. F. Stevens, "Neurophysiology: A Primer," Wiley, New York, 1966, p. 24.*)



occur. The axon of one cell is connected to the dendritic inputs of other cells via a synaptic termination. This synapse is a chemical connection which employs a transmitter substance to convey information across its boundary. The action potential pulses conducted along the axon are converted by the synapse to a voltage in the dendrite, referred to as the "postsynaptic potential" (PSP). The PSP is proportional to the amount of transmitter released but becomes saturated for large amounts of transmitter substance. Because the junction has a much larger time constant than the spacing between pulses, a temporal summation occurs. Hence new potentials are simply added to what remains of the previous, now partially decayed PSP, thereby yielding the so-called slow potential. The resulting magnitude of this dendritic depolarization is proportional to the average frequency at which pulses arrive at the synapse, a form of frequency-voltage coding. This slow potential is shown in Figure 3.27; a typical neuron may have $10^3$ to $10^5$ synapses of the type shown.

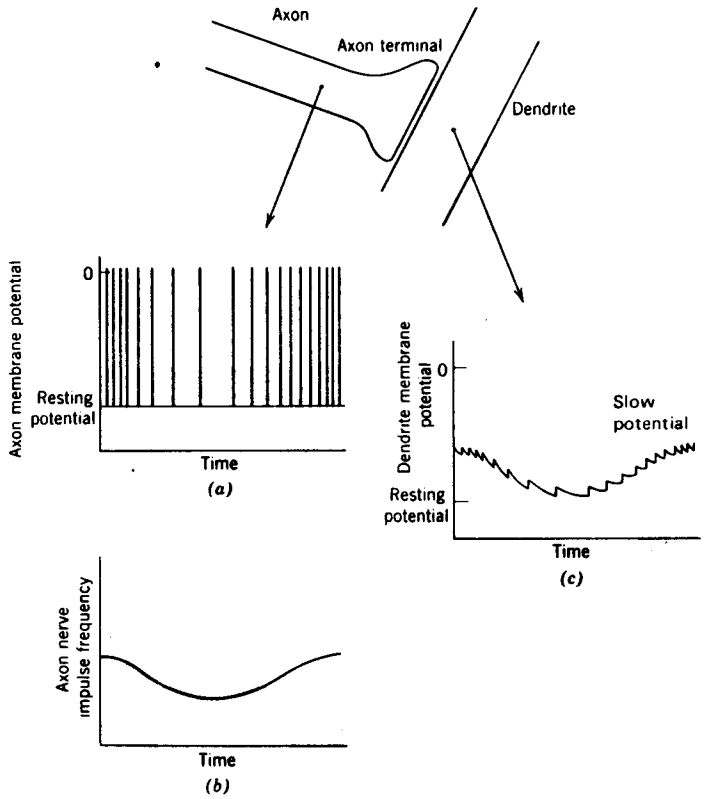



**Figure 3.27** The process of frequency-voltage coding at the synaptic junction of an axon and a dendrite. The input frequency-modulated pulse train (*a*) represents an average frequency versus time relationship (*b*), which is converted by the synapse into a slow potential on the dendrite (*c*). (*Adapted from C. F. Stevens, "Neurophysiology: A Primer," Wiley, New York, 1966, p. 33.*)

However, these originate from perhaps only about 10 other neurons [28]. Thus it would appear that each neuron must project a large number of synapses to each of the neurons to which it is connected. Because of this large set of input paths, an average of $10^4$ to $10^5$ PSPs per second will be conducted over all the synapses. Synaptic junctions usually occur between axons and dendrites but can also appear between axon and axon, dendrite and dendrite, and axon and cell body.

An interesting aspect from the point of view of computation is that there can exist two types of synapses at the junction, excitory and inhibitory, and these are not necessarily synonomous with positive and negative increments, respectively. Furthermore the amount of depolarization or hyperpolarization is weighted by each particular synapse, thereby providing multiplication by a parameter. The magnitude of the parameter may depend on such fixed anatomical features as the physical size of the synapse and the distance of the synapse from the soma, that is, the length of the dendrite. It turns out that adaptation is also possible at the synaptic input if it possesses an additional connection just prior to the same junction, thereby exhibiting an axo-axonal synapse. This less frequently observed synapse is used to transmit a slow potential (with respect to the frequency of the arriving pulses at the junction), which controls the incremental amount of transmitter substance secreted by the terminal. Therefore the presynaptic potential magnitude, which depends on this axon terminal membrane potential, is slowly altered, and in this way it controls the increments contributing to the amplitude of the slow potentials at the junction. In other words, incremental additions are multiplicatively modulated by the input from this axon. Since only depolarizing axo-axonal synapses have been found to date, this mechanism is referred to as "presynaptic inhibition."

So far we have discussed the mechanism by which frequency-coded information originating on the axon is transmitted via a synapse to a dendrite of a cell. We observed that the dendrite conveys a slow potential to the soma or cell body of the neuron. We note another interesting computational aspect that occurs at this point. The cell output projected along its axon only appears if the input slow potential is above a certain threshold. Under these circumstances, as we might expect, this output is none other than the stream of nerve impulses which we have previously called action potentials. In that experiment we had artificially stimulated the axon to respond. Here we observe that the dendritic inputs cause the pulses to form at the axon hillock, which is the junction between the soma and the cell axon. The action potential frequency is proportional to the input slow potential, as shown in Figure 3.28, where only one input to neuron N is shown.

What happens with the multitude of dendritic inputs to a cell? A process of weighted spatial summation in the cell body yields a weighted average of all the input signals, which is then linearly converted to the appropriate output frequency. Figure 3.29 demonstrates this process for the situation where one neuron input is inhibitory and the other excitory, resulting in a weighted



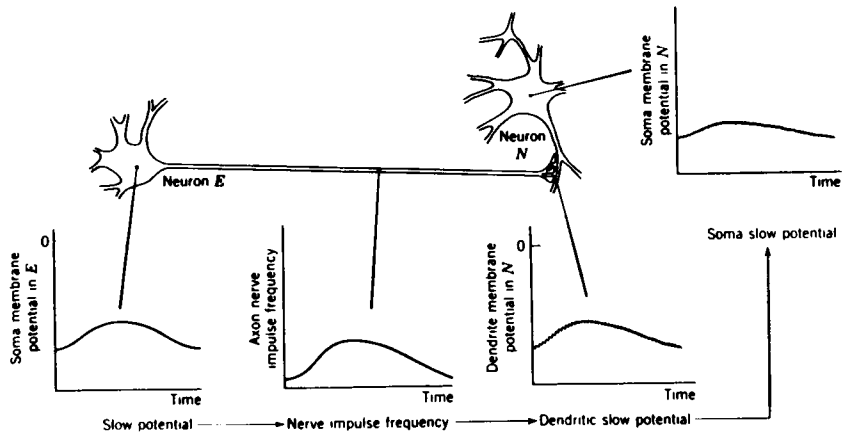



**Figure 3.28** The steps involved in communicating a signal from a neuron E to a neuron N. (*From C. F. Stevens, "Neurophysiology: A Primer," Wiley, New York, 1966, p. 50.*)

difference between the two slow potentials. We remind the reader that this is a simplification and that the significance of deviations from this model is not yet entirely clear. Additionally, in the retina only the ganglion cells generate pulse trains in the fashion described above; the other cells generate slow potentials.

Having described elemental cellular processing, we now consider how the nervous system communicates with the external visual environment. This process of photoreception is in fact quite similar to that of the other receptors in the body which are responsible for detecting other physical properties. The light intensity is photochemically transformed by such a specialized cell into a graded or slow potential. This is in distinction to the ganglion cells we have already discussed, which generate nerve impulses along their axons. The magnitude of the slow potential at any instant in time is proportional to the logarithm of the input light intensity. This mechanism for enhancing the dynamic range will be discussed in detail in Chapter 4. We have already mentioned how the amount of light that impinges upon the retina is quite rapidly controlled by the pupillary reflex over a range of about 100 to 1. Another, slower mechanism for modulating the input is the process of adaptation to a constant signal. The resulting slow potential decreases with time for constant input, thereby allowing the photoreceptor to operate in a range where it is more sensitive to the incoming light.

There is evidence that vertebrate receptors are hyperpolarized by a light stimulus. The rods and cones of the carp, frog, gekko, and mud puppy possess this property. On the other hand, receptors in invertebrate eyes are depolarized by the incoming light. In most cases the visual receptors are specialized in their function; however, in certain primitive living things the same cell may be responsible for the dual action of sensing the environment

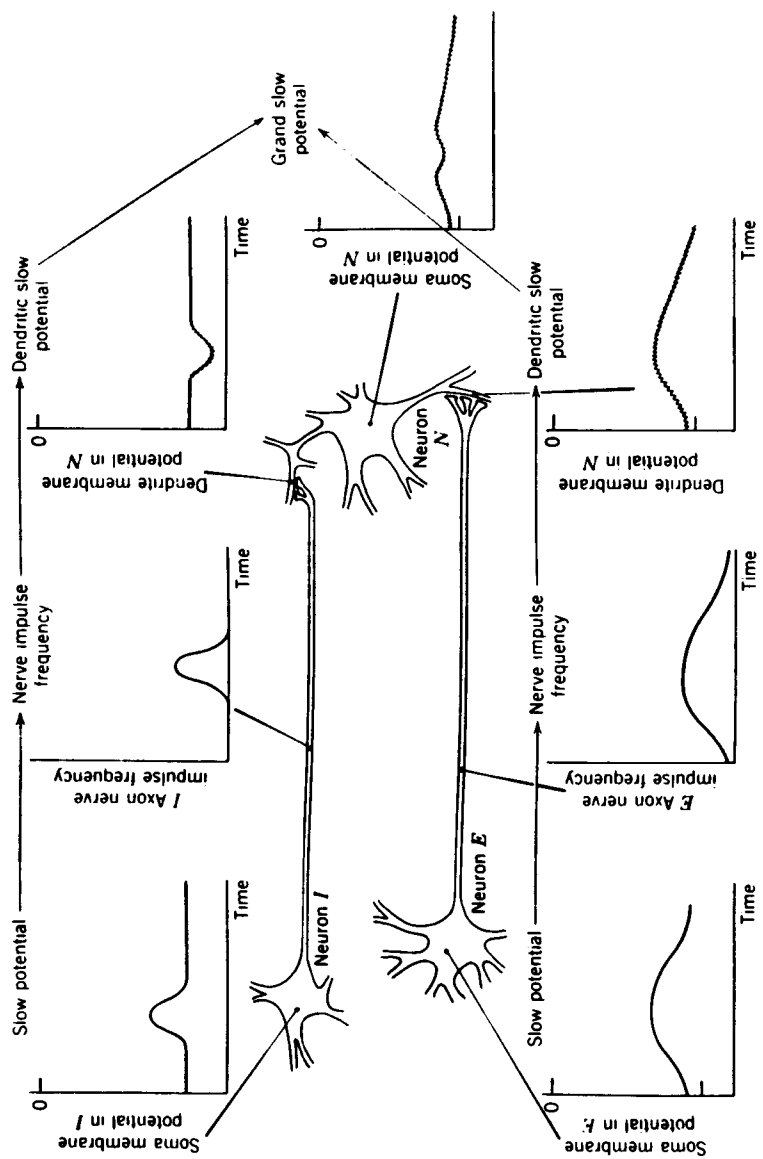



**Figure 3.29** A small neuronal network in which the weighted difference between two input signals is computed. (*From C. F. Stevens, "Neurophysiology: A Primer," Wiley, New York, 1966, p. 54.*)





and responding to it by controlling an affector. The higher organisms tend to possess cells which are more focused in their roles, thereby providing enhanced sensitivity for each independent function. We see therefore, that the sensory input to the "vision computer," the receptor mosaic, is capable of providing information about the light input with regard to both its intensity and spatial location.

Based on the above discussion, McCulloch and Pitts [13] postulated a simple mathematical model of a neuron, shown in Figure 3.30. The dendritic inputs to the neuron are given by $y_1, y_2, \ldots, y_n$ and the corresponding synaptic weights, either positive or negative, by $w_1, w_2, \ldots, w_n$. The axonal output $x$ is given by

$$x = \text{sign}\left[\sum_{i=1}^{n} (w_i y_i - \theta)\right] \tag{3.3}$$

where $\theta$ is a threshold value and obviously only takes on values of either $-1$ or $+1$. A more realistic model, in which the input and output signals represent actual firing rates, can also be derived [1]:

$$\tau \frac{du(t)}{dt} = -u(t) + \sum_{i=1}^{n} w_i y_i(t) \tag{3.4}$$

$$x(t) = f[u(t) - \theta] \tag{3.5}$$

Here the function $f$ may be considered to be monotonic with saturation at both negative and positive values, a slight relaxation of the stringent condition imposed by the sign function above. It turns out that as far as the macroscopic behavior of neural nets is concerned, these two models are equivalent. Thus the simpler McCulloch–Pitts representation is more convenient computation-

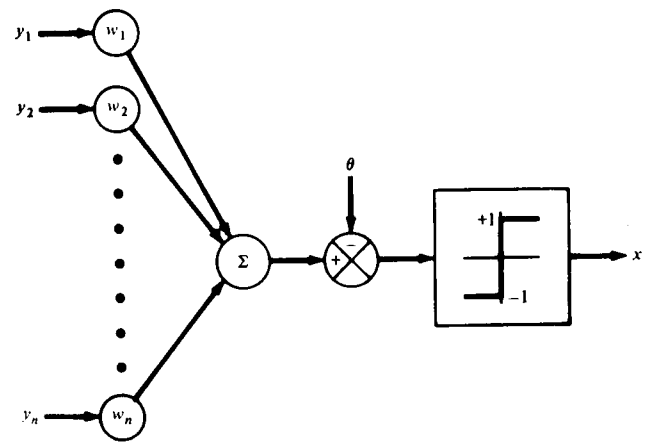



**Figure 3.30** The classical McCulloch-Pitts model of the neuron.

ally. Readers are referred to [1] for a discussion of the analysis of neuron networks using these models. Some interesting mathematical idealizations of these models, suitable for pattern analysis and recognition, are given in [16]. Adaptive neural networks are discussed in [29].

It is the connectivity of these neurons in the cortex which, of course, defines the particular functions being computed. The degree of complexity of this network in the mammalian brain turns out to be related to intelligence [28].

## REFERENCES


1. Amari, S. I., "A Mathematical Approach to Neural Systems," in Metzler, J. (ed.), "Systems Neuroscience," Academic, New York, 1977, pp. 67–117.
2. Baylor, D. A., Fuortes, M. G. F., and O'Bryan, P. M., "Receptive Fields of Cones in the Retina of the Turtle," *Journal of Physiology (London)*, vol. 173, no. 3, 1964, pp. 377–407.
3. Borwein, B., Borwein, D., Medeiros, J., and McGowan, J. W., "The Ultrastructure of Monkey Foveal Photoreceptors, with Special Reference to the Structure, Shape, Size, and Spacing of the Foveal Cones," *American Journal of Anatomy*, vol. 159, 1980, pp. 125–146.
4. Boycott, B. B., and Dowling, J. E., "Organization of the Primate Retina; Light Microscopy," *Philosophical Transactions of the Royal Society (London)*, ser. B, vol. 255, 1969, pp. 109–176.
5. Cajal, S. R., "La Rétine des Vertebres," *La Cellule*, vol. 9, 1893, pp. 17–257.
6. Cajal, S. R., "Histologie des Systèmes Nerveux de l'Homme et des Vertebres," vol. 2, Maloine, Paris, 1911.
7. Dowling, J. E., and Boycott, B. B., "Organization of the Primate Retina: Electron Microscopy," *Proceedings of the Royal Society (London)*, ser. B, vol. 166, no. 1002, 1966, pp. 80–111.
8. Graham, C. H., "Vision and Visual Perception," Wiley, New York, 1965.
9. Julesz, B., "Foundations of Cyclopean Perception," University of Chicago Press, Chicago, 1971.
10. Kolb, H., "Organization of the Outer Plexiform Layer of the Primate Retina: Electron Microscopy of Golgi-Impregnated Cells," *Philosophical Transactions of the Royal Society (London)*, ser. B, vol. 258, 1970, pp. 261–283.
11. Lowenstein, O., and Loewenfeld, I. E., "The Pupil," in Davson, H. (ed.), "The Eye," vol. 3, Academic, New York, 1962, pp. 231–267.
12. Marcelja, S., "Electrical Coupling of Photoreceptors in Retinal Network Models," *Biological Cybernetics*, vol. 39, no. 1, 1980, pp. 15–20.
13. McCulloch, W. S., and Pitts, W. H., "A Logical Calculus of the Ideas Imminent in Neural Nets," *Bulletin of Mathematical Biophysics*, vol. 5, 1943, pp. 115–133.
14. Missothen, L., "The Ultrastructure of the Human Retina," Arscia Uitgaven, N.V., Brussels, 1965.
15. Newman, E. A., and Hartline, P. H., "The Infrared 'Vision' of Snakes," *Scientific American*, vol. 246, no. 3, March 1982, pp. 116–127.
16. Nilsson, N. J., "Learning Machines, Foundations of Trainable Pattern Classifying Systems," McGraw-Hill, New York, 1965.
17. Rodieck, R. W., "The Vertebrate Retina," Freeman, San Francisco, 1973.
18. Rolls, P., "Photographic Optics," in Engels, C. E. (ed.), "Photography for the Scientist," Academic, New York, 1968, pp. 67–174.
19. Rushton, W. A. H., "Visual Pigments in Man," *Scientific American*, vol. 207, no. 5, November 1962, pp. 120–132.
20. Schiller, P. H., and Stryker, M., "Single-Unit Recording and Stimulation in Superior Colliculus of the Alert Rhesus Monkey," *Journal of Neurophysiology*, vol. 35, no. 6, 1972, pp. 915–924.
21. Semmlow, J., and Stark, L., "Simulation of a Biomechanical Model of the Human Pupil," *Mathematical Biosciences*, vol. 11, 1971, pp. 109–128.
22. Semmlow, J., and Chen, D. C., "A Simulation Model of the Human Pupil Light Reflex," *Mathematical Biosciences*, vol. 33, 1977, pp. 5–24.
23. Shantz, M., and Naka, K. I., "The Bipolar Cell," *Vision Research*, vol. 16, 1976, pp. 1517–1518.
24. Stark, L., "Stability, Oscillations, and Noise in the Human Pupil Servomechanism," *Proceeding of the IRE*, vol. 47, 1959, pp. 1925–1939.
25. Stark, L., "Pupillary Control System; Its Nonlinear Adaptive and Stochastic Engineering Design Characteristics," *Automatica*, vol. 5, 1969, pp. 655–676.
26. Stone, J., Dreher, B., and Leventhal, A., "Hierarchical and Parallel Mechanisms in the Organization of Visual Cortex," *Brain Research Reviews*, vol. 1, 1979, pp. 345–394.
27. Stubbs, D. F., "Frequency and the Brain," *Life Sciences*, vol. 18, 1976, pp. 1–14.
28. Stubbs, D. F., "Connectivity and the Brain," *Kybernetes*, vol. 7, no. 2, 1978, pp. 93–98.
29. Sutton, R. S., and Barto, A. G., "Toward a Modern Theory of Adaptive Networks: Expectation and Prediction," *Psychological Review*, vol. 88, no. 2, 1981, pp. 135–170.
30. Szentagothai, J., and Arbib, M. A., "Conceptual Models of Neural Organization," *Neurosciences Research Progress Bulletin*, vol. 12, no. 1, 1975, pp. 90–93.
31. Tanimoto, S. L., "Image Data Structures," in Tanimoto, S., and Klinger, A. (eds.), "Structured Computer Vision, Machine Perception Through Hierarchical Computation Structures," Academic, New York, 1980, pp. 31–55.
32. Tryon, W. T., "Pupillometry: A Survey of Sources of Variation," *Psychophysiology*, vol. 12, no. 1, 1975, pp. 90–93.
33. Webster, J. G., "Pupillacy Light Reflex: Development of Teaching Models," *IEEE Transactions on Biomedical Engineering*, vol. BME-18, no. 3, 1971, pp. 25–33.
34. Weiman, C. F. R., and Chaikin, G. M., "Logarithmic Spiral Grids for Image Processing," *Proceedings 1979 IEEE Computer Society Conference on Pattern Recognition and Image Processing*, Chicago, 1979, pp. 25–31.
35. Wurtz, R. H., and Goldberg, M. E., "Activity of Superior Colliculus in Behaving Monkey, IV. Effects of Lesions on Eye Movements," *Journal of Neurophysiology*, vol. 35, no. 4, 1972, pp. 587–596.


## BIBLIOGRAPHY

A good introduction to the neurophysiology and neuroanatomy discussed in this chapter is provided by the many very well written articles which have appeared in *Scientific American* over the years. In particular, we recommend the collection of offprints on the mechanisms and models of perception [7] and the September 1979 issue which deals specifically with the brain. A very complete early history of the development of man's knowledge of the eye is given in [20].

Two more advanced and often cited books which will still be understandable to the lay reader are by Brindley [3] and Pirenne [19], both distinguished workers in the field. Absolutely everything you would ever want to know about the vertebrate retina can be found in the book by Rodieck [22], a scholarly tome but one containing much accessible material. The "Handbook of Sensory Physiology," published by Springer-Verlag, and the "Handbook of Perception," published by Academic Press, have both issued many volumes over the years which contain articles of a survey nature by active researchers. These are

particularly useful for obtaining a detailed knowledge about the history and development of a specific research problem.

For the reader with no particular background in the biological sciences, Kuffler and Nicholls [10] explain how the nervous system functions. State-of-the-art papers including discussions on the neurosciences can be found in the *Neuroscience Research Progress Bulletins* published by the MIT Press in Cambridge, Massachusetts. These are edited and written by various scholars and are concerned with how the central nervous system controls the behavior and thought processes of man. An excellent and succinct introductory book on neural physiology is that by Stevens [25]. Most of the material in Section 3.5 is adapted from this source. A more advanced treatment can be found in [16] and [17]. A recent review of the literature on the visual pathway appears in [11].

In contrast to the conventional approach to the subject in the above sources, the reader may also wish to consult two other books which view the brain from completely different standpoints. The first, by Michael Arbib, is concerned with the brain as a cybernetic system and is aimed at the intelligent layman [1]. The second, by Uttal, is a unique and detailed exposition of the relationship between the psychology and the biology of the human mind [26].

The study of the simpler visual system of insects is also intriguing and was originated in modern times by Muller in 1829. The existence of a compound eye in insects led him to propose a theory of mosaic vision involving an array sensor and a processor. Readers interested in the subject of insect vision may consult [5, 8, 9, 13, 21, 23, 24, 27].

Research into the visual systems of simpler and smaller animals has provided significant clues to the properties of the more complex human visual system. Section 3.5 deals with certain elementary aspects of the neural networks which necessarily constitute the biological vision "computer." Further discussions may be found in [2, 4, 6, 12, 14, 15, 18].


1. Arbib, M. A., "The Metaphorical Brain, An Introduction to Cybernetics as Artificial Intelligence and Brain Theory," Wiley, New York, 1972.
2. Arbib, M. A., Kilmer, W. L., and Spinelli, D. N., "Neural Models and Memory," in Rosenzweig, M. R., and Bennet, E. L. (eds.), "Neural Mechanisms of Learning and Memory," MIT Press, Cambridge, Mass., 1976, pp. 109–132.
3. Brindley, G. S., "Physiology of the Retina and Visual Pathway," Physiological Society Monograph No. 6, Edward Arnold Ltd., London, 1970.
4. Caianiello, E. R. (ed.), "Neural Networks," Springer-Verlag, New York, 1968.
5. Dethier, V. G., "The Physiology of Insect Senses," Wiley, New York, 1963.
6. Harmon, L. D., and Lewis, E. R., "Neural Modelling," *Physiological Review*, vol. 46, no. 5, July 1966, pp. 513–591.
7. Held, R., and Richards, W., (eds.), "Perception: Mechanisms and Models," Freeman, San Francisco, 1971.
8. Horridge, E. A. (ed.), "The Compound Eye and Vision of Insects," Clarendon Press, Oxford, 1975.
9. Horridge, E. A., "The Compound Eye of Insects," *Scientific American*, July 1977, pp. 108–120.
10. Kuffler, S. W., and Nicholls, J. G., "From Neuron to Brain," Sinauer Associates, Sunderland, Mass., 1976.
11. Lennie, P., "Parallel Visual Pathways: A Review," *Vision Research*, vol. 20, 1980, pp. 561–594.





12. MacGregor, R. J., and Lewis, E. R., "Neural Modelling," Plenum, New York, 1977.
13. Mazokhin-Porshnyakov, G. A., "Insect Vision," Masiromi, R., and Masiromi, L. (trans.), Goldsmith, T. H., trans. ed., Plenum, New York, 1969.
14. Morishita, I., and Yajima, A., "Analysis and Simulation of Networks of Mutually Inhibiting Neurons," *Kybernetik*, vol. 11, no. 3, 1972, pp. 154–156.
15. Mountcastle, V. B., "The Problem of Sensing and the Neural Coding of Sensory Events," in Quarton, G. C., Melnechnuk, T., and Schmidt, F. O. (eds.), "The Neuro Sciences," Rockefeller University Press, New York, 1967, pp. 393–408.
16. Mountcastle, V. B., "Medical Physiology," vol. II, Mosby, St. Louis, 1968.
17. Ochs, S., "Elements of Neurophysiology," Wiley, New York, 1965.
18. Perkel, D. H., and Bullock, T. H., "Neural Coding," *Neurosciences Research Progress Bulletin*, vol. 6, no. 3, 1968, pp. 221–348.
19. Pirenne, M. H., "Vision and the Eye," 2d ed., Associated Book Publishers, London, 1967.
20. Polyak, S. L., "The Retina," University of Chicago Press, Chicago, 1941.
21. Rockstein, M. (ed.), "The Physiology of Insects," 2d ed., vol. 1, Academic, New York, 1973.
22. Rodieck, R. W., "The Vertebrate Retina, Principles of Structure and Function," Freeman, San Francisco, 1973, p. 366.
23. Smith, D. S., "Insect Cells, Their Structure and Function," Oliver and Boyd, Edinburgh, 1968.
24. Smythe, R. H., "Vision in the Animal World," Macmillan, New York, 1975.
25. Stevens, C. F., "Neurophysiology: A Primer," Wiley, New York, 1966.
26. Uttal, W. R., "The Psychobiology of the Mind," Lawrence Erlbaum Associates, Hillsdale, N.J., 1978.
27. Wiggleworth, V. B., "The Principles of Insect Physiology," Halsted, New York, 1972.