Stetten, G. "Vision System."

The Biomedical Engineering Handbook: Second Edition.

Ed. Joseph D. Bronzino

Boca Raton: CRC Press LLC, 2000

4

Vision System

George Stetten
Duke University

David Marr, an early pioneer in computer vision, defined *vision* as extracting "… from images of the external world, a description that is useful for the viewer and not cluttered with irrelevant information" [Marr, 1982]. Advances in computers and video technology in the past decades have created the expectation that artificial vision should be realizable. The nontriviality of the task is evidenced by the continuing proliferation of new and different approaches to computer vision without any observable application in our everyday lives. Actually, computer vision is already offering practical solutions in industrial assembly and inspection, as well as for military and medical applications, so it seems we are beginning to master some of the fundamentals. However, we have a long way to go to match the vision capabilities of a 4-year-old child. In this chapter we will explore what is known about how nature has succeeded at this formidable task — that of interpreting the visual world.

4.1   Fundamentals of Vision Research

Research into biologic vision systems has followed several distinct approaches. The oldest is psychophysics, in which human and animal subjects are presented with visual stimuli and their responses recorded. Important early insights also were garnered by correlating clinical observations of visual defects with known neuroanatomic injury. In the past 50 years, a more detailed approach to understanding the mechanisms of vision has been undertaken by inserting small electrodes deep within the living brain to monitor the electrical activity of individual neurons and by using dyes and biochemical markers to track the anatomic course of nerve tracts. This research has led to a detailed and coherent, if not complete, theory of a visual system capable of explaining the discrimination of form, color, motion, and depth. This theory has been confirmed by noninvasive radiologic techniques that have been used recently to study the physiologic responses of the visual system, including positron emission tomography [Zeki et al., 1991] and functional magnetic resonance imaging [Belliveau et al., 1992; Cohen and Bookheimer, 1994], although these noninvasive techniques provide far less spatial resolution and thus can only show general regions of activity in the brain.

4.2   A Modular View of the Vision System

The Eyes

Movement of the eyes is essential to vision, not only allowing rapid location and tracking of objects but also preventing stationary images on the retina, which are essentially invisible. Continual movement of the image on the retina is essential to the visual system.

The eyeball is spherical and therefore free to turn in both the horizontal and vertical directions. Each eye is rotated by three pairs of mutually opposing muscles, innervated by the oculomotor nuclei in the brainstem. The eyes are coordinated as a pair in two useful ways: turning together to find and follow objects and turning inward to allow adjustment for parallax as objects become closer. The latter is called *convergence.*

The optical portion of the eye, which puts an image on the retina, is closely analogous to a photographic or television camera. Light enters the eye, passing through a series of transparent layers — the cornea, the aqueous humor, the lens, and the vitreous body — to eventually project on the retina.

The *cornea,* the protective outer layer of the eye, is heavily innervated with sensory neurons, triggering the blink reflex and tear duck secretion in response to irritation. The cornea is also an essential optical element, supplying two-thirds of the total refraction in the eye. Behind the cornea is a clear fluid, the *aqueous humor,* in which the central aperture of the iris, the pupil, is free to constrict or dilate. The two actions are accomplished by opposing sets of muscles.

The *lens,* a flexible transparent object behind the iris, provides the remainder of refraction necessary to focus an image on the retina. The ciliary muscles surrounding the lens can increase the lens' curvature, thereby decreasing its focal length and bringing nearer objects into focus. This is called *accommodation.* When the ciliary muscles are at rest, distant objects are in focus. There are no contradictory muscles to flatten the lens. This depends simply on the elasticity of the lens, which decreases with age. Behind the lens is the *vitreous humor,* consisting of a semigelatinous material filling the volume between the lens and the retina.

The Retina

The retina coats the back of the eye and is therefore spherical, not flat, making optical magnification constant at 3.5 degrees of scan angle per millimeter. The retina is the neuronal front end of the visual system, the image sensor. In addition, it accomplishes the first steps in edge detection and color analysis before sending the processed information along the optic nerve to the brain. The retina contains five major classes of cells, roughly organized into layers. The dendrites of these cells each occupy no more than 1 to 2 mm^2 in the retina, limiting the extent of spatial integration from one layer of the retina to the next.

First come the *receptors,* which number approximately 125 million in each eye and contain the light-sensitive pigments responsible for converting photons into chemical energy. Receptor cells are of two general varieties: *rods* and *cones.* The cones are responsible for the perception of color, and they function only in bright light. When the light is dim, only rods are sensitive enough to respond. Exposure to a single photon may result in a measurable increase in the membrane potential of a rod. This sensitivity is the result of a chemical cascade, similar in operation to the photo multiplier tube, in which a single photon generates a cascade of electrons. All rods use the same pigment, whereas three different pigments are found in three separate kinds of cones.

Examination of the retina with an otoscope reveals its gross topography. The yellow circular area occupying the central 5 degrees of the retina is called the *macula lutea,* within which a small circular pit called the *fovea* may be seen. Detailed vision occurs only in the fovea, where a dense concentration of cones provides visual activity to the central 1 degree of the visual field.

On the inner layer of the retina one finds a layer of *ganglion cells,* whose axons make up the optic nerve, the output of the retina. They number approximately 1 million, or less than 1% of the number of receptor cells. Clearly, some data compression has occurred in the space between the receptors and the ganglion cells. Traversing this space are the *bipolar cells,* which run from the receptors through the retina to the ganglion cells. Bipolar cells exhibit the first level of information processing in the visual system; namely, their response to light on the retina demonstrates "center/surround" receptive fields. By this I mean that a small dot on the retina elicits a response, while the area surrounding the spot elicits the opposite response. If both the center and the surround are illuminated, the net result is no response.

Thus bipolar cells respond only at the border between dark and light areas. Bipolar cells come in two varieties, on-center and off-center, with the center respectively brighter or darker than the surround.

The center response of bipolar cells results from direct contact with the receptors. The surround response is supplied by the *horizontal cells,* which run parallel to the surface of the retina between the receptor layer and the bipolar layer, allowing the surrounding area to oppose the influence of the center. The *amacrine cells,* a final cell type, also run parallel to the surface but in a different layer, between the bipolar cells and the ganglion cells, and are possibly involved in the detection of motion.

Ganglion cells, since they are triggered by bipolar cells, also have center/surround receptive fields and come in two types, on-center and off-center. On-center ganglion cells have a receptive field in which illumination of the center increases the firing rate and a surround where it decreases the rate. Off-center ganglion cells display the opposite behavior. Both types of ganglion cells produce little or no change in firing rate when the entire receptive field is illuminated, because the center and surround cancel each other. As in many other areas of the nervous system, the fibers of the optic nerve use frequency encoding to represent a scalar quantity.

Multiple ganglion cells may receive output from the same receptor, since many receptive fields overlap. However, this does not limit overall spatial resolution, which is maximum in the fovea, where two points separated by 0.5 minutes of arc may be discriminated. This separation corresponds to a distance on the retina of 2.5 μm, which is approximately the center-to-center spacing between cones. Spatial resolution falls off as one moves away from the fovea into the peripheral vision, where resolution is as low as 1 degree of arc.

Several aspects of this natural design deserve consideration. Why do we have center/surround receptive fields? The ganglion cells, whose axons make up the optic nerve, do not fire unless there is meaningful information, i.e., a border, falling within the receptive field. It is the edge of a shape we see rather than its interior. This represents a form of data compression. Center/surround receptive fields also allow for relative rather than absolute measurements of color and brightness. This is essential for analyzing the image independent of lighting conditions. Why do we have both on-center and off-center cells? Evidently, both light and dark are considered information. The same shape is detected whether it is lighter or darker than the background.

Optic Chiasm

The two optic nerves, from the left and right eyes, join at the optic chiasm, forming a *hemidecussation,* meaning that half the axons cross while the rest proceed uncrossed. The resulting two bundles of axons leaving the chiasm are called the *optic tracts.* The left optic tract contains only axons from the left half of each retina. Since the images are reversed by the lens, this represents light from the right side of the visual field. The division between the right and left optic tracts splits the retina down the middle, bisecting the fovea. The segregation of sensory information into the contralateral hemispheres corresponds to the general organization of sensory and motor centers in the brain.

Each optic tract has two major destinations on its side of the brain: (1) the superior colliculus and (2) the lateral geniculate nucleus (LGN). Although topographic mapping from the retina is scrambled within the optic tract, it is reestablished in both major destinations so that right, left, up, and down in the image correspond to specific directions within those anatomic structures.

Superior Colliculus

The *superior colliculus* is a small pair of bumps on the dorsal surface of the midbrain. Another pair, the *inferior colliculus,* is found just below it. Stimulation of the superior colliculus results in contralateral eye movement. Anatomically, output tracts from the superior colliculus run to areas that control eye and neck movement. Both the inferior and superior colliculi are apparently involved in locating sound. In the bat, the inferior colliculus is enormous, crucial to that animal's remarkable echolocation abilities.

The superior colliculus processes information from the inferior colliculus, as well as from the retina, allowing the eyes to quickly find and follow targets based on visual and auditory cues.

Different types of eye movements have been classified. The *saccade* (French, for "jolt") is a quick motion of the eyes over a significant distance. The saccade is how the eyes explore an image, jumping from landmark to landmark, rarely stopping in featureless areas. *Nystagmus* is the smooth pursuit of a moving image, usually with periodic backward saccades to lock onto subsequent points as the image moves by. *Microsaccades* are small movements, several times per second, over 1 to 2 minutes of arc in a seemingly random direction. Microsaccades are necessary for sight; their stabilization leads to effective blindness.

LGN

The thalamus is often called "the gateway to the cortex" because it processes much of the sensory information reaching the brain. Within the thalamus, we find the *lateral geniculate nucleus* (LGN), a peanut-sized structure that contains a single synaptic stage in the major pathway of visual information to higher centers. The LGN also receives information back from the cortex, so-called reentrant connections, as well as from the nuclei in the brainstem that control attention and arousal.

The cells in the LGN are organized into three pairs of layers. Each pair contains two layers, one from each eye. The upper two pairs consist of parvocellular cells (*P cells*) that respond with preference to different colors. The remaining lower pair consists of magnocellular cells (*M cells*) with no color preference (Fig. 4.1). The topographic mapping is identical for all six layers; i.e., passing through the layers at a given point yields synapses responding to a single area of the retina. Axons from the LGN proceed to the primary visual cortex in broad bands, the *optic radiations,* preserving this topographic mapping and displaying the same center/surround response as the ganglion cells.

FIGURE 4.1 Visual pathways to cortical areas showing the separation of information by type. The lateral geniculate nucleus (LGN) and areas V1 and V2 act as gateways to more specialized higher areas.

Area V1

The LGN contains approximately 1.5 million cells. By comparison, the *primary visual cortex,* or *striate cortex,* which receives the visual information from the LGN, contains 200 million cells. It consists of a thin (2-mm) layer of gray matter (neuronal cell bodies) over a thicker collection of white matter (myelinated axons) and occupies a few square inches of the occipital lobes. The primary visual cortex has been called *area 17* from the days when the cortical areas were first differentiated by their cytoarchitectonics (the microscopic architecture of their layered neurons). In modern terminology, the primary visual cortex is often called *visual area 1,* or simply *V1.*

Destroying any small piece of V1 eliminates a small area in the visual field, resulting in *scotoma,* a local blind spot. Clinical evidence has long been available that a scotoma may result from injury, stroke, or tumor in a local part of V1. Between neighboring cells in V1's gray matter, horizontal connections are at most 2 to 5 mm in length. Thus, at any given time, the image from the retina is analyzed piecemeal in V1. Topographic mapping from the retina is preserved in great detail. Such mapping is seen elsewhere in the brain, such as in the somatosensory cortex [Mountcastle, 1957]. Like all cortical surfaces, V1 is a highly convoluted sheet, with much of its area hidden within its folds. If unfolded, V1 would be roughly
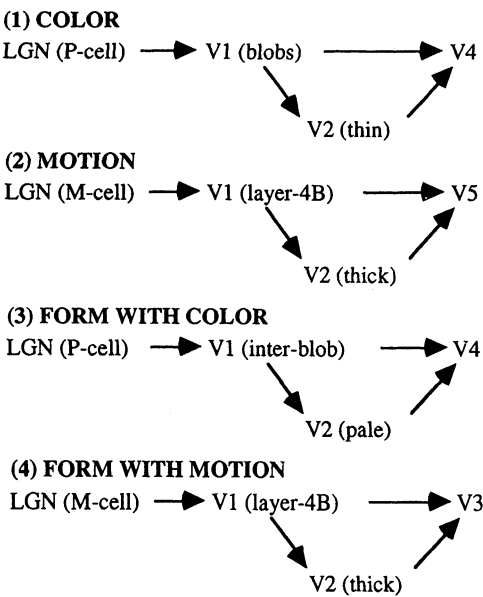
pear-shaped, with the top of the pear processing information from the fovea and the bottom of the pear processing the peripheral vision. Circling the pear at a given latitude would correspond roughly to circling the fovea at a fixed radius.

The primary visual cortex contains six layers, numbered 1 through 6. Distinct functional and anatomic types of cells are found in each layer. Layer 4 contains neurons that receive information from the LGN. Beyond the initial synapses, cells demonstrate progressively more complex responses. The outputs of V1 project to an area known as *visual area 2* (*V2*), which surrounds V1, and to higher visual areas in the occipital, temporal, and parietal lobes as well as to the superior colliculus. V1 also sends reentrant projections back to the LGN. Reentrant projections are present at almost every level of the visual system [Felleman and Essen, 1991; Edelman, 1978].

Cells in V1 have been studied extensively in animals by inserting small electrodes into the living brain (with surprisingly little damage) and monitoring the individual responses of neurons to visual stimuli. Various subpopulations of cortical cells have thus been identified. Some, termed *simple cells,* respond to illuminated edges or bars at specific locations and at specific angular orientations in the visual field. The angular orientation must be correct within 10 to 20 degrees for the particular cell to respond. All orientations are equally represented. Moving the electrode parallel to the surface yields a smooth rotation in the orientation of cell responses by about 10 degrees for each 50 μm that the electrode is advanced. This rotation is subject to reversals in direction, as well as "fractures," or sudden jumps in orientation.

Other cells, more common than simple cells, are termed *complex cells.* Complex cells respond to a set of closely spaced parallel edges within a particular receptive field. They may respond specifically to movement perpendicular to the orientation of the edge. Some prefer one direction of movement to the other. Some complex and simple cells are *end-stopped,* meaning they fire only if the illuminated bar or edge does not extend too far. Presumably, these cells detect corners, curves, or discontinuities in borders and lines. End-stopping takes place in layers 2 and 3 of the primary visual cortex. From the LGN through the simple cells and complex cells, there appears to be a sequential processing of the image. It is probable that simple cells combine the responses of adjacent LGN cells and that complex cells combine the responses of adjacent simple cells.

A remarkable feature in the organization of V1 is binocular convergence, in which a single neuron responds to identical receptive fields in both eyes, including location, orientation, and directional sensitivity to motion. It does not occur in the LGN, where axons from the left and right eyes are still segregated into different layers. Surprisingly, binocular connections to neurons are present in V1 at birth. Some binocular neurons are equally weighted in terms of responsiveness to both eyes, while others are more sensitive to one eye than to the other. One finds columns containing the latter type of cells in which one eye dominates, called *ocular dominance columns,* in uniform bands approximately 0.5 mm wide everywhere in V1. Ocular dominance columns occur in adjacent pairs, one for each eye, and are prominent in animals with forward-facing eyes, such as cats, chimpanzees, and humans. They are nearly absent in rodents and other animals whose eyes face outward.

The topography of orientation-specific cells and of ocular dominance columns is remarkably uniform throughout V1, which is surprising because the receptive fields near the fovea are 10 to 30 times smaller than those at the periphery. This phenomenon is called magnification. The fovea maps to a greater relative distance on the surface of V1 than does the peripheral retina, by as much as 36-fold [Daniel and Whitteridge, 1961]. In fact, the majority of V1 processes only the central 10 degrees of the visual field. Both simple and complex cells in the foveal portion can resolve bars as narrow as 2 minutes of arc. Toward the periphery, the resolution falls off to 1 degree of arc.

As an electrode is passed down through the cortex *perpendicular* to the surface, each layer demonstrates receptive fields of characteristic size, the smallest being at layer 4, the input layer. Receptive fields are larger in other layers due to lateral integration of information. Passing the electrode *parallel* to the surface of the cortex reveals another important uniformity to V1. For example, in layer 3, which sends output fibers to higher cortical centers, one must move the electrode approximately 2 mm to pass from one collection of receptive fields to another that does not overlap. An area approximately 2 mm across thus represents the smallest unit piece of V1, i.e., that which can completely process the visual information.

Indeed, it is just the right size to contain a complete set of orientations and more than enough to contain information from both eyes. It receives a few tens of thousands of fibers from the LGN, produces perhaps 50,000 output fibers, and is fairly constant in cytoarchitectonics whether at the center of vision, where it processes approximately 30 minutes of arc, or at the far periphery, where it processes 7 to 8 degrees of arc.

The topographic mapping of the visual field onto the cortex suffers an abrupt discontinuity between the left and right hemispheres, and yet our perception of the visual scene suffers no obvious rift in the midline. This is due to the *corpus collousum,* an enormous tract containing at least 200 million axons, that connects the two hemispheres. The posterior portion of the corpus collousum connects the two halves of V1, linking cells that have similar orientations and whose receptive fields overlap in the vertical midline. Thus a perceptually seamless merging of left and right visual fields is achieved. Higher levels of the visual system are likewise connected across the corpus collousum. This is demonstrated, for example, by the clinical observation that cutting the corpus collousum prevents a subject from verbally describing objects in the left field of view (the right hemisphere). Speech, which normally involves the left hemisphere, cannot process visual objects from the right hemisphere without the corpus collousum.

By merging the information from both eyes, V1 is capable of analyzing the distance to an object. Many cues for depth are available to the visual system, including occlusion, parallax (detected by the convergence of the eyes), optical focusing of the lens, rotation of objects, expected size of objects, shape based on perspective, and shadow casting. Stereopsis, which uses the slight difference between images due to the parallax between the two eyes, was first enunciated in 1838 by Sir Charles Wheatstone and its probably the most important cue [Wheatstone, 1838]. Fixating on an object causes it to fall on the two foveas. Other objects that are nearer become outwardly displaced on the two retinas, while objects that are farther away become inwardly displaced. About 2 degrees of horizontal disparity is tolerated, with fusion by the visual system into a single object. Greater horizontal disparity results in double vision. Almost no vertical displacement (a few minutes of arc) is tolerated. Physiologic experiments have revealed a particular class of complex cells in V1 which are *disparity tuned.* They fall into three general classes. One class fires only when the object is at the fixation distance, another only when the object is nearer, and a third only when it is farther away [Poggio and Talbot, 1981]. Severing the corpus collousum leads to a loss of stereopsis in the vertical midline of the visual field.

When the inputs to the two retinas cannot be combined, one or the other image is rejected. This phenomenon is known as *retinal rivalry* and can occur in a piecewise manner or can even lead to blindness in one eye. The general term *amblyopia* refers to the partial or complete loss of eyesight not caused by abnormalities in the eye. The most common form of amblyopia is caused by *strabismus,* in which the eyes are not aimed in a parallel direction but rather are turned inward (cross-eyed) or outward (wall-eyed). This condition leads to habitual suppression of vision from one of the eyes and sometimes to blindness in that eye or to *alternation,* in which the subject maintains vision in both eyes by using only one eye at a time. Cutting selected ocular muscles in kittens causes strabismus, and the kittens respond by alternation, preserving functional vision in both eyes. However, the number of cells in the cortex displaying binocular responses is greatly reduced. In humans with long-standing alternating strabismus, surgical repair making the eyes parallel again does not bring back a sense of depth. Permanent damage has been caused by the subtle condition of the images on the two retinas not coinciding. This may be explained by the Hebb model for associative learning, in which temporal association between inputs strengthens synaptic connections [Hebb, 1961].

Further evidence that successful development of the visual system depends on proper input comes from clinical experience with children who have *cataracts* at birth. Cataracts constitute a clouding of the lens, permitting light, but not images, to reach the retina. If surgery to remove the cataracts is delayed until the child is several years old, the child remains blind even though images are restored to the retina. Kittens and monkeys whose eyelids are sown shut during a critical period of early development stay blind even when the eyes are opened. Physiologic studies in these animals show very few cells responding in the visual cortex. Other experiments depriving more specific elements of an image, such as certain orientations or motion in a certain direction, yield a cortex without the corresponding cell type.

Color

Cones, which dominate the fovea, can detect wavelengths between 400 and 700 nm. The population of cones in the retina can be divided into three categories, each containing a different pigment. This was established by direct microscopic illumination of the retina [Wald, 1974; Marks et al., 1964]. The pigments have a bandwidth on the order of 100 nm, with significant overlap, and with peak sensitivities at 560 nm (yellow-green), 530 nm (blue-green), and 430 nm (violet). These three cases are commonly known as red, green, and blue. Compared with the auditory system, whose array of cochlear sensors can discriminate thousands of different sonic frequencies, the visual system is relatively impoverished with only three frequency parameters. Instead, the retina expends most of its resolution on spatial information. Color vision is absent in many species, including cats, dogs, and some primates, as well as in most nocturnal animals, since cones are useless in low light.

By having three types of cones at a given locality on the retina, a simplified spectrum can be sensed and represented by three independent variables, a concept known as *trichromacy.* This model was developed by Thomas Young and Hermann von Helmholtz in the 19th century before neurobiology existed and does quite well at explaining the retina [Young, 1802; Helmholtz, 1889]. The model is also the underlying basis for red-green-blue (RGB) video monitors and color television [Ennes, 1981]. Rods do not help in discriminating color, even though the pigment in rods does add a fourth independent sensitivity peak.

Psychophysical experimentation yields a complex, redundant map between spectrum and perceived color, or *hue,* including not only the standard red, orange, yellow, green, and blue but hues such as pink, purple, brown, and olive green that are not themselves in the rainbow. Some of these may be achieved by introducing two more variables: *saturation,* which allows for mixing with white light, and *intensity,* which controls the level of color. Thus three variables are still involved: hue, saturation, and intensity.

Another model for color vision was put forth in the 19th century by Ewald Hering [Hering, 1864]. This theory also adheres to the concept of trichromacy, espousing three independent variables. However, unlike the Young-Helmholtz model, these variables are signed; they can be positive, negative, or zero. The resulting three axes are *red-green, yellow-blue,* and *black-white.* The Hering model is supported by the physiologic evidence for the center/surround response, which allows for positive as well as negative information. In fact, two populations of cells, activated and suppressed along the red-green and yellow-blue axes, have been found in monkey LGN. Yellow is apparently detected by a combination of red and green cones.

The Hering model explains, for example, the perception of the color brown, which results only when orange or yellow is surrounded by a brighter color. It also accounts for the phenomenon of color constancy, in which the perceived color of an object remains unchanged under differing ambient light conditions provided background colors are available for comparison. Research into color constancy was pioneered in the laboratory of Edwin Land [Land and McCann, 1971]. As David Hubel says, "We require color borders for color, just as we require luminance borders for black and white" [Hubel, 1988, p. 178]. As one might expect, when the corpus collousum is surgically severed, color constancy is absent across the midline.

Color processing in V1 is confined to small circular areas, known as *blobs,* in which *double-opponent cells* are found. They display a center/surround behavior based on the red-green and yellow-blue axes but lack orientation selectivity. The V1 blobs were first identified by their uptake of certain enzymes, and only later was their role in color vision discovered [Livingstone and Hubel, 1984]. The blobs are especially prominent in layers 2 and 3, which receive input from the P cells of the LGN.

Higher Cortical Centers

How are the primitive elements of image processing so far discussed united into an understanding of the image? Beyond V1 are many higher cortical centers for visual processing, at least 12 in the occipital lobe and others in the temporal and parietal lobes. Areas V2 receives axons from both the blob and interblob areas of V1 and performs analytic functions such as filling in the missing segments of an edge. V2 contains three areas categorized by different kinds of stripes: *thick stripes* which process relative horizontal position

and stereopsis, *thin stripes* which process color without orientations, and *pale stripes* which extend the process of end-stopped orientation cells.

Beyond V2, higher centers have been labeled V3, V4, V5, etc. Four parallel systems have been delineated [Zeki, 1992], each system responsible for a different attribute of vision, as shown in Fig. 4.1. This is obviously an oversimplification of a tremendously complex system.

Corroborative clinical evidence supports this model. For example, lesions in V4 lead to *achromatopsia,* in which a patient can only see gray and cannot even recall colors. Conversely, a form of poisoning, *carbon monoxide chromatopsia,* results when the V1 blobs and V2 thin stripes selectively survive exposure to carbon monoxide thanks to their rich vasculature, leaving the patient with a sense of color but not of shape. A lesion in V5 leads to *akinetopsia,* in which objects disappear.

As depicted in Fig. 4.1, all visual information is processed through V1 and V2, although discrete channels within these areas keep different types of information separate. A total lesion of V1 results in the perception of total blindness. However, not all channels are shown in Fig. 4.1, and such a "totally blind" patient may perform better than randomly when forced to guess between colors or between motion in different directions. The patient with this condition, called *blindsight,* will deny being able to see anything [Weiskrantz, 1990].

Area V1 preserves retinal topographic mapping and shows receptive fields, suggesting a piecewise analysis of the image, although a given area of V1 receives sequential information from disparate areas of the visual environment as the eyes move. V2 and higher visual centers show progressively larger receptive fields and less defined topographic mapping but more specialized responses. In the extreme of specialization, neurobiologists joke about the "grandmother cell," which would respond only to a particular face. No such cell has yet been found. However, cortical regions that respond to faces in general have been found in the temporal lobe. Rather than a "grandmother cell," it seems that face-selective neurons are members of ensembles for coding facts [Gross and Sergen, 1992].

Defining Terms

Binocular convergence: The response of a single neuron to the same location in the visual field of each eye.

Color constancy: The perception that the color of an object remains constant under different lighting conditions. Even though the spectrum reaching the eye from that object can be vastly different, other objects in the field of view are used to compare.

Cytoarchitectonics: The organization of neuron types into layers as seen by various staining techniques under the microscope. Electrophysiologic responses of individual cells can be correlated with their individual layer.

Magnification: The variation in amount of retinal area represented per unit area of V1 from the fovea to the peripheral vision. Even though the fovea takes up an inordinate percentage of V1 compared with the rest of the visual field, the scale of the cellular organization remains constant. Thus the image from the fovea is, in effect, magnified before processing.

Receptive field: The area in the visual field that evokes a response in a neuron. Receptive fields may respond to specific stimuli such as illuminated bars or edges with particular directions of motion, etc.

Stereopsis: The determination of distance to objects based on relative displacement on the two retinas because of parallax.

Topographic mapping: The one-to-one correspondence between location on the retina and location within a structure in the brain. Topographic mapping further implies that contiguous areas on the retina map to contiguous areas in the particular brain structure.

References

Belliveau JH, Kwong KK, et al. 1992. Magnetic resonance imaging mapping of brain function: Human visual cortex. Invest Radiol 27(suppl 2):S59.

Cohen MS, Bookheimer SY. 1994. Localization of brain function using magnetic resonance imaging. Trends Neurosci 17(7):268.

Daniel PM, Whitteridge D. 1961. The representation of the visual field on the cerebral cortex in monkeys. J Physiol 159:203.

Edelman GM. 1978. Group selection and phasic reentrant signalling: A theory of higher brain function. In GM Edelman and VB Mountcastle (eds), The Mindful Brain, pp 51–100, Cambridge, MIT Press.

Ennes HE. 1981. NTSC color fundamentals. In Television Broadcasting: Equipment, Systems, and Operating Fundamentals. Indianapolis, Howard W. Sams & Co.

Felleman DJ, V Essen DC. 1991. Distributed hierarchical processing in the primate cerebral cortex. Cerebral Cortex 1(1):1.

Gross CG, Sergen J. 1992. Face recognition. Curr Opin Neurobiol 2(2):156.

Hebb DO. 1961. The Organization of Behavior. New York, Wiley.

Helmholtz H. 1889, Popular Scientific Lectures. London, Longmans.

Hering E. 1864. Outlines of a Theory of Light Sense. Cambridge, Harvard University Press.

Hubel DH. 1995. Eye, Brain, and Vision. New York, Scientific American Library.

Land EH, McCann JJ. 1971. Lightness and retinex theory. J. Opt Soc Am 61:1.

Livingstone MS, Hubel DH. 1984. Anatomy and physiology of a color system in the primate visual cortex. J Neurosci 4:309.

Marks WB, Dobelle WH, MacNichol EF. 1964. Visual pigments of single primate cones. Science 143:1181.

Marr D. 1982. Vision. San Francisco, WH Freeman.

Mountcastle VB. 1957. Modality and topographic properties of single neurons of cat's somatic sensory cortex. J Neurophysiol 20(3):408.

Poggio GF, Talbot WH. 1981. Mechanisms of static and dynamic stereopsis in foveal cortex of the rhesus monkey. J Physiol 315:469.

Wald G. 1974. Proceedings: Visual pigments and photoreceptors — Review and outlook. Exp Eye Res 18(3):333.

Weiskrantz L. 1990. The Ferrier Lecture: Outlooks for blindsight: explicit methodologies for implicit processors. Proc R Soc Lond B239:247.

Wheatstone SC. 1838. Contribution to the physiology of vision. Philosoph Trans R Soc Lond

Young T. 1802. The Bakerian Lecture: On the theory of lights and colours. Philosoph Trans R Soc Lond 92:12.

Zeki S. 1992. The visual image in mind and brain. Sci Am, Sept. 1992, p. 69.

Zeki S, Watson JD, Lueck CJ, et al. 1991. A direct demonstration of functional specialization in human visual cortex. J Neurosci 11(3):641.

Further Reading

An excellent introductory text about the visual system is *Eye, Brain, and Vision,* by Nobel laureate, David H. Hubel (1995, Scientific American Library, New York). A more recent general text with a thorough treatment of color vision, as well as the higher cortical centers, is *A Vision of the Brain,* by Semir Zeki (1993, Blackwell Scientific Publications, Oxford).

Other useful texts with greater detail about the nervous system are *From Neuron to Brain,* by Nicholls, Martin, Wallace, and Kuffler (3rd ed., 1992, Sinauer Assoc., Sunderand Mass.), *The Synaptic Organization of the Brain,* by Shepherd (4th ed., 1998, Oxford Press, New York), and *Fundamental Neuroanatomy,* by Nauta and Feirtag (1986, Freeman, New York).

A classic text that laid the foundation of computer vision by *Vision,* by David Marr (1982, Freeman, New York). Other texts dealing with the mathematics of image processing and image analysis are *Digital Image Processing,* by Pratt (1991, Wiley, New York), and *Digital Imaging Processing and Computer Vision,* by Schalkoff (1989, Wiley, New York).