Chapter 2
Neurological Substrate of the HVS

Considerable information may be gleaned from the vast neuroscientific literature regarding the functionality (and limitations) of the human visual system (HVS). It is often possible to qualitatively predict observed psychophysical results by studying the underlying visual "hardware." For example, visual spatial acuity may be roughly estimated from knowledge of the distribution of retinal photoreceptors. Other characteristics of human vision may also be estimated from the neural organization of deeper brain structures.

Neurophysiological and psychophysical literature on the human visual system suggests the field of view is inspected *minutatim* through brief fixations over small regions of interest. This allows perception of detail through the fovea. Central foveal vision subtends $1^\circ - 5^\circ$ (visual angle) allowing fine scrutiny of only a small portion of the entire visual field, for example only 3% of the size of a large (21 in.) computer monitor (seen at $\sim$60 cm viewing distance). Approximately 90% of viewing time is spent in fixations. When visual attention is directed to a new area, fast eye movements (saccades) reposition the fovea. The dynamics of visual attention probably evolved in harmony with (or perhaps in response to) the perceptual limitations imposed by the neurological substrate of the visual system.

The brain is composed of numerous regions classified by their function (Zeki 1993). A simplified representation of brain regions is shown in Fig. 2.1, with lobe designations stylized in Fig. 2.2. The human visual system is functionally described by the connections between retinal and brain regions, known as visual pathways. Pathways joining multiple brain areas involved in common visual functions are referred to as streams. Figure 2.1 highlights regions and pathways relevant to selective visual attention. For clarity, many connections are omitted. Of particular importance to dynamic visual perception and eye movements are the following neural regions, summarized in terms of relevance to attention.

- Superior colliculus (SC): involved in programming eye movements and contributes to eye movement target selection for both saccades and smooth pursuits (possibly

A.T. Duchowski, *Eye Tracking Methodology*, DOI 10.1007/978-3-319-57883-5_2

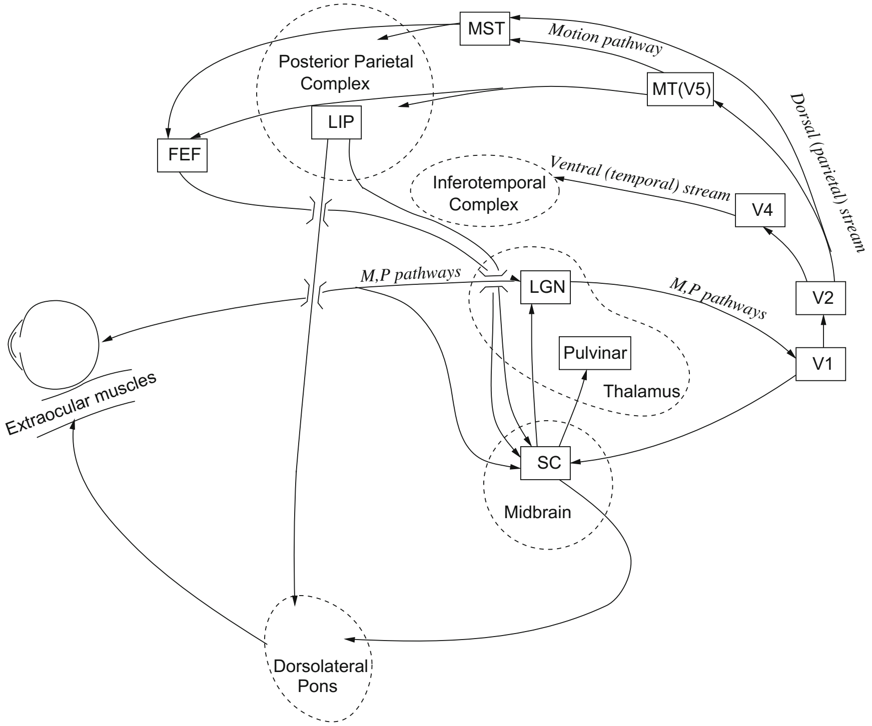

Fig. 2.1 A simplified view of the brain and the visual pathways relevant to eye movements and attention

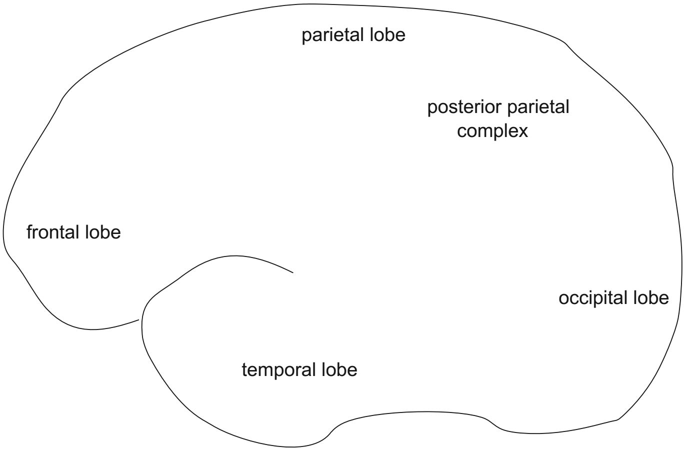

Fig. 2.2 Stylized classification of cortical lobes

in concert with the frontal eye fields (FEF) and area lateral intraparietal (LIP)); also remaps auditory space into visual coordinates (presumably for target foveation); with input of motion signals from area MT (see below), the SC is involved in pursuit target selection as well as saccade target selection.
- Area V1 (primary visual cortex): detection of range of stimuli, e.g., principally orientation selection and possibly to a lesser extent color; cellular blob regions (double-opponent color cells) respond to color variations and project to areas V2 and V4 (Livingstone and Hubel 1988).
- Areas V2, V3, V3A, V4, MT: form, color, and motion processing.
- Area V5/MT (middle temporal) and MST (middle superior temporal): furnish large projections to Pons; hence possibly involved in smooth pursuit movements; involved in motion processing: area MT also projects to the colliculus, providing it with motion signals from the entire visual field.
- Area lateral intra parietal (LIP): contains receptive fields that are corrected (reset) before execution of saccadic eye movements.
- Posterior parietal complex (PPC): involved in fixations.

Connections made to these areas from area V1 can be generally divided into two streams: the dorsal and ventral streams. Loosely, their functional description can be summarized as

- Dorsal stream: sensorimotor (motion, location) processing (e.g., the attentional "where")
- Ventral stream: cognitive processing (e.g., the attentional "what")

In general attentional terms, the three main neural regions implicated in eye movement programming and their functions are Palmer (1999):

- Posterior parietal complex: disengages attention,
- SC: relocates attention,
- Pulvinar: engages, or enhances, attention.

In a very simplified view of the brain, it is possible to identify the neural mechanisms involved in visual attention and responsible for the generation of eye movements. First, by examining the structure of the eye, it becomes clear why only the central or foveal region of vision can be perceived at high resolution. Second, signals from foveal and peripheral regions of the eye's retina can be roughly traced along pathways in the brain showing how the brain may process the visual scene. Third, regions in the brain can be identified which are thought to be involved in moving the eyes so that the scene can be examined piecemeal. In this simplified view of the brain, one can in a sense obtain a complete picture of an "attentional feedback loop," which creates the attentional cycles of disengaging attention, shifting of attention and (usually) the eyes, and for processing the region of interest currently being attended to, re-engaging attention and brain regions.

The neural substrate of the human visual system is examined in this chapter from the intuitive attentional perspective given above. The human neural hardware responsible for visual processing is presented in order roughly following the direction of light and hence information entering the brain. That is, the discussion is presented "front-to-back" starting with a description of the eye and ending with a summary of the visual cortex located at the back of the brain. Emphasis is placed on differentiating the processing capability of foveal and peripheral vision, i.e., the simplified "what" and "where" of visual attention, respectively. However, the reader must be cautioned against underestimating the complexity of the visual system as presented in this text. The apparent "what" and "where" dual pathways are most probably not independent functional channels. There is a good deal of interconnection and "crosstalk" between these and other related visual centers which deems the dichotomous analysis overly simplistic. Nevertheless, there is a great deal of valuable information to be found in the neurological literature as human vision is undoubtedly the most studied human sense.

2.1 The Eye

Often called "the world's worst camera," the eye, shown in Fig. 2.3, suffers from numerous optical imperfections, for example,

- Spherical aberrations: prismatic effect of peripheral parts of the lens
- Chromatic aberrations: shorter wavelengths (blue) refracted more than longer wavelengths (red)
- Curvature of field: a planar object gives rise to a curved image

However, the eye is also endowed with various mechanisms that reduce degradive effects, e.g.,

- To reduce spherical aberration, the iris acts as a stop, limiting peripheral entry of light rays,
- To overcome chromatic aberration, the eye is typically focused to produce sharp images of intermediate wavelengths,
- To match the effects of curvature of field, the retina is curved compensating for this effect.

The eye is schematically shown in Fig. 2.3.

2.2 The Retina

At the rear interior surface of the eye, the retina contains receptors sensitive to light (photoreceptors) which constitute the first stage of visual perception. Photoreceptors can effectively be thought of as "transducers" converting light energy to electrical impulses (neural signals). Neural signals originating at these receptors lead to deeper visual centers in the brain. Photoreceptors are functionally classified into rods and cones. Rods are sensitive to dim and achromatic light (night vision), whereas cones

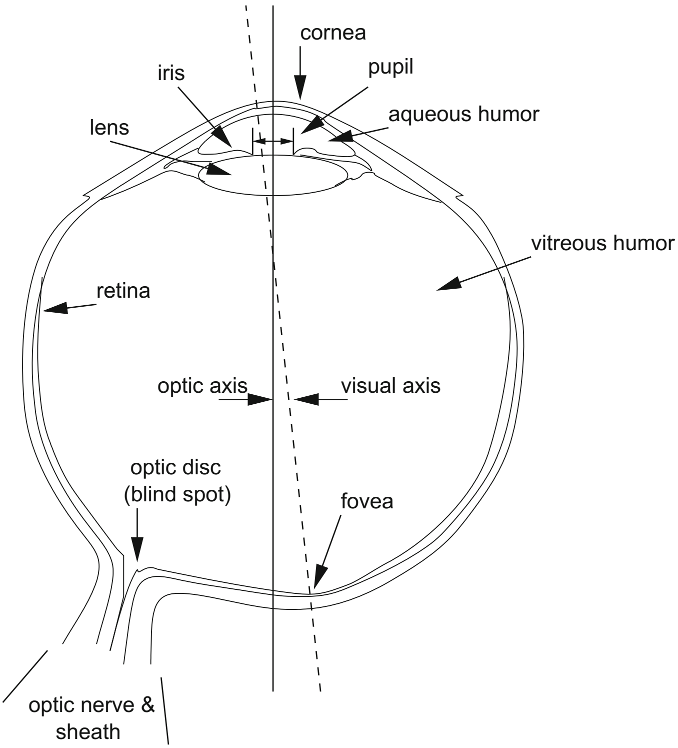

Fig. 2.3 The eye. Adapted from Visual Perception, 1st edition, by Cornsweet (1970) ©1970. Reprinted with permission of Wadsworth, a division of Thomson Learning: `www.thomsonrights.com`

respond to brighter chromatic light (daylight vision). The retina contains approximately 120 million rods and 7 million cones.

The retina is composed of multiple layers of different cell types (Valois and Valois 1988). Surprisingly, the "inverted" retina is constructed in such a way that photoreceptors are found at the bottom layer. This construction is somewhat counterintuitive inasmuch as rods and cones are farthest away from incoming light, buried beneath a layer of cells. The retina resembles a three-layer cell sandwich, with connection bundles between each layer. These connectional layers are called plexiform or synaptic layers. The retinogeniculate organization is schematically depicted in Fig. 2.4. The outermost layer (w.r.t. incoming light) is the outer nuclear layer which contains the photoreceptor (rod/cone) cells. The first connectional layer is the outer plexiform layer which houses connections between receptor and bipolar nuclei. The next outer layer of cells is the inner nuclear layer containing bipolar (amacrine, bipolar, horizontal) cells. The next plexiform layer is the inner plexiform layer where connections between inner nuclei cells and ganglion cells are formed. The top layer, or the ganglion layer, is composed of ganglion cells.

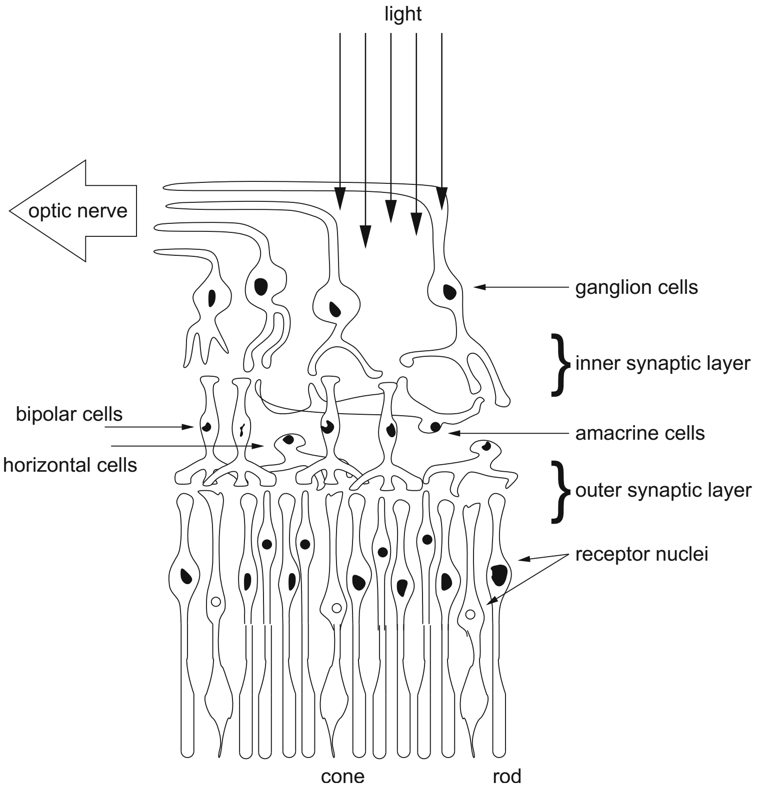

Fig. 2.4 Schematic diagram of the neural interconnections among receptors and bipolar, ganglion, horizontal, and amacrine cells. Adapted from Dowling and Boycott (1966) with permission ©1966 The Royal Society (London)

The fovea's photoreceptors are special types of neurons, the nervous system's basic elements (see Fig. 2.5). Retinal rods and cones are specific types of dendrites. In general, individual neurons can connect to as many as 10,000 other neurons. Comprised of such interconnected building blocks, as a whole, the nervous system behaves as a large neural circuit. Certain neurons (e.g., ganglion cells) resemble a "digital gate," sending a signal (firing) when the cell's activation level exceeds a threshold. The myelin sheath is an axonal cover providing insulation which speeds up conduction of impulses. Unmyelinated axons of the ganglion cells converge to the optic disk (an opaque myelin sheath would block light). Axons are myelinated at the optic disk, and connect to the Lateral Geniculate Nuclei (LGN) and the Superior Colliculus (SC).

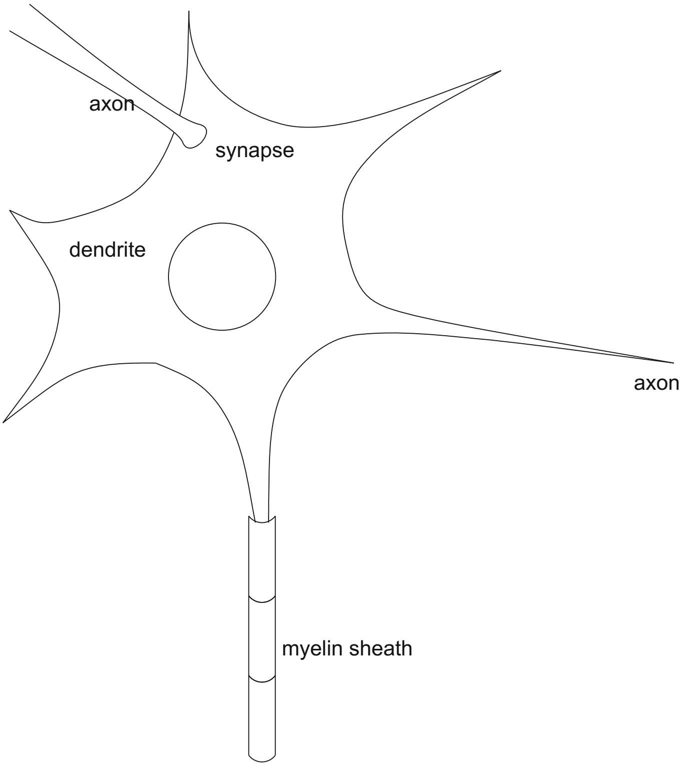

Fig. 2.5 Schematic of the neuron. From Brain, Mind, and Behavior by Floyd E. Bloom and Arlyne Lazerson ©1985, 1988, 2001 by Educational Broadcasting Corporation. Used with the permission of Worth Publishers

2.2.1 The Outer Layer

Rods and cones of the outer retinal layer respond to incoming light. A simplified account of the function of these cells is that rods provide monochromatic scotopic (night) vision, and cones provide trichromatic photopic (day) vision. Both types of cells are partially sensitive to mesopic (twilight) light levels.

2.2.2 The Inner Nuclear Layer

Outer receptor cells are laterally connected to the horizontal cells. In the fovea, each horizontal cell is connected to about 6 cones, and in the periphery to about 30–40 cones. Centrally, the cone bipolar cells contact one cone directly, and several cones indirectly through horizontal or receptor–receptor coupling. Peripherally, cone bipolar cells directly contact several cones. The number of receptors increases eccentrically. The rod bipolar cells contact a considerably larger number of receptors than cone bipolars. There are two main types of bipolar cells: ones that depolarize to increments of light (+), and others that depolarize to decrements of light (−). The signal profile (cross-section) of bipolar receptive fields is a "Mexican Hat," or center-surround, with an on-center, or off-center signature.

2.2.3 The Ganglion Layer

In a naive view of the human visual system, it is possible to inaccurately think of the retina (and thus the HVS as a whole) acting in a manner similar to that of a camera. Although it is true that light enters the eye and is projected through the lens onto the retina, the camera analogy is only accurate up to this point. In the retina, ganglion cells form an "active contrast-enhancing system," not a camera like plate. Centrally, ganglion cells directly contact one bipolar. Peripherally, ganglion cells directly contact several bipolars. Thus the retinal "camera" is not composed of individual "pixels." Rather, unlike isolated pixels, the retinal photoreceptors (rods and cones in the base layer) form rich interconnections beyond the retinal outer layer. With about 120 million rods and cones and only about 1 million ganglion cells eventually innervating at the LGN, there is considerable convergence of photoreceptor output. That is, the signals of many (on the order of about 100) photoreceptors are combined to produce one type of signal. This interconnecting arrangement is described in terms of receptive fields, and this arrangement functions quite differently from a camera.

Ganglion cells are distinguished by their morphological and functional characteristics. Morphologically, there are two types of ganglion cells, the α and β cells. Approximately 10% of retinal ganglion cells are α cells possessing large cell bodies and dendrites, and about 80% are β cells with small bodies and dendrites (Lund et al. 1995). The α cells project to the magnocellular (M-) layers of the LGN and the β cells project to the parvocellular (P-) layers. A third channel of input relays through narrow, cell-sparse laminae between the main M- and P-layers of the LGN. Its origin in the retina is not yet known. Functionally, ganglion cells fall into three classes, the X, Y, and W cells (Valois and Valois 1988; Kaplan 1991). X cells respond to sustained stimulus, location and fine detail, and innervate along both M- and P-projections. Y cells innervate only along the M-projection, and respond to transient stimulus, coarse features and motion. W cells respond to coarse features, and motion, and project to the Superior Colliculus.

The receptive fields of ganglion cells are similar to those of bipolar cells (center-surround, on-center, off-center). Center-on and center-off receptive fields are depicted in Fig. 2.6. Plus signs (+) denote illumination stimulus, minus signs (−) denote lack of stimulus. The vertical bars below each receptive field depict the firing response of the receptive field. This signal characteristic (series of "ticks") is usually obtained by inserting an electrode into the brain. The signal profile of receptive fields resembles the "Mexican hat" operator, often used in image processing.

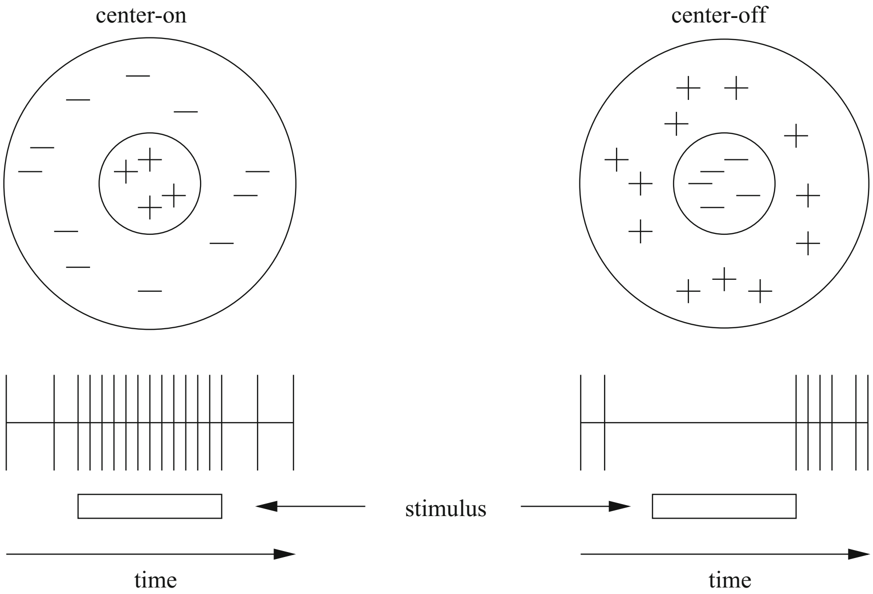

Fig. 2.6 Schematic of receptive fields

2.3 The Optic Tract and M/P Visual Channels

Some (but not all) neural signals are transmitted from the retina to the occipital (visual) cortex through the optic tract, crossing in the optic chiasm, making connections to the LGN along the way. The physiology of the optic tract is often described functionally in terms of visual pathways, with reference to specific cells (e.g., ganglion cells). It is interesting to note the decussation (crossing) of the fibers from the nasal half of the retina at the optic chiasm, i.e., nasal retinal signals cross, temporal signals do not.

M and P ganglion cells in the retina connect to M and P channels, respectively. Along the optic pathways, the superior colliculus and the lateral geniculate nucleus are of particular importance. The SC is involved in programming eye movements and also remaps auditory space into visual coordinates. As shown in Fig. 2.1, some neural signals along the optic tract project to the SC. The SC is thought to be responsible for directing the eyes to a new region of interest for subsequent detailed visual inspection. Like other regions in the thalamus serving similar functions, the LGN is a crossover point, or relay station, for α and β ganglion cells. The physiological organization of the LGN, with respect to nervations of these cells, produces a visual field topography of great clinical importance. Here, the magnocellular and the parvocellular ganglionic projections are clearly visible (under microscope), forming junctions within two distinct layers of the LGN, correspondingly termed the M- and P-layers. Thalamic

Table 2.1 Functional characteristics of ganglionic projections

Characteristics	Magnocellular	Parvocellular
Ganglion size	Large	Small
Transmission time	Fast	Slow
Receptive fields	Large	Small
Sensitivity to small objects	Poor	Good
Sensitivity to change in light levels	Large	Small
Sensitivity to contrast	Low	High
Sensitivity to motion	High	Low
Color discrimination	No	Yes

axons from the M- and P-layers of the LGN terminate in area V1 (the primary visual center) of the striate cortex.

The functional characteristics of ganglionic projections to the LGN and the corresponding magno- and parvocellular pathways are summarized in Table 2.1. The parvocellular pathway in general responds to signals possessing the following attributes: high contrast (the parvocellular pathway is less sensitive to luminance), chromaticity, low temporal frequency, and high spatial frequency (due to the small receptive fields). Conversely, the magnocellular pathway can be characterized by sensitivity to the following signals: low contrast (the magnocellular pathway is more sensitive to luminance), achromaticity, moderate-to-high temporal frequency (e.g., sudden onset stimuli), and low spatial frequency (due to the large receptive fields). Zeki (1993) suggests the existence of four functional pathways defined by the M and P channels: motion, dynamic form, color, and form (size and shape). It is thought that fibers reaching the superior colliculus represent retinal receptive fields in rod-rich peripheral zones, whereas the fibers reaching the LGN represent cone-rich areas of high acuity (Bloom and Lazerson 1988). It seems likely that, in a general sense, the M ganglion cells correspond to rods, mainly found in the periphery, and the P cells correspond to cones, which are chromatic cells concentrated mainly in the foveal region.

2.4 The Occipital Cortex and Beyond

Thalamic axons from the M- and P-layers of the LGN terminate mainly in the lower and upper halves (β, α divisions, respectively) of layer 4C in middle depth of area V1 (Lund et al. 1995). Cell receptive field size and contrast sensitivity signatures are distinctly different in the M- and P- inputs of the LGN, and vary continuously through the depth of layer 4C. Unlike the center-surround receptive fields of retinal ganglion and LGN cells, cortical cells respond to orientation-specific stimulus (Hubel 1988). Cortical cells are distinguished by two classes: simple and complex.

In area V1, the size of a simple cell's receptive field depends on its relative retinal position. The smallest fields are in and near the fovea, with sizes of about 1/4 × 1/4 degree. This is about the size of the smallest diameters of the smallest receptive field centers of retinal ganglion or LGN cells. In the far periphery, simple cell receptive field sizes are about 1× 1 degree. The relationship between small foveal receptive fields and large peripheral receptive fields is maintained about everywhere along the visual pathway.

Simple cells fire only when a line or edge of preferred orientation falls within a particular location of the cell's receptive field. Complex cells fire wherever such a stimulus falls into the cell's receptive field (Lund et al. 1995). The optimum stimulus width for either cell type is, in the fovea, about two minutes of arc. The resolving power (acuity) of both cell types is the same.

About 10–20% of complex cells in the upper layers of the striate cortex show marked directional selectivity (Hubel 1988). Directional Selectivity (DS) refers to the cell's response to a particular direction of movement. Cortical Directional Selectivity (CDS) contributes to motion perception and to the control of eye movements (Grzywacz and Norcia 1995). CDS cells establish a motion pathway from V1 projecting to MT and V2 (which also projects to MT) and to MST. In contrast, there is no evidence that retinal directional selectivity (RDS) contributes to motion perception. RDS contributes to oculomotor responses (Grzywacz et al. 1995). In vertebrates, it is involved in optokinetic nystagmus, a type of eye movement discussed in Chap. 4.

2.4.1 Motion-Sensitive Single-Cell Physiology

There are two somewhat counterintuitive implications of the visual system's motion-sensitive single-cell organization for perception. First, due to motion-sensitive cells, eye movements are never perfectly still but make constant tiny movements called *microsaccades* (Hubel 1988). The counterintuitive fact regarding eye movements is that if an image were artificially stabilized on the retina, vision would fade away within about a second and the scene would become blank. Second, due to the response characteristics of single (cortical) cells, the camera like "retinal buffer" representation of natural images is much more abstract than intuition suggests. An object in the visual field stimulates only a tiny fraction of the cells on whose receptive field it falls (Hubel 1988). Perception of the object depends mostly on the response of (orientation-specific) cells to the object's borders. For example, the homogeneously shaded interior of an arbitrary form (e.g., a kidney bean) does not stimulate cells of the visual system. Awareness of the interior shade or hue depends on only cells sensitive to the borders of the object. In Hubel (1988) words, "...our perception of the interior as black, white, gray, or green has nothing to do with cells whose fields are in the interior—hard as that may be to swallow.... What happens at the borders is the only information you need to know: the interior is boring."

2.5 Summary and Further Reading

This chapter presented a simplified view of the brain with emphasis on regions and structures of the brain responsible for attentional and visual processing, including those regions implicated in eye movement generation. Starting with the structure of the eye, the most salient observation is the structure of the retina which clearly shows the limited scope of the high resolution fovea. The division between foveo–peripheral vision is maintained along the visual pathways and can be clearly seen under microscope in the LGN. Of particular relevance to attention and eye movements is the physiological and functional duality of the magno- and parvocellular pathways and of their apparent mapping to their attentional "what" and "where" classification. Although this characterization of the M- and P-pathways is admittedly overly simplistic, it provides an intuitive functional distinction between foveal and peripheral vision.

An interesting visual example of foveo–peripheral processing is shown in Fig. 2.7. To notice the curious difference between foveal and peripheral processing, foveate one corner of the image in Fig. 2.7 and, without moving your eyes, shift your attention to the opposing corner of the image. Interestingly, you should perceive white dots at the line crossings in the foveal region, but black dots should appear at the line crossings in the periphery.

Examining regions in the brain along the visual pathways, one can obtain insight into how the brain processes visual information. The notion that attention may be driven by certain visual features (e.g., edges) is supported to an extent by the identi-

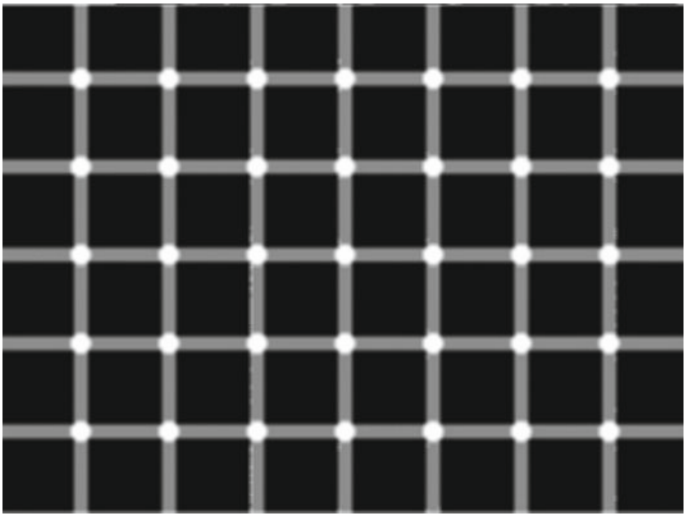

Fig. 2.7 Foveo–peripheral illusion: scintillation effect produced by a variation of the standard Hermann grid illusion (attributed to L. Hermann (1870)), first discovered by Elke Lingelbach (at home). Adapted from Ninio and Stevens ©2000, Pion, London

fication of neural regions which respond to these features. How certain features are perceived, particularly within and beyond the fovea, is the topic covered in the next chapter.

For an excellent review of physiological optics and visual perception in general, see Hendee and Wells (1997). For an introduction to neuroscience, see Hubel (1988) very readable text. For a more recent description of the brain with an emphasis on color vision, see Zeki (1993). Apart from these texts on vision, several "handbooks" have also been assembled describing current knowledge of the brain. Arbib (1995) handbook is one such example. It is an excellent source summarizing current knowledge of the brain, although it is somewhat difficult to read and to navigate through.[1] Another such well-organized but rather large text is Gazzaniga (2000).

---

[1] A new edition of Arbib (1995)'s book has recently been announced.