

# 1 What is known about vision experimentally

Vision is one of the most studied brain functions, in some sense, offering a window to study the brain. There is thus a vast knowledge about the physiology and anatomy of the brain responsible for vision, as well as about the visual behavior particularly in human vision. A good book for beginners to learn such knowledge is “Foundations of vision”<sup>128</sup> by Brain A Wandell, and readers will find it easy to read whether or not they are from the life science background. Other very useful books are: “Visual perception, physiology, psychology and ecology” by Bruce, Green, and Georgeson,<sup>13</sup> and “Vision Science, photons to phenomenology” by Palmer.<sup>96</sup> The book “Theoretical Neuroscience” by Dayan and Abbott<sup>23</sup> also provides a good introduction to early visual system and its receptive fields to the modellers. Meanwhile, here I give a brief review of the parts of these knowledge most relevant to the topics in this book. Most of the reviewed are about the human or primate visual system. Most materials presented in this section are the results of my paraphrasing the general knowledge in the vision science community, and hence I often omit the detailed references which can be obtained from typical textbooks such as the ones above, and from the two volumn book “The Visual Neurosciences” edited by Chalupa and Werner.<sup>21</sup>

## 1.1 Neurons, neural circuits, cortical areas, and the brain

Neurons are cells in the nervous system that receive, process, and transmit information. There are billions of neurons in the human brain, each is typically composed of dendrites, axons and a soma or cell body. Dendrites receive inputs from other neurons or from the external sensory world through some signal transduction process. Axons send output signals to other neurons or effectors such as muscle fibers. Typically, the output signals are in the form of electrical pulses or spikes called action potentials, each is about 1 millisecond (ms) in duration and dozens of millivolts in amplitude. Through synapses, which are contacts between neurons, action potentials cause electric current to flow across the membrane of the target neuron and change the target neuron’s membrane potential. The electric potentials within a neuron determine the state of a neuron and its production of action potentials. Action potentials are near identical to each other, hence, information are conveyed by their timing and rates, i.e., when they are fired or how many of them per unit time, rather than their individual voltage profiles. They can propagate long distances along axons without appreciable decays before reaching their destination neurons, and so are adequate for communication between neurons far apart from each other. Sometimes, very nearby neurons can also influence each other’s states without action potentials.

A simple model<sup>39</sup> of a neuron is as follows: a neuron’s internal or membrane potential is modelled by a single variable  $u$ , which can change in time  $t$  by input current  $I$  as  $du/dt = -u/\tau + I$ , where  $\tau$  is the membrane time constant to model the neuron as a leaky integrator or input current; the rate of the action potentials can be viewed as the output of the neuron; the output can be modelled as a nonlinear function  $g(u) \geq 0$  of the membrane potential  $u$ , such that this function  $g(u)$  is monotonously increasing with  $u$ , is zero for small  $u$ , and saturating for  $u \rightarrow \infty$ . The effect of this output  $g(u)$  is to contribute to the input current to the target neuron by an amount  $w \cdot g(u)$ , where  $w$  models the strength of the synaptic connection.

Each neuron has synaptic connections with hundreds or thousands of other neurons, forming neural circuits for computation. There are micro-circuits between nearby neurons, and macro-circuits between neural groups. Neurons with similar functional properties are aggregated together, and a cortical area, such as one of the visual cortical areas in Fig. (1), is defined by these locally connected and functionally similar groups of neurons. Nearby neurons are more likely con-



does some information processing within itself while receiving signals from, and sending signals to, other areas. Each neuron can be seen as an information processor transforming its input signals to output signals. It is typically composed of dendrites which receive input signals, axons which send output signals to other neurons or muscle cells, and the cell body. Often, the output signals are in the form of electrical pulses or spikes called action potentials, each is about 1 millisecond (ms) in duration and dozens of millivolts in amplitude. Action potentials are near identical to each other, hence, information are conveyed by their timing and rates, i.e., when the spikes are fired or how many of them per unit time, rather than their individual voltage profiles. They can propagate long distances along axons without appreciable decays before reaching their destination neurons, and so are ideal for communication between neurons far apart from each other, particular between different brain areas. Physically nearby cortical areas are more likely connected by the axons, as expected from the design to minimize the brain volume occupied by the inter-area neural axons to transmit the signals. Within a cortical area, some neurons which are physically near each other can also communicate without action potentials. Note from Fig. (1A) that about half of the brain areas are involved with vision. Most brain regions are denoted by their abbreviated names in Fig. (1). For instance, V1 denotes visual area 1, the primary visual cortex and the largest visual area containing detailed representation of the visual input; V2 for visual area 2 which receives most of its inputs from V1; LGN for lateral geniculate nucleus which is often viewed as the relay station between the retina and V1 by our ignorance; FEF for frontal eye field, SC for superior colliculus, and both FEF and SC control eye movements. IT for inferotemporal cortex, whose neurons respond to complex spatial shapes in visual inputs; MT for middle temporal area whose neurons are particularly sensitive to visual motion; LIP for lateral intra-parietal area, implicated for decision making for eye movements. The lower case letters ending some of the abbreviations often denote spatial locations of the cortical areas, e.g., v for ventral, d for dorsal.

The term visual pathway implies that there is a hierarchy of levels for information processing, starting from the retina, as shown schematically in Fig (1B). Information processing progresses from lower stages, starting at retina (and excluding the SC and gaze control stages in the pink shaded area), to higher stages, ending at FEF within this figure. Each neuron typically responds to, or is excited by, visual inputs in a limited extent of the visual space called its receptive field. The receptive field is small for retinal neurons, with a diameter only 0.06 degree in visual angle near the center of vision,<sup>110</sup> too small to cover most recognizable visual objects, e.g., an apple, in a typical scene. As one ascends along the visual hierarchy, the neural receptive field gets progressively larger, with a diameter of (in order of magnitudes) 10 degree in visual angle in V4, and 20-50 degrees in IT,<sup>102</sup> making it possible to hope that a single neuron in higher visual areas can signal the recognition of a visual object, e.g., one's grandmother. In the early stages such as the retina and V1, the receptive fields are relatively invariant to the animal's state of arousal. They become increasingly variable in the later stages, for instance, the sizes of the receptive fields depend on the animal's attention and on the complexity of the visual scenes.<sup>85</sup>

The connections between stages or brain regions in Fig (1B) symbolize the existence of neural connections between the regions. Most of these connections are non-directional, indicating that the connections are reciprocal or that each of the two areas connected receive signals from the other. This figure shows not only the flow of sensory information through various processing stages in the hierarchy, but also that of information flow towards visually induced action of eye movements. It also reflects the view shared by many others (e.g., Findlay and Gilchrist 2003) that understanding the motor actions associated with vision is very important to understanding the sensory processing. After all, the main purpose of recognizing and localizing objects in the scene

is to act on them; meanwhile, actions, such as directing the gaze to conspicuous locations in the scene, in turn facilitate sensing and sensory information processing. In this light, it is noteworthy that signals from as early as the retina and V1 in this hierarchy already influence the motor outputs of vision.

Physiologically and anatomically, much more is known about the early visual stages, in particular the retina, LGN, and V1, than higher visual areas. This is partly because it is often easier to access these early stages and is easier to determine how neural responses are related to the visual inputs. Behaviorally, one can probe how sensitive an animal is to various simple or complex visual inputs, ranging from the image of a simple small bar to that of an emotionally looking face. One can also measure how quickly and easily visual objects are localized or identified, e.g., in finding a tomato among many apples. Often, behavioral findings using simple visual stimuli could be linked with physiological and anatomical findings about the early visual stages. However, our relative ignorance of the higher visual areas means that our knowledge of more complex visual behavior is much less associated with the neural bases. In particular, the hierarchy of visual cortical areas shown in Fig (1B) is inferred mostly from anatomical evidences. They may suggest but not precisely determine the hierarchy of information processing, and different anatomical or physiological evidences<sup>11</sup> can give different interpretation as to which level in the hierarchy a particular visual cortical area should be. As understanding vision necessarily means understanding both the neural and behavioral aspects, theoretical and modeling studies on visual functions are much easier for early visual processes. This book reflects this by focusing on the retina and V1 and their associated visual behavior.

### 1.3 Retina

The retina is the first stage in the visual pathway. The three dimensional visual scene is imaged on the retina, where the lights in the images are absorbed by the photoreceptors at the image plane, see Fig. (2). In primate retina, there are about  $5 \times 10^6$  cones responsible for the day time color vision,  $10^8$  rods, which are mainly functional in dim light.<sup>128</sup> Each photoreceptor absorbs the local light in the image to electrical response signals. These signals are transformed through several intermediate cell types called bipolar cells, horizontal cells, and amacrine cells, before they are finally received by about  $10^6$  retinal ganglion cells, the output neurons from the retina. By firing voltage impulses, each about 1 millisecond (ms) in duration and dozens of milli-volts in amplitude, at up to about 100 spikes per second for each neuron, the  $10^6$  ganglion cells send the visual signals via their axons, bundled together into the optic nerve, on to the brain. Note that the blood vessels in the eye ball are also imaged onto the back of the retina together with the visual scene. Nevertheless, we seldom see them since they are static in the images. Human vision is insensitive to static or non-changing inputs. Voluntary and involuntary eye movements, many of them are ever-present small jitters of our eyes that we are unaware of, keep us not blind to the part of the visual world which is motionless.

#### 1.3.1 Receptive fields of the retinal ganglion cells

If one quantifies the response of a retinal ganglion cells by the firing rate, i.e., the number of neural spikes per seconds, and the visual input at any image location by the contrast, i.e., the ratio between input intensity at this location and the mean input intensity, then for most ganglion cells, the response is approximately a linear function of the input.<sup>110</sup> This means, if such an input at location

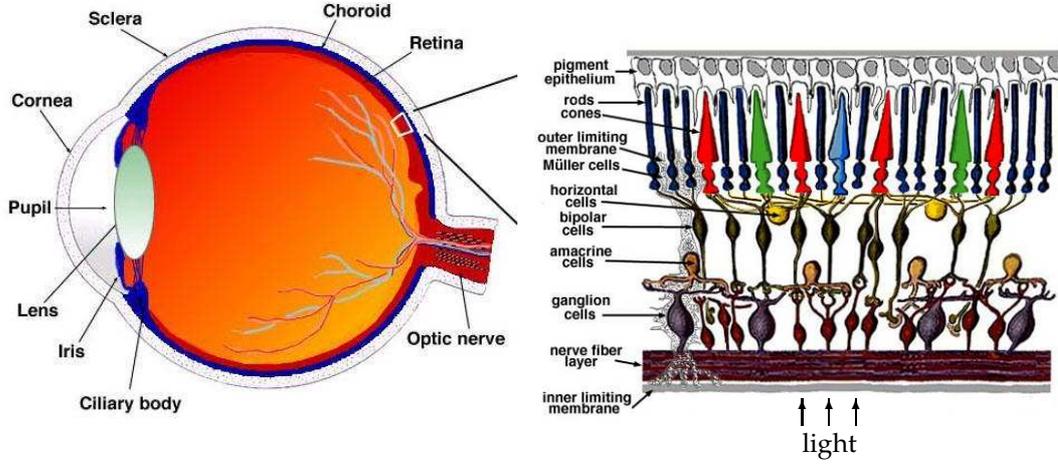


Figure 2: The schematic illustration of the retina and its neurons, adapted from figures in “Simple Anatomy of the retina” from <http://www.webvision.med.utah.edu/sretina.html>. In the left part, light enters the eye and the retinal neural responses are transmitted by the optic nerve to the rest of the brain. The right half is a zoomed up view of a patch of the retina on the left, with imaging light entering from the bottom, passing through ganglion and other cell layers before hitting the rods and cones.

$x$  and time  $t$  is  $S(x, t)$ , and the firing rate of a neuron at time  $t$  is  $O(t)$ , then

$$O(t) = \int dx dt' K(x, t - t') S(x, t') + \text{spontaneous firing rate} \quad (1)$$

where  $K(x, t - t')$  denotes the spatiotemporal filter, or the spatio-temporal receptive field, that transforms input  $S$  to output  $O$ . In many physiological experiments, the input  $S(x, t)$  is set to zero before a time designated as  $t = 0$ , the stimulus onset time, and then stays unchanged  $S(x, t) = S_x(x)$  for a sufficiently long time. To this input, the response  $O(t)$  at large enough  $t >$  is called the sustained response of the neuron and can reveal the spatial shape of the receptive field  $K$  as  $K_x(x) \equiv \int K(x, t) dt$ , when the temporal filter is integrated out, thus

$$O(t \rightarrow \infty) = \int dx K_x(x) S_x(x) \quad (2)$$

The filter value  $K_x(x)$  is non-zero for a limited spatial range of  $x$ , typically only a fraction of a degree, and this range is then the range of the receptive field of the neuron. The center of this receptive field varies from neuron to neuron, such that the whole population of the retinal ganglion cells can adequately sample the whole visual field. Let the center of a receptive field be  $x = 0$ , it is often found that  $K_x(x)$  has a shape which can be modelled by a difference of two gaussians, which in two dimensional space  $x$  is

$$K_x(x) = \frac{w_c}{\sigma_c^2} \exp[-x^2/(2\sigma_c^2)] - \frac{w_s}{\sigma_s^2} \exp[-x^2/(2\sigma_s^2)] \quad (3)$$

where the first and the second terms denote the two gaussian shapes respectively, with  $w_c$  and  $w_s$  indicating their strengths, and  $\sigma_c$  and  $\sigma_s$  their spatial extents, as illustrated in Fig. (3). Typically, the  $\sigma_c < \sigma_s$  and  $w_c \approx w_s$  such that  $K_x(x)$  has a spatially opponent shape. In the example in Fig.

(3),  $w_c$  and  $w_s$  are both positive, the neuron will increase its output  $O$  by a bright spot near the center of the receptive field but decrease its output when this bright spot is farther from the center, and the optimal visual input to excite this cell would be a bright center disk surrounded by a dark ring. Hence, such a receptive field is called a center-surround receptive field. If both  $w_c$  and  $w_s$  are negative, then the optimal stimulus would be a dark central spot surrounded by a bright ring, and a bright central spot in a dark ring would decrease the neural response. The two kinds of receptive fields, or neural types, corresponding to positive or negative values for  $w_c$  and  $w_s$ , are called on-center or off-center cells respectively. As the firing rates are never negative, to make room for firing rate decrease, the spontaneous firing rates in response to no inputs, or spatially uniform inputs, are high enough, around 50 and 20 spikes/second for the majority (i.e., the X or P cells, see later) of ganglion cells in the cat and monkeys respectively.<sup>120,121</sup>

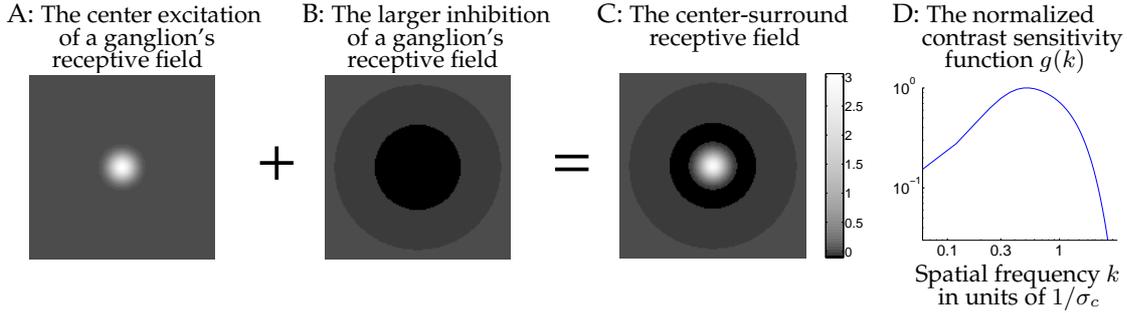


Figure 3: A-C: The receptive field shape of a retinal ganglion cell is modelled as a difference of two gaussians shown in A and B (as an inhibition), giving a center-surround shape of the receptive field in C. In each plot, the value of the receptive field  $K_x(x)$ , or its components, is visualized by the gray scale at image location  $x$ , with bright and dark pixels for excitation and inhibition, and the gray level near the image corners for zero  $K_x(x)$ . Parameters used are:  $\sigma_s/\sigma_c = 5$ ,  $w_c/w_s = 1.1$ . D: the normalized contrast sensitivity  $g(k)$  vs spatial frequency  $k$  in the units of  $1/\sigma_c$ , for the receptive field in C. Also see Fig. (??).

An input

$$S(x, t) = S_x(x)\delta(t) \quad (4)$$

(where  $\delta(t \neq 0) = 0$ , and  $\int dt\delta(t) = 1$ ) which appears only momentarily at time  $t = 0$  can reveal the impulse response

$$O(t) = \int dx K(x, t) S_x(x) \quad (5)$$

of the filter  $K(x, t)$ , or how  $K(x, t)$  changes with  $t$ . The impulse response to an optimal center-surround spatial pattern  $S_x(x)$  is typically an increase followed by a decrease of responses  $O(t)$  from the spontaneous response level, like that shown in Fig. (??D) by approximating the impulse response function

$$O(t) = e^{-\alpha t} [(\alpha t)^5/5! - (\alpha t)^7/7!], \quad (6)$$

with  $\alpha = 70/\text{second}$ . If a stimulus

$$S(x, t) = S_x(x)H(t) \quad (7)$$

$$\text{(where } H(t) \text{ is a step function with } H(t) = \begin{cases} 1, & t \geq 0 \\ 0, & \text{otherwise} \end{cases} \quad (8)$$

is the onset of a spatial pattern  $S_x(x)$  which stays on, the response  $O(t)$  typically has a transient component near the onset time resembling the impulse response, and a sustained response for  $t \rightarrow \infty$ .

The two best known classes of the retinal ganglion cells in primates are called the parvocellular cells and the magnocellular cells, or P and M cells for short. The P cells are about of the order 10 times as numerous, and have smaller receptive fields and longer impulse responses compared to the M cells. Hence, the P cells can have a better spatial resolution while the M cells better temporal resolution.

### 1.3.2 Contrast sensitivity to the sinusoidal gratings

One can also investigate how a ganglion cell responds to a sinusoidal input pattern

$$S_x(x) = S_k \cos(kx + \phi) + \text{constant} \quad (9)$$

a grating of spatial frequency  $k/(2\pi)$  cycles/degree, with an amplitude  $S_k$  and phase  $\phi$ . If one decompose the spatial receptive field  $K_x(x)$  into Fourier waves by

$$K_x(x) \sim \int dk g(k) e^{ikx}, \quad (10)$$

then  $g(k)$  is the sensitivity or gain of the neuron to the sinusoidal wave of frequency  $k$ , and it is called the contrast sensitivity curve. For  $K_x(x)$  in equation (3),  $g(k)$  in two dimensional space  $k = (k_x, k_y)$  is,

$$g(k) \sim w_c \exp[-k^2 \sigma_c^2 / 2] - w_s \exp[-k^2 \sigma_s^2 / 2] \quad (11)$$

which is another difference of two gaussians. For  $|w_c| \geq |w_s|$ , we have  $|g(k)|$  slowly increasing with  $k$  until reaching a peak value at some frequency  $k_p$  before decreasing with  $k$ . Thus  $K_x(x)$  is a band pass filter which is insensitive to low spatial frequency signals or spatially smooth signals, or to high frequency signals which vary in a scale much finer than the scale  $\sigma_c$  and  $\sigma_s$  of the receptive field, but is most sensitive to spatial frequency on the order of  $k_p \sim 1/\sigma_c$ , or to spatial variations on a scale comparable to the size of the center of the receptive field. See Fig. (3ACD).

Equation (2) implies that the sustained response level of the cell should be  $O(t \rightarrow \infty) \sim g(k) \cos(\phi)$ . Hence, by using spatial grating with zero phase  $\phi = 0$ , one can obtain  $g(k)$ , from which one can quite easily construct the shape of the spatial filter  $K_x(x)$ . Experiments often use a drifting grating

$$S(x, t) \propto \cos(kx + \omega t) + \text{constant}, \quad (12)$$

leading to a response

$$O(t) \propto \cos(\omega t + \phi) \quad (13)$$

which scales with  $g(k, \omega)$  the Fourier transform of the spatiotemporal filter

$$K(x, t) \propto \int dk d\omega g(k, \omega) e^{ikx + i\omega t}. \quad (14)$$

The response to the static grating is simply the special case when  $\omega = 0$ . Typically, the monkey retinal ganglion cells are most sensitive to temporal frequency on the order of 10 Hz. This means that the impulse response to a momentary sinusoidal spatial wave is typically a transient wave form lasting about  $\sim 100$  ms. The contrast sensitivity functions of the ganglion cells in monkeys correspond quite well to the human observers' sensitivity to the same gratings.<sup>64</sup> Comparing the P and M ganglion cells, the P cells are more sensitive to higher spatial frequencies while the M cells to higher temporal frequencies.

### 1.3.3 Spatial sampling on the retina

For each unit area of visual space, more cones and retinal ganglion cells are devoted to the central than the periphery visual space. Fig. (4A) shows that the density  $D$  of cones per unit area decreases rapidly with eccentricity  $e$ , the distance in visual angle from the center of vision. Roughly,

$$D \propto \alpha / (e_o + e) \quad (15)$$

with  $e_o \sim 1 - 2^\circ$ .<sup>123</sup> Consequently, visual acuity drops drastically with eccentricity  $e$ , as demonstrated in Fig. (4B), the size of the smallest recognizable letter increases roughly linearly with  $e$ . The sizes of the receptive fields of the ganglion cells also scale up with  $e$  accordingly.<sup>123</sup> Hence, humans have to use eye movements to bring objects of interests to fovea in order to scrutinize them. Such eye movements, or saccades, occur at a rate of about three times a second, although we are typically unaware that we saccade this frequently, suggesting that many of the saccades are carried out more or less involuntarily. Related to this is the problem for the human visual system to decide where in the visual space to saccade to next, or which object in the visual scene to pay attention to. This is the problem of visual attention, which we will discuss extensively in the book.

Rods belong to another class of photoreceptors that function mainly in dim light due to their higher sensitivity to light. Because the cones are packed so density in the fovea, there is no rods in the center of fovea, and rod density peaks around  $20^\circ$  eccentricity, as shown in Fig. (4A). As cones are not functional in very dim light, one often has to not look at something directly in such an environment in order to make it visible by bringing the image of the object to the rods on the retina. This maybe necessary to see a dim star in the night sky.

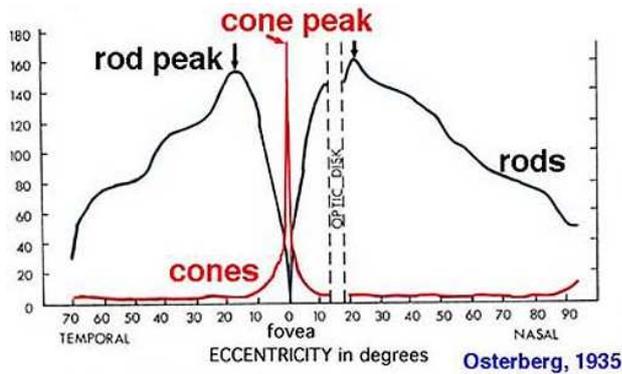
### 1.3.4 Color processing in the retina

Cones belong to the class of photoreceptors which are activate by day light. In human vision, there are red, green, and blue cone types, defined by their selective sensitivity to the predominantly red, green, or blue parts of the visible light spectrum, so that they are most activated by image locations emitting light that are more red, green, or blue respectively, see Fig. (5A). It is interesting to note that the sensitivity curves of the red and green cones overlap a lot, making the responses of the two cones very correlated. At the ganglion cell level, the different cones can contribute to different spatial regions of the receptive fields. For example, the red cone input can excite the center of the receptive field and the green cone inhibit the surround, giving red-on-center and green-off-surround receptive field, see Fig. (5B), making this cell most sensitive to a small red disk of light. It will be explained later in the book (section ??) that such a receptive fiel organization serves a computational goal of efficient color coding, decorrelating the responses from the red and green cones. Other ganglion cells can be of the type blue-center-yellow-surround, giving blue-yellow opponency. The color tuned ganglion cells are the P cells, while the M cells are not color tuned.

## 1.4 The primary visual cortex (V1)

The optic nerve carries the responses of the retinal ganglion cells to a region of the thalamus called the lateral geniculate nucleus, or LGN for short, see Fig. (6). As mentioned above, the function of the LGN is unclear. It has been seen as a relay station for retinal signals on route to the primary visual cortex mainly because the receptive fields of the LGN cells resemble very much those of the retinal ganglion cells in aneathetized animals, and because there is a lack of concensus regarding its function due to our current ignorance, except that the brain is unlikely to waste resources on a relay

A: Density of photoreceptors ( $\times 10^3 / \text{mm}^2$ ) vs. eccentricity



B: Visual acuity illustrated in an eye chart

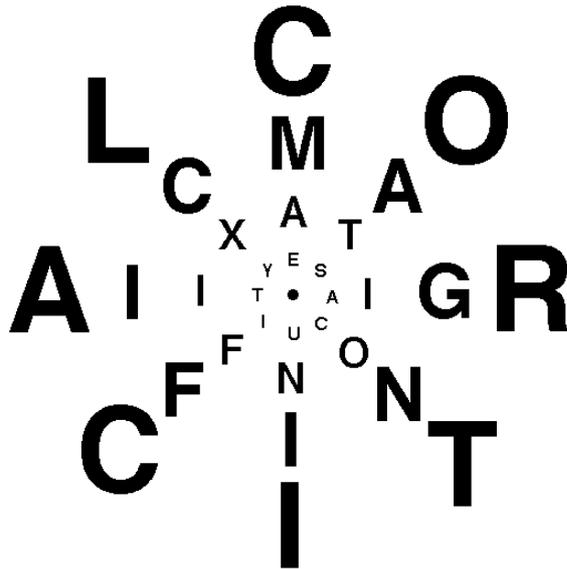


Figure 4: A: The density of human cones and rods versus visual angle from the center of vision according to Osterberg<sup>95</sup> (1935), adapted from <http://www.webvision.med.utah.edu/phphoto2.html#cones>. Note that sampling density of cones drops dramatically with eccentricity, densest at the fovea where there is no room for the rods, whose density peaks slightly off fovea. B: visual acuity drops dramatically with increasing eccentricity: fixating at the center of the eye chart, all the letters are equally visible, from Stuart Anstis, <http://www.psy.ucsd.edu/~sanstis/SABlur.html>.

station for no other reasons. More details about the LGN can be found in a chapter by Sherman and Guillery (2004).<sup>108</sup> The primary visual cortex receives retinal inputs via LGN.

### 1.4.1 The retinotopic map

Neighboring points in a visual image evoke activity in neighboring regions of the primary visual cortex. The retinotopic map refers to the transformation from the coordinates of the visual world to that on the cortical surface, see Fig. (7). It is clear that the cortex devote more surface areas to the central part of visual field, just as the retina devote more receptors and ganglion cells to the fovea region. There is also a transformation of the visual space in angles eccentricity  $e$  and azimuth

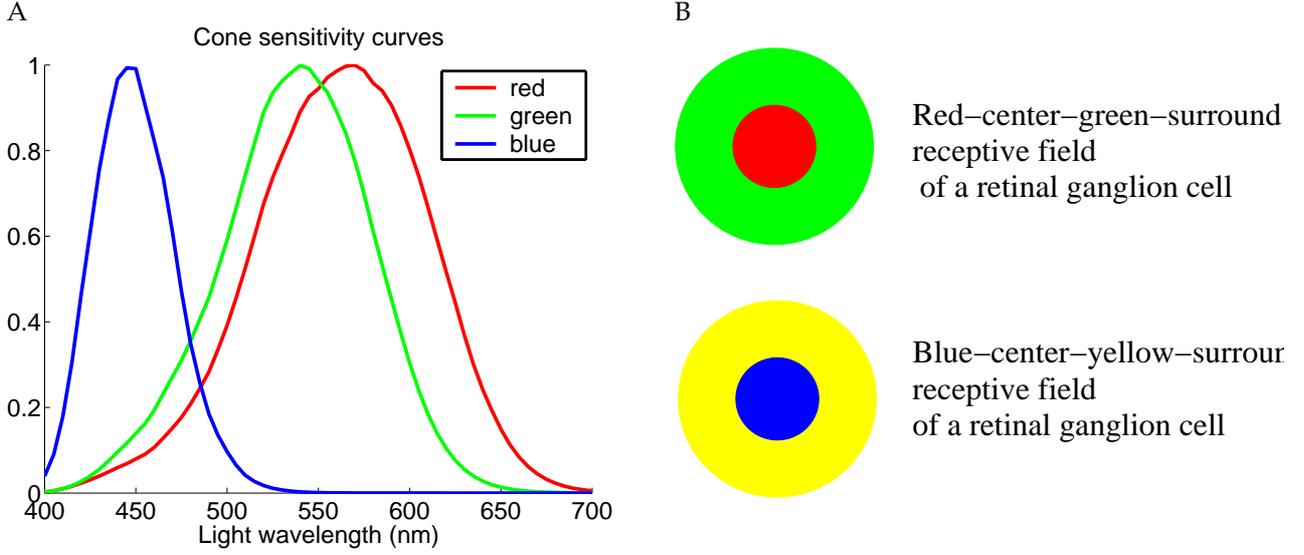


Figure 5: A: Spectrum sensitivity of the cones as a function of the wavelength of light. B: schematics of two retinal ganglion cells with center-surround color opponency in their receptive fields.

$a$  into the cortical Cartesian coordinates  $X$  going along the horizon and  $Y$  going perpendicular to it. The correspondence between the visual space in degrees  $e$  and  $a$  and cortical surface  $X$  and  $Y$  in millimeters (mm) is approximately:

$$X = \lambda \ln(1 + e/e_0) \quad Y = -\frac{\lambda e a \pi}{(e_0 + e)180^\circ} \quad (16)$$

where  $\lambda \approx 12$  mm and  $e_0 \approx 1^\circ$ , and the negative sign in the expression for  $Y$  comes from the inversion of visual image in the image formation process. For visual locations much beyond the foveal region, i.e.,  $e \gg e_0 \approx 1$ , we have  $X \approx \lambda \ln(e/e_0)$  growing linearly with log eccentricity  $\ln e$  and  $Y \approx -\lambda \pi a / 180^\circ$  growing linearly with azimuth  $a$ . Denoting  $z \equiv (e/e_0) \exp(-i\pi a / 180^\circ)$  and  $Z \equiv X + iY$  (with  $i = \sqrt{-1}$ ), we have  $Z = \lambda \ln(z)$  for large eccentricity locations. Hence, the cortical map is sometimes called a complex logarithmic map. A scaling of image  $e \rightarrow \gamma e$  on the retina corresponds to a shift on the cortex  $X \rightarrow X + \lambda \ln(\gamma)$  for large  $e$ . This of course applies only approximately for large  $e$ . The cortical magnification factor

$$M(e) \equiv \frac{dX}{de} = \frac{\lambda}{(e + e_0)} \quad (17)$$

characterizes the degree to which cortical areas are devoted to visual space at different eccentricity  $e$ . Its similarity to how retinal receptor density  $D \propto 1/(e + e_0)$  depends on  $e$  in equation (12), perhaps with a different but similar numerical value of  $e_0$ , is apparent.

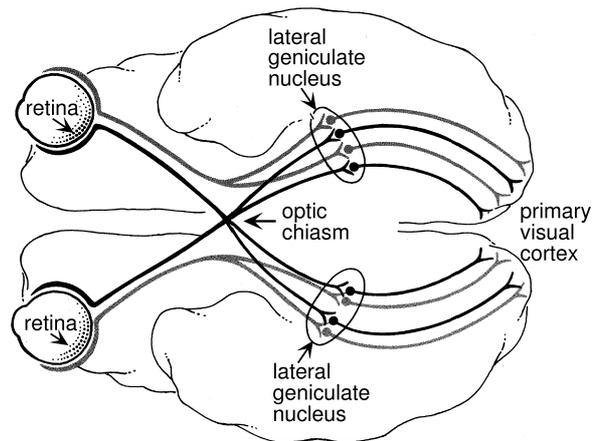


Figure 6: The retina sends information the primary visual cortex via LGN, from Fig. 2.5 of Dayan and Abbott's book.<sup>23</sup> Information from the two eyes are separated in separate layers within LGN, but combined in the primary visual cortex. Information from two different hemifields of the visual space, left and right hemifields, are sent to right and left part of the primary visual cortical regions.

## 2 The higher visual areas

## 3 Behavioral studies on vision

## 4 Etc

## References

- [1] Adelson EH and Bergen JR, (1985) Spatiotemporal energy models for the perception of motion *Journal of Optical Society of America. A* 2, 284-299.
- [2] Andrews T. J. & Coppola, D. M. (1999) Idiosyncratic characteristics of saccadic eye movements when viewing different visual environments. *Vision Research* 39. p. 2947-2953.
- [3] Allman J, Miezin F, McGuinness E. Stimulus specific responses from beyond the classical receptive field: neurophysiological mechanisms for local-global comparisons in visual neurons. *Annu Rev. Neurosci.* 8:407-30 (1985).
- [4] Atick JJ and Redlich AN Towards a theory of early visual processing (1990) *Neural Computation* 2:308-320.
- [5] Atick JJ, Li Z, Redlich AN (1990) Color coding and its interaction with spatiotemporal processing in the retina. Preprint IASSNS-HEP-90-75 of Institute for Advanced Study, Princeton, USA.
- [6] Atick JJ. Could information theory provide an ecological theory of sensory processing. *Network:Computation and Neural Systems* 3: 213-251. (1992)

A: visual space

B: retinotopic map in V1

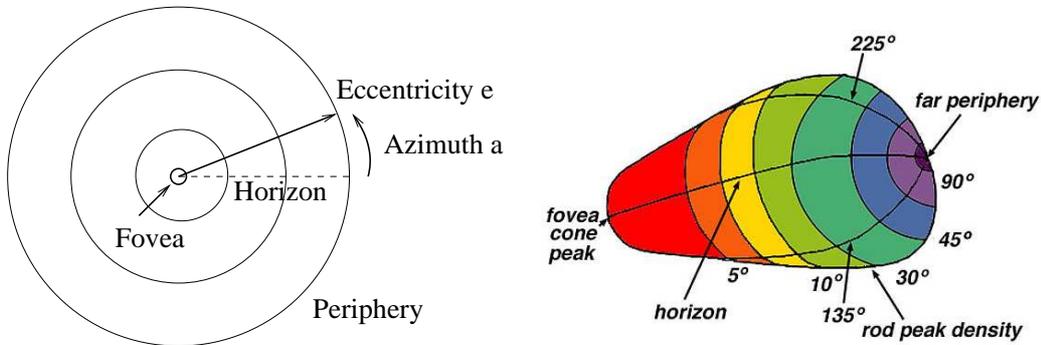


Figure 7: A: definitions of the visual angles eccentricity and azimuth in the visual space. B: The retinotopic map of the visual space onto the primary visual cortex (from The primary visual cortex by Matthew Schmolesky at <http://webvision.med.utah.edu/VisualCortex.html> ) showing higher magnification to the more central part of the visual field. The angles  $5^\circ$ ,  $10^\circ$ ,  $30^\circ$ ,  $45^\circ$ , and  $90^\circ$  mark eccentricity  $e$ , while angles  $135^\circ$  and  $225^\circ$  mark azimuth  $a$ .

- [7] Atick J.J. Li, Z., and Redlich A. N. Understanding retinal color coding from first principles *Neural Computation* 4(4):559-572 (1992).
- [8] Atick J. J., Li, Z., and Redlich A. N. What does post-adaptation color appearance reveal about cortical color representation? (1993) *Vision Res.* 33(1):123-9.
- [9] Becker W. (1991). Saccades. In R.H.S. Carpenter (Eds), *Eye movements* (pp. 95-137). London: Macmillan.
- [10] Bahill, A.T., Adler D., & Stark , L. (1975) Most naturally occurring human saccades have magnitudes of 15 degrees or less. *Investigative Ophthalmology*, 14, p.468-469.
- [11] Bullier J. (2004) Communications between cortical areas of the visual system. Page 522-540 of *The Visual Neuroscience* Eds. Chalupa L.M. and Werner JS. MIT press.
- [12] Bruce CJ, Friedman HR. , Kraus MS, and Stanton GB. (2004) The primate frontal eye field. page 1429-1448. of *The Visual Neuroscience* Eds. Chalupa L.M. and Werner JS. MIT press.
- [13] Bruce V. Green PR., Georgeson MA "Visual perception, physiology, psychology, and ecology" Fourth Edition, Psychology Press, New York, 2003.
- [14] Buchsbaum G, Gottschalk A. (1983) Trichromacy, opponent colours coding and optimum colour information transmission in the retina. *Proc R Soc Lond B Biol Sci.* 220(1218):89-113.
- [15] Barlow HB, 1961, "Possible principles underlying the transformations of sensory messages." In: Sensory Communication W.A. Rosenblith, ed., Cambridge MA, MIT Press, pp. 217-234.
- [16] Barlow H.B. (1981) The Ferrier Lecture, 1980: Critical limiting factors in the design of the eye and visual cortex. *Proc. R. Soc. London B* 212, 1-34.

- [17] Barlow HB, Fitzhugh R, Kuffler SW (1957) Change of organization in the receptive fields of the cat retina during dark adaptation. *J. Physiol.* 137: 338-354.
- [18] Bell AJ Sejnowski TJ (1997) The 'independent components' of natural scenes are edge filters. *Vision Res.* 23: 3327-38.
- [19] Bressloff PC, Cowan JD, Golubitsky M, Thomas PJ, Wiener MC. (2002) What geometric visual hallucinations tell us about the visual cortex. *Neural Comput.* 14(3):473-91.
- [20] Casagrande VA, Guillery RW, and Sherman SM (2005) Eds. *Cortical function: a view from the thalamus*, volumn 149. of *Progress in Brain Research*, Elsevier 2005.
- [21] Chalupa L.M. and Werner JS. (Editors of) *The visual neurosciences* MIT press, 2004.
- [22] Dan Y, Atick JJ, Reid RC. (1996) Efficient coding of natural scenes in the lateral geniculate nucleus: experimental test of a computational theory *J Neurosci.* 16(10):3351-62.
- [23] Dayan P and Abbott L.F. (2001) *Theoretical Neuroscience: Computational and Mathematical Modeling of Neural Systems*, the MIT press, 2001.
- [24] Daubechies I. *The lectures on wavelets*, SIAM 1992.
- [25] Dong DW, Atick JJ 1995 "Temporal decorrelation: a theory of lagged and non-lagged responses in the lateral geniculate nucleus," Network: Computation in Neural Systems, 6:159-178.
- [26] Duncan J., Humphreys G.W. 1989 Visual search and stimulus similarity, *Psychological Rev.* 96(3): 433-58.
- [27] Douglas RJ, Koch C, Mahowald M, Martin KA, Suarez HH. (1995) Recurrent excitation in neocortical circuits. *Science* 269(5226):981-5.
- [28] Egeth HE, Yantis S. (1997) Visual attention: control, representation, and time course. *Annual Rev. Pyschol.* 48:269-97.
- [29] Felleman DJ, Van Essen DC. (1991) Distributed hierarchical processing in the primate cerebral cortex. *Cereb Cortex.* 1(1):1-47.
- [30] Field DJ 1987 Relations between the statistics of natural images and the response properties of cortical cells. Journal of Optical Society of America, A 4(12):2379-94. 1987
- [31] Field DJ 1989 What the statistics of natural images tell us about visual coding. *SPIE* vol. 1077 Human vision, visual processing, and digital display, 269-276
- [32] Findlay JM and Gilchrist ID (2003) *Active vision, the psychology of looking and seeing*. Oxford University Press, UK.
- [33] Freeman WT and Adelson EH (1991) The design and use of steerable filters. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 13(9):891-906.
- [34] Gegenfurtner, K. R., Kiper, D. C., & Fenstemaker, S. B. (1996). Processing of color, form, and motion in macaque area V2. *Visual Neuroscience* 13, 161172.
- [35] Gilbert C.D., Wiesel T.N., Clustered intrinsic connections in cat visual cortex. *J. Neurosci.* 3(5):1116-33 (1983)

- [36] Gottlieb JP, Kusunoki M, Goldberg ME. (1998) The representation of visual salience in monkey parietal cortex. *Nature* 391(6666):481-4.
- [37] Hamilton DB, Albrecht DG, Geisler WS. (1989) Visual cortical receptive fields in monkey and cat: spatial and temporal phase transfer function. *Vision Research* 29(10):1285-308.
- [38] Holub RA, Morton-Gibson M. (1981) Response of Visual Cortical Neurons of the cat to moving sinusoidal gratings: response-contrast functions and spatiotemporal interactions. *Journal of Neurophysiology* 46(6):1244-59
- [39] Hopfield J.J. (1984) Neurons with graded response have collective computational properties like those of two-state neurons. *Proc. Natl. Acad. Sci. USA* 81:3088-3092.
- [40] Horton JC and Hocking DR (1996) Anatomical demonstration of ocular dominance columns in striate cortex of the squirrel monkey *J. Neurosci.* 16(17):5510-5522.
- [41] Horowitz GD, Albright TD. (2005) Paucity of chromatic linear motion detectors in macaque V1. *J Vis.* 5(6):525-33.
- [42] Hubel, D. H. and Wiesel, T. N. 1965. "Binocular interaction in striate cortex of kittens reared with artificial squint." *J. Neurophysiol* 28: 1041-1059.
- [43] Hubel DH Wiesel TN Receptive fields and functional architecture of monkey striate cortex. *J Physiol.* 195(1):215-43. (1968).
- [44] Hubel, D. H. and Wiesel, T. N. LeVay, S 1977. "Plasticity of ocular dominance columns in monkey striate cortex." *Philosophical Transactions of the Royal Society of London, Series B*, 278: 377-409.
- [45] Itti L. and Baldi P. (2006) "Bayesian surprise attracts human attention." In *Advances in Neural Information Processing Systems*, Vol. 19 (NIPS2005), pp. 1-8, Cambridge, MA:MIT Press.
- [46] Itti L., Koch C. A saliency-based search mechanism for overt and covert shifts of visual attention. *Vision Res.* 40(10-12):1489-506, (2000).
- [47] Jones HE, Grieve KL, Wang W, Sillito AM. (2001) Surround suppression in primate V1. *J Neurophysiol.* 86(4):2011-28.
- [48] Jonides J. (1981) Voluntary versus automatic control over the mind's eye's movement. In J. B. Long & A. D. Baddeley (Eds.) *Attention and Performance IX* (pp. 187-203). Hillsdale, NJ. Lawrence Erlbaum Associates Inc.
- [49] Julesz B. (1981) Textons, the elements of texture perception, and their interactions. *Nature* 290(5802):91-7.
- [50] Kadir T. Brady M. (2001) Saliency, scale, and image description. *International J. of Computer Vision* 45(2):83-105.
- [51] Kapadia MK, Ito M, Gilbert CD, Westheimer G. Improvement in visual sensitivity by changes in local context: parallel studies in human observers and in V1 of alert monkeys. *Neuron* 15(4):843-56 (1995).
- [52] Kaplan E, Marcus S, So YT. Effects of dark adaptation on spatial and temporal properties of receptive fields in cat lateral geniculate nucleus. (1979) *J. Physiol.* 294:561-80.

- [53] Kelly D. H. Information capacity of a single retinal channel. *IEEE Trans. Information Theory* 8:221-226, 1962.
- [54] Knierim JJ., Van Essen DC, Neuronal responses to static texture patterns in area V1 of the alert macaque monkey. *J. Neurophysiol.* 67(4): 961-80 (1992)
- [55] Koch C., Ullman S. Shifts in selective visual attention: towards the underlying neural circuitry. *Hum. Neurobiol.* 4(4): 219-27 (1985).
- [56] Koene AR and Zhaoping L. (2007) Feature-specific interactions in salience from combined feature contrasts: Evidence for a bottom-up saliency map in V1. , *Journal of Vision*, 7(7):6, 1-14, <http://journalofvision.org/7/7/6/>, doi:10.1167/7.7.6
- [57] Laughlin S B (1981) A simple coding procedure enhances a neuron's information capacity. *Z. Naturforsch [C]* 36: 910-2.
- [58] Laughlin SB, Howard J, Blakeslee B. (1987) Synaptic limitations to contrast coding in the retina of the blowfly *Calliphora*. *Proc. R. Soc. Lond. B Biol. Sci.* 231(1265):437-67.
- [59] Lennie P. (2003) The cost of cortical computation. *Curr Biol.* 13(6):493-7.
- [60] Lewis A, Garcia R, Zhaoping L. (2003) The distribution of visual objects on the retina: connecting eye movements and cone distributions. *Journal of vision* 3(11):893-905.
- [61] Lewis AS and Zhaoping L. (2005) Saliency from natural scene statistics. Program No. 821.11. *2005 Abstract Viewer/Itinerary Planner*. Washington, DC: Society for Neuroscience, 2005. Online.
- [62] Lewis AS and Zhaoping L. (2006) Are cone sensitivities determined by natural color statistics? *Journal of Vision.* 6(3):285-302. <http://www.journalofvision.org/6/3/8/>.
- [63] Levy WB and Baxter RA 1996 Energy efficient neural codes. *Neural Computation*, 8(3) 531-43.
- [64] Lee BB, Pokorny J., Smith VC, Martin PR, Valberg A. (1990) Luminance and chromatic modulation sensitivity of macaque ganglion cells and human observers. *J. Opt. Soc. Am. A* 7(12):2223-2236.
- [65] Lee TS (1996) Image representation using 2D Gabor wavelets. *IEEE Trans. Pattern Analysis and Machine Intelligence* 18, 959-971.
- [66] Lee TS (2003) Computations in the early visual cortex. *J. Physiology, Paris* 97(2-3):121-39.
- [67] Li Zhaoping (1992) Different retinal ganglion cells have different functional goals. *International Journal of Neural Systems*, 3(3):237-248.
- [68] Li Zhaoping and Atick J. J. 1994a, "Towards a theory of striate cortex" *Neural Computation* **6, 127-146**
- [69] Li Zhaoping and Atick J. J. 1994b, "Efficient stereo coding in the multiscale representation" *Network: computation in neural systems* 5(2):157-174.
- [70] Li, Zhaoping Understanding ocular dominance development from binocular input statistics. in *The neurobiology of computation* (Proceeding of Computational Neuroscience Conference, , July 21-23, 1994, Monterey, California, 1994), p. 397-402. Ed. J. Bower, Kluwer Academic Publishers, 1995.

- [71] Li Zhaoping 1996 "A theory of the visual motion coding in the primary visual cortex" Neural Computation vol. 8, no.4, p705-30.
- [72] Li Z. (1998a) Primary cortical dynamics for visual grouping. *Theoretical aspects of neural computation* Eds. Wong KM, King I, Yeung D-Y. pages 155-164. Springer-verlag, Singapore, January 1998.
- [73] Li Z. (1998b) A neural model of contour integration in the primary visual cortex. *Neural Comput.* 10(4):903-40.
- [74] Li Z. (1998c) Visual segmentation without classification: A proposed function for primary visual cortex. *Perception* Vol. 27, supplement, p 45. (Proceedings of ECVF, 1998, Oxford, England).
- [75] Li Z. Visual segmentation by contextual influences via intra-cortical interactions in primary visual cortex. *Network: Computation and neural systems* 10(2):187-212, (1999a).
- [76] Li Z. Contextual influences in V1 as a basis for pop out and asymmetry in visual search. *Proc. Natl Acad. Sci USA*, 96(18):10530-5. (1999b)
- [77] Li, Zhaoping and Dayan P. (1999) "Computational differences between asymmetrical and symmetrical networks" *Network: Computation in Neural Systems* Vol. 10, 1, 59-77.
- [78] Li Z, Pre-attentive segmentation in the primary visual cortex. *Spatial Vision*, 13(1) 25-50. (2000)
- [79] Li Z. (2001) Computational design and nonlinear dynamics of a recurrent network model of the primary visual cortex. *Neural Computation* 13(8):1749-1780.
- [80] Li Zhaoping 2002, "A saliency map in primary visual cortex " Trends in Cognitive Sciences Vol 6. No.1. Jan. 2002, page 9-16
- [81] Linsker R. 1990 Perceptual neural organization: some approaches based on network models and information theory. Annu Rev Neurosci. 13:257-81.
- [82] Livingstone MS, Hubel DH. Anatomy and physiology of a color system in the primate visual cortex. *J. Neurosci.* 4(1):309-56 (1984).
- [83] Meister M, Berry MJ (1999) The neural code of the retina *NEURON* 22(3):435-450.
- [84] Mitchison G. (1992) Axonal trees and cortical architecture. *Trends Neurosci.* 15(4):122-6
- [85] Moran J, Desimone R. (1985) Selective attention gates visual processing in the extrastriate cortex. *Science* 229(4715):782-4.
- [86] Nadal J.P. and Parga N. 1993. Information processing by a perceptron in an unsupervised learning task. *Network:Computation and Neural Systems* 4(3), 295-312.
- [87] Nadal J.P. and Parga N. 1994. Nonlinear neurons in the low-noise limit: a factorial code maximizes information transfer. *Network:Computation and Neural Systems* 5:565-581.
- [88] Nakayama K, Mackeben M. (1989) Sustained and transient components of focal visual attention. *Vision Res.* 29(11):1631-47.
- [89] Nelson JJ, and Frost BJ. (1985) Intracortical facilitation among co-oriented, co-axially aligned simple cells in cat striate cortex. *Exp Brain Res.* 61(1):54-61.

- [90] Nirenberg, S. Carcieri, S. M. Jacobs A. L. and Latham P. E. (2001) Retinal ganglion cells act largely as independent encoders *Nature* 411:698-701.
- [91] Nothdurft HC (1991) Texture segmentation and pop-out from orientation contrast. *Vision Res.* 31(6):1073-8
- [92] Nothdurft HC, Gallant JL, Van Essen DC. Response modulation by texture surround in primate area V1: correlates of "popout" under anesthesia. *Vis. Neurosci.* 16, 15-34 (1999).
- [93] Olshausen BA and Field DJ (1997) Sparse coding with an overcomplete basis set: a strategy employed by V1? *Vision Research* 37:3311-3325.
- [94] Olshausen BA and Field DJ (2005) How Close Are We to Understanding V1? *Neural Computation* 17:1665-1699
- [95] Osterberg GA (1935) Topography of the layer of rods and cones in the human retina. *Acta Ophthalmol.* 6 (suppl. 13), 1-102.
- [96] Palmer SE "Vision Science: photons to phenomenology", 1999, MIT press.
- [97] Pashler H. (1998) *Attention* Editor. East Sussex, UK. Psychology Press Ltd.
- [98] Petrov Y. Zhaoping L. Local correlations, information redundancy, and sufficient pixel depth in natural images. *J. Opt Soc. Am. A Opt Image Sci. Vis.* 20(1):56-66. (2003).
- [99] Puchalla JL, Schneidman E, Harris RA, Berry MJ. 2005 Redundancy in the population code of the retina. *Neuron* 46(3):493-504.
- [100] Rao RPN and Ballard DH Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive field effects. *Nature Neuroscience* 2: 79-87 (1999).
- [101] Rockland KS., Lund JS., Intrinsic laminar lattice connections in primate visual cortex. *J. Comp. Neurol.* 216(3):303-18 (1983).
- [102] Rolls E.T. (2004) Invariant Object and Face recognition. *The Visual Neuroscience* Eds. Chalupa LM. and Werner JS. p. 1165-1178. MIT press.
- [103] Rubenstein B.S., Sagi D., Spatial variability as a limiting factor in texture-discrimination tasks: implications for performance asymmetries. *J Opt Soc Am A.*;7(9):1632-43, (1990)
- [104] Ruderman DL, Cronin TW, Chiao C-C. Statistics of cone responses to natural images: implications for visual coding. *Journal of Optical Society of America. A* 15(8):2036-45.
- [105] Salinas E, Abbott LF. (2000) Do simple cells in primary visual cortex form a tight frame? *Neural Comput.* 12(2):313-35.
- [106] Schreiber W. 1956 The measurement of third order probability distributions of television signals *IEEE Trans. Information Theory* 2(3):94-105.
- [107] Schwartz O, Simoncelli EP. (2001) Natural signal statistics and sensory gain control. *Nat Neurosci* 4(8): 819-25.
- [108] Sherman SM Guillery RW (2004) The visual relays in the thalamus *The Visual Neuroscience* Eds. Chalupa LM. and Werner JS. p. 565-591 MIT press.

- [109] Shannon CE and Weaver W (1949) The mathematical theory of communication (Urbana IL, University of Illinois Press)
- [110] Shapley R and Perry VH (1986) Cat and monkey retinal ganglion cells and their visual functional roles. *Trends in Neuroscience* 9: 229-235.
- [111] Sillito AM, Grieve KL, Jones HE, Cudeiro J, Davis J. Visual cortical mechanisms detecting focal orientation discontinuities. *Nature* 378, 492-496 (1995).
- [112] Simons D.J. & Chabris C.F. (1999) Gorillas in our midst: sustained inattention blindness for dynamic events. *Perception* 28: 1059-1074
- [113] Simoncelli E and Olshausen B. 2001 "Natural image statistics and neural representation" Annual Review of Neuroscience, 24, 1193-216.
- [114] Srinivasan MV, Laughlin SB, Dubs A. (1982) Predictive coding: a fresh view of inhibition in the retina. *Proc R Soc Lond B Biol Sci.* 216(1205):427-59.
- [115] Stryker M. P. 1986 "The role of neural activity in rearranging connections in the central visual system." In Ruben, R.J. Vater, T.R. Van De, and Rubel, E.W. (Eds.), *The biology of Change in Otolaryngology*, pp. 211-224, Elsevier Science B.V. Amsterdam.
- [116] Sziklai G (1956) Some studies in the speed of visual perception *IEEE Transactions on Information Theory* 2(3):125-8
- [117] Tehovnik EJ, Slocum WM, Schiller PH. Saccadic eye movements evoked by microstimulation of striate cortex. *Eur J. Neurosci.* 17(4):870-8 (2003).
- [118] Treisman A. M., Gelade G. A feature-integration theory of attention. *Cognit Psychol.* 12(1), 97-136, (1980).
- [119] Treisman A, Gormican S. (1988) Feature analysis in early vision: evidence from search asymmetries. *Psychol Rev.* 95(1):15-48.
- [120] Troy JB, Lee BB.(1994) Steady discharges of macaque retinal ganglion cells. *Vis.Neurosci.* 11(1):111-8.
- [121] Troy JB, Robson JG. (1992) Steady discharges of X and Y retinal ganglion cells of cat under photopic illuminance. *Vis. Neurosci.* 9(6):535-53.
- [122] Tsotsos, J.K. (1990) Analyzing Vision at the Complexity Level. *Behavioral and Brain Sciences* 13-3:423 - 445.
- [123] Van Essen DC, and Anderson CH. (1995) Information processing strategies and pathways in the primate visual system. *An Introduction to Neural and Electronic Networks* 2nd edition. S. Zornetzer, J.L. Davis, C. Lau, T. McKenna, eds. Academic Press, Florida.
- [124] Van Essen DC, Anderson CH, Felleman DJ.(1992) Information processing in the primate visual system: an integrated systems perspective. *Science* 255(5043):419-23.
- [125] van Hateren J. (1992) A theory of maximizing sensory information. *Biol Cybern* 68(1):23-9.
- [126] van Hateren J. Ruderman D.L. (1998) Independent component analysis of natural image sequences yields spatio-temporal filters similar to simple cells in primary visual cortex *Proc. Biol. Sciences.* 265(1412):2315-20

- [127] Wachtler T, Sejnowski TJ, Albright TD. (2003) Representation of color stimuli in awake macaque primary visual cortex. *Neuron* 37(4):681-91.
- [128] Wandell B. A. (1995) *Foundations of Vision* Sinauer Associates. Inc.
- [129] Wolfe J.M., Cave K.R., Franzel S. L. Guided search: an alternative to the feature integration model for visual search. *J. Experimental Psychol.* 15, 419-433, (1989).
- [130] Wolfe J. M. (1998) "Visual Search, a review" In H. Pashler (ed.) *Attention* p, 13-74. Hove, East Sussex, UK. Psychology Press Ltd.
- [131] Zhaoping L. (2005) The primary visual cortex creates a bottom-up saliency map. In *Neurobiology of Attention* Eds L. Itti, G. Rees and J.K. Tsotsos, Elsevier, 2005, Chapter 93, page 570-575
- [132] Zhaoping L. Hubner M., and Anzai A. (2006) Efficient stereo coding in the primary visual cortex and its experimental tests based on optical imaging and single cell data. Presented at Annual Meeting of Computational Neuroscience, Edinburgh, summer, 2006.
- [133] Zhaoping L. and May K.A. (2004) Irrelevance of feature maps for bottom up visual saliency in segmentation and search tasks. Program No. 20.1 *2004 Abstract Viewer/Itinerary Planner*, Washington D. C., Society for Neuroscience, 2004. Online.
- [134] Zhaoping L. and May K.A. (2007) Psychophysical tests of the hypothesis of a bottom-up saliency map in primary visual cortex, *Public Library of Science, Computational Biology* 3(4):e62. doi:10.1371/journal.pcbi.0030062.
- [135] Zhaoping L. and Snowden RJ (2006) A theory of a saliency map in primary visual cortex (V1) tested by psychophysics of color-orientation interference in texture segmentation. in *Visual Cognition* 14(4-8):911-933.
- [136] Zhaoping L. (2006) Overcomplete representation for fast attentional selection by bottom up saliency in the primary visual cortex. Presented at *European Conference on Visual Perception*, August, 2006.
- [137] Zhaoping L. (2007) Popout by unique eye of origin: A fingerprint of the role of primary visual cortex in bottom-up saliency. Submitted to Annual meeting of Society for Neuroscience 2007.
- [138] Zigmond, M.J., Bloom F. E., Lndis S. C., Roberts J. L., Squire L. R. *Fundamental neuroscience* Academic Press, New York, 1999.