# Chapter 3

# The Visual System

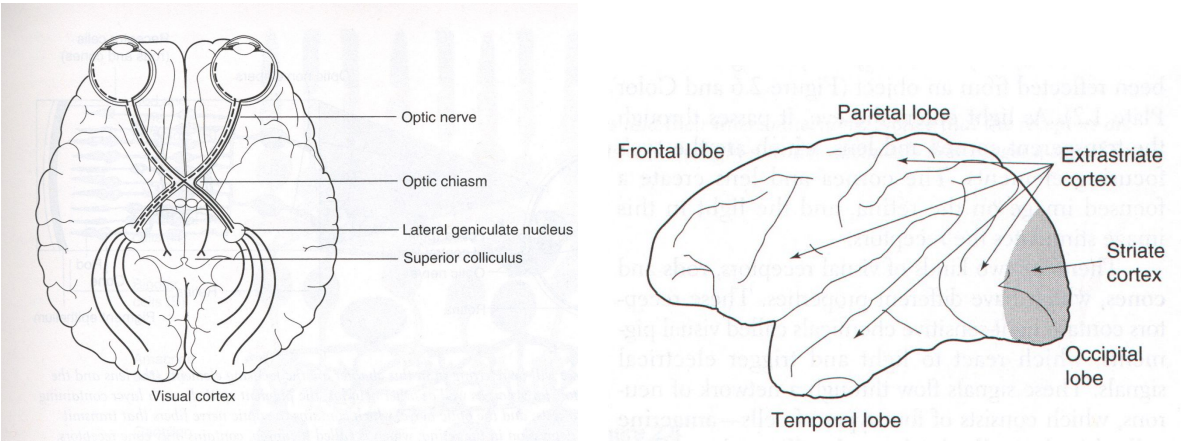



Figure 3.1: Left: The visual system seen from underneath the brain showing the nerve fibres from the retina cross over to the opposite sides of the brain. A small part of the fibres go to superior colliculus. Rest of it go to the visual receiving area of the brain through the lateral genuiculate nucleus. Right: The monkey cortex showing the visual processing areas in the brain including the striate and the extra striate cortex.

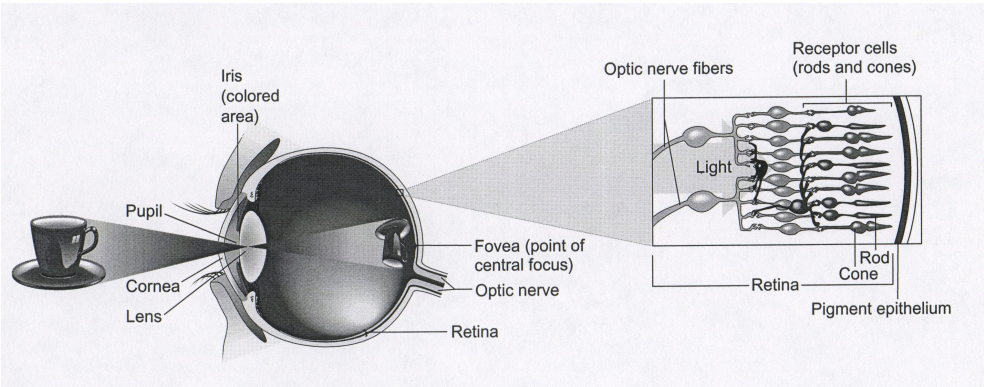



Figure 3.2: The cross section of the human eye. On the right is the zoomed in view of the receptors in the retina.

The visual system consist of the *eye*, the *lateral geniculate nucleus (LGN)* and the *visual receiving area* in the

occipital lobe of the brain. The visual receiving area in the brain is also called *striate cortex* because of its striped appearance. There are also some other processing area in the brain for visual signals which is called the *extra striate cortex*. Figure 3.1 illustrates the visual system.

## 3.1 The Eye

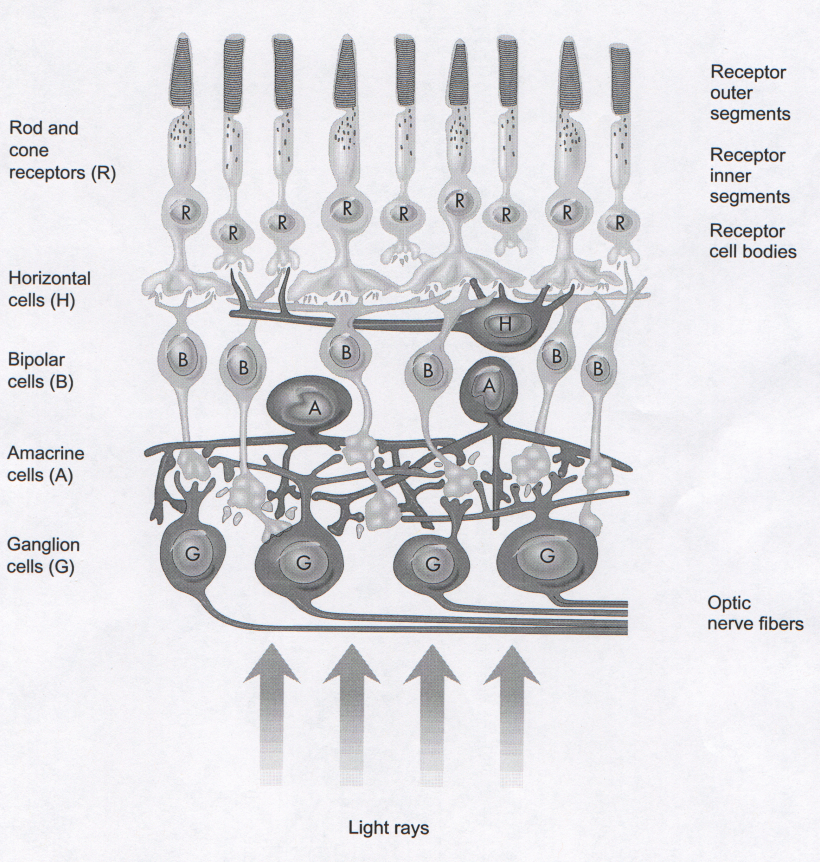



Figure 3.3: A cross section of the primate retina showing the five different kinds of cells and their interconnections.

Light is reflected from objects in the environment and enters our eye and forms an image on the retina. Figure 3.2 shows a cross section of the eye. Light first enters the *cornea* which is responsible for $80\%$ of eye's focussing power. However, since it is held in place, it cannot change its focus. To do this, the *lens* is used which provides the remaining $20\%$ of the focussing ability helping us to adapt to stimulus at different distances. The ciliary muscles of the eye that hold the lens can tighten or loosen depending on the distance of the stimulus. This affects the curvature of the lens and focuses stimulus from different distances on the retina. This ability of the eye to change the focussing power is called *accommodation*. However, accommodation has its limits. Objects too close to the eye or too far away cannot be brought to focus at all. More formally, if objects are closer than the *near point* and farther than the *far point*, they cannot be brought to focus. These near and far points are different for different individuals. If the far point is too close, distant objects always looks blurred. This is called *myopia*. If the near point is far, closer objects look blurred. This is called *presbyopia* and is a common problem for elderly people. Both of these can be rectified by corrective glasses.

The iris is the sphincter muscles that controls the pupil size and hence defines the level of illumination on the

retina. In practical situation, the pupil diameter varies from 3mm to 7mm. This results in approximately five times change in the pupil area and therefore the retinal illuminance. However, the change in visual sensitivity with pupil area is restricted by the Stiles-Crawford Effect by which the marginal rays stimulate the receptors on the retina lesser than the central rays. Hence, the pupil diameter alone cannot explain the excellent human visual function over prevailing illuminance levels that can vary over 10 orders of magnitude.

The optical image of the eye is formed on the retina – a thin layer of cells approximately the thickness of the tissue paper, located at the back of eye and rich in photoreceptors (Details in next section). Behind the retina is a dark pigmented layer called the *pigmented epithelium*. This absorbs the light that happens to pass through the retina without being absorbed by the photoreceptors. The function of this layer is to prevent scattering of light behind the retina so that a sharp high-contrast image is enabled.

The most important structural area on the retina is the *fovea* that subtends about 2 degrees of visual angle in the central field of vision (about the width of your thumbnail). The fovea has the best color and spatial acuity. The fovea is located on the light of sight, so that anytime we look directly at an object the image is falls on the fovea. When we look, we fixate on an object with our fovea. This period of fixation changes very rapidly. To demonstrate how rapidly the spatial acuity falls off as you move away from the fovea, try to read a text in the next paragraph while fixating on this line. The task is very difficult, if not impossible. The area surrounding the fovea is called the *periphery*.

### 3.1.1 Receptors

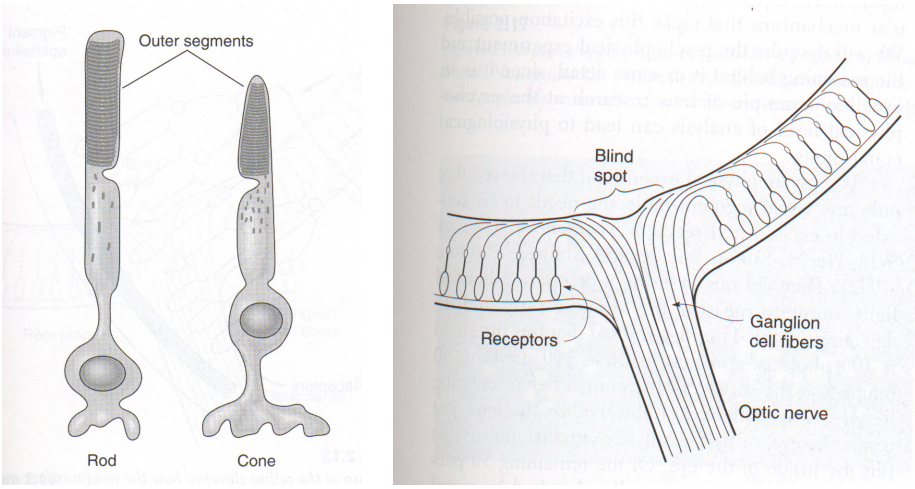



Figure 3.4: Left: This shows the drawings of the rod and cone receptors. Right: There is no receptor at the place where the optic nerve leaves the eye. This enables the ganglion cell fibres to flow into the optical nerve. The absence of receptors creates a blind spot.

Once the image is formed on the retina, the next step is to stimulate the receptors in the retina through this image. The retina has many layers as shown in Figure 3.3. The first layer is the receptors that contain light-sensitive chemicals to convert the light stimulus to electrical signal. There are two kinds of visual receptors, the *rods* and the *cones*. The electrical signals thus generated in the receptors then flow through a network of neurones that consist of four types of cells - *amacrine*, *bipolar*, *horizontal* and *ganglion cells*. The axons of the ganglion cells forms the *optic nerve* which runs from the eye to the LGN.

The rods and cones get their name from their structural differences as shown in Figure 3.4. Note that the receptors are separated into an outer and inner segments. The outer segment contains the light sensitive chemicals that

trigger electrical signals from light stimulus.

### Distribution

The rods and cones are not uniformly distributed in the retina as shown in Figure 3.5. The eye contains almost $50,000$ cones, only $1\%$ of which lies in the fovea while the rest are in the periphery. There are about $120$ million rod cells in the eye, all of which is in the periphery. Hence, the cones are outnumbered by the rods in the periphery by about $20:1$ ratio. However, as shown in Figure 3.3, the receptors are turned away from light which has to travel through all the other four types of cells before reaching the eye. This is to enable the receptors to be in contact with the *pigment epithelium* which provides essential nutrition for the receptors. To enable enough light reaching the receptors, the layers of other four types of cells are transparent. However, this backward facing receptors create one problem. They block the ganglion cells from leaving the eye through the optic nerve. Figure 3.4 shows how eye solves this problem. The ganglion cell fibres form an optic nerve that leaves the eye through an area on the retina that is devoid of any receptors. This area is called the *blind spot*. The reason we are not aware of this spot is due to the fact that higher level processing in the brain extrapolates the information in this portion of the image from the other areas. Further, the blood vessels nourishing the retina are also present in the layers of the cells preceding the receptors. These cast shadows on the retinal image. The brain also fills in for these shadows thus making them imperceptible in the images we see.

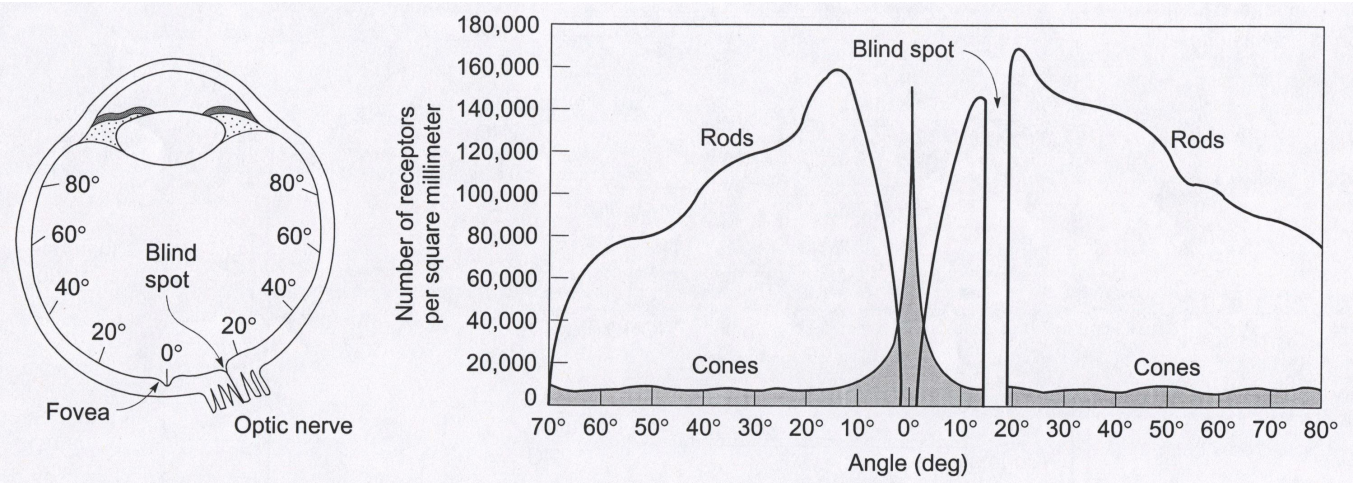



Figure 3.5: Distribution of rods and cones in the retina of eye. The eye on the left shows the degrees relative to the fovea which are repeated along the x-axis of the chart on the right. Note that there are no receptors in the blind spot.

### Function

The more important difference between the rods and the cones are in their visual functions. The cones are responsible for visual experiences at high light levels called the *photopic conditions*, while the rods are responsible for vision at very low light levels (less than 1 $cd/m^2$) called *scotopic conditions*. At very high light levels rods saturate and the cones only function. In the intermediate light conditions, called *mesopic condition*, both rods and cones function and contribute to our vision.

The obvious question at this point is how the light is converted to electric signal in the receptors. The inner segment of the receptors contain the nucleus and other cellular machinery. The outer segment contains billions

of light sensitive pigment molecules. The pigments in the rods is called *rhodopsin*. When the pigment molecule absorbs the photon, it changes its shape through a complex biochemical process in such a way that the flow of electric current in and around the pigment molecule is changed. The results of many photons being absorbed is integrated by an overall change in potential between the inside and outside of the cell. This change is graded, i.e. continuous instead of discrete. This change is usually logarithmic in nature, which is why the response of the eye to increasing brightness can be expressed by Steven's power function. This graded potential is then transmitted down the outer membrane to the synaptic region where it is transmitted to the next neuron. The signal is thus transmitted through the different layers of the retina. Part of the neural processing happens right here in the retina.

This complex set of events in the outer membrane is called *pigment bleaching* since this is associated with a change in the color of the pigment molecule. Once bleached, the pigments cannot absorb any more photons. They are restored to their prior unbleached state by actions of enzymes from the pigment epithelium after which they can again absorb photons. Full regeneration takes about 30 minutes in rods and only about 6 minutes in cones. This time decides the maximum firing rates of the receptors.

**Effects of Visual Pigments on Perception**

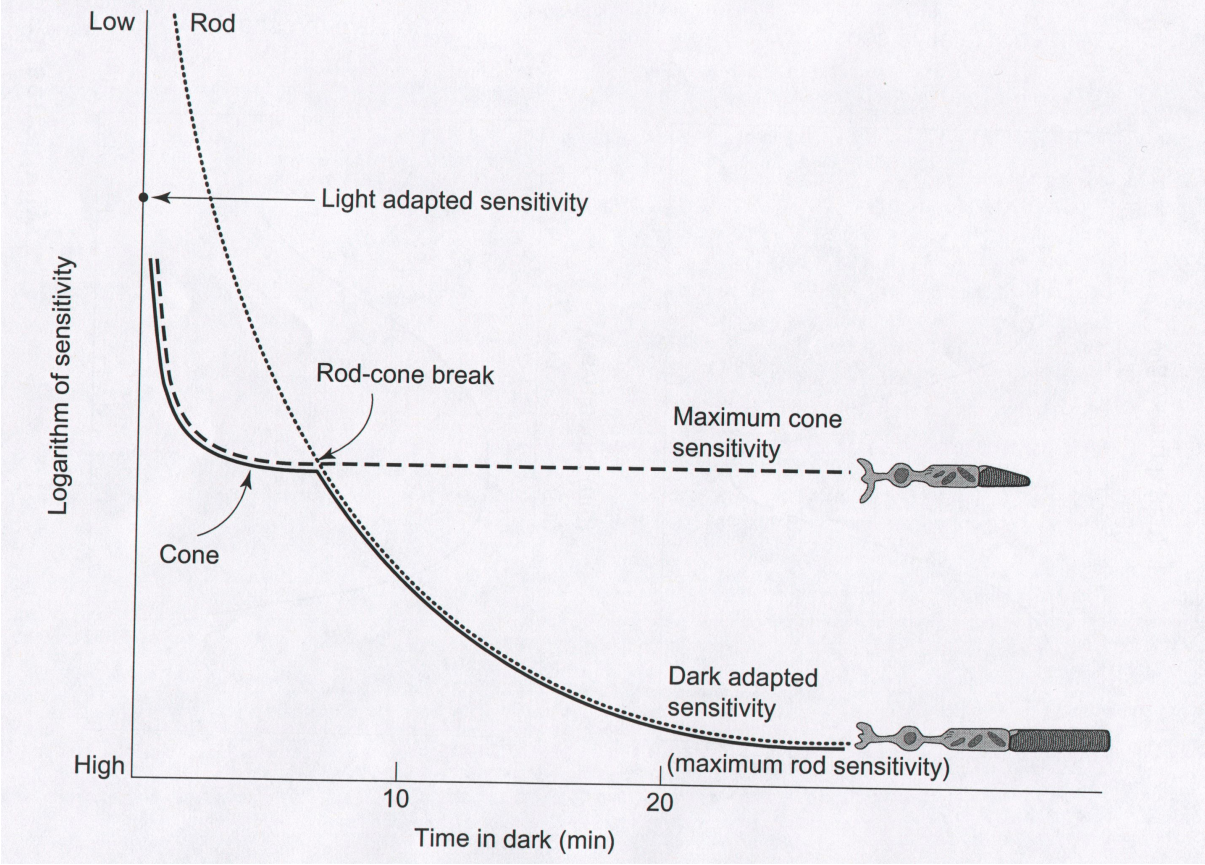



Figure 3.6: The dark adaptation curve. Three curves are shown here. The solid line shows the two stage dark adaptation curve. The dashed line shows the cone adaptation curve and the dotted line shows the rod adaptation curve.

Now, let us see how the visual pigments of the receptors affect our perception. The first topic we will deal with here is *dark adaptation*. Most of you must have experienced the following phenomenon. When you walk into

a dark room from bright sunlight, initially you can hardly see anything. However, slowly with time you start to see better. This increased sensitivity of the eye to dark is called dark adaptation. The plot increased sensitivity of the eye undergoing dark adaptation with time is shown in Figure 3.6. Several psychophysical experiments show that this increase in sensitivity occurs in two distinct stage: a initial rapid stage due to the cone adaptation and a second slower stage due to rod adaptation. This showed that the time taken for the cones to reach their maximum sensitivity is 6 minutes, while that for rods is close to 30 minutes. In fact, evidences were found that showed that this is due to the time taken for pigment regeneration (note the similar timings). Cones have their pigments regenerated soon and hence reach their peak sensitivity earlier, whereas rods take a longer time. Also note that the peak sensitivity of the rods in dark are much higher than the cones. This shows that rods are responsible for scotopic vision.

The spectral sensitivity (the sensitivity to the different wavelengths of light) of the pigments in rods and cones also affect our perception. A simple experiment is performed to determine the spectral sensitivity. The user is presented with monochromatic light (light of a single wavelength) one at a time. For each of the wavelengths, method of adjustment is used by which the user communicates the amount of light required to detect each wavelength. This determines the relative threshold at each wavelength. The sensitivity is determined by the reciprocal of threshold. Thus, the wavelength for which less light is required for detection are the ones to which the eye is more sensitive. The curve thus generated is called the *spectral sensitivity curve*. These curves are illustrated in Figure 3.7. This curve is taken when subjects are looking directly at a stimulus. Hence, this is the sensitivity curve for the cones. Experiments were performed to find the sensitivity curve for the rods as illustrated in Figure 3.8.

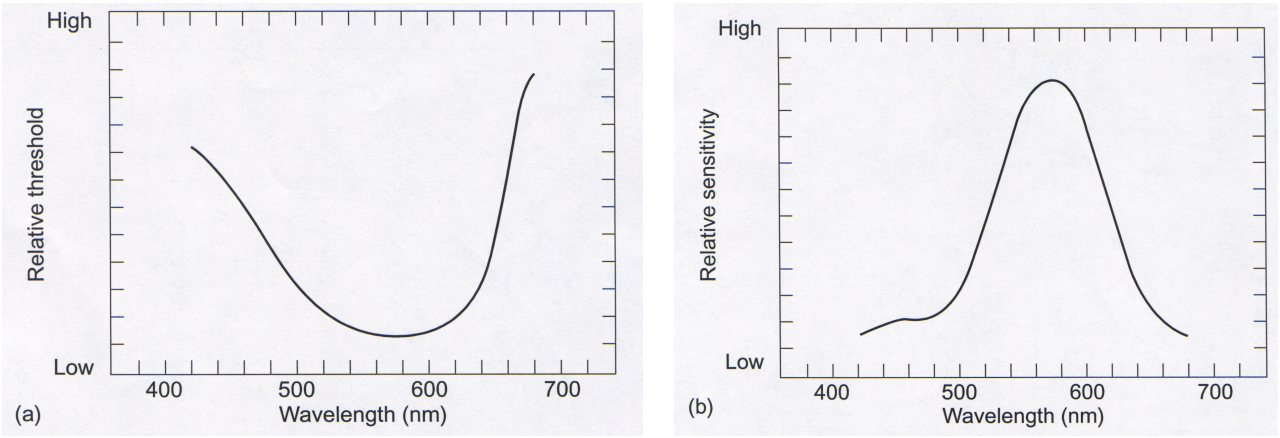



Figure 3.7: Left: The threshold for seeing light versus wavelength for cones. Right: The spectral sensitivity curve for cones obtained by taking the reciprocals of the thresholds for each wavelength in the threshold curve on the left.

Note that the rods are more sensitive to shorter wavelength light than the cones. This means that when we go to low illumination environment (scotopic vision) where rods are more sensitive, we are more sensitive to shorter wavelengths. You may have noticed this shift in the sensitivity. Things appear more bluish in dark, or the green foliage seems to stand out more in the low light of dusk. This shift is called the Purkinje shift, after Johann Purkinje who described the effect in 1825. The spectral sensitivity of the rods and cones are due to the spectral sensitivity of their pigments. The rod-pigment absorbs best near $500\ nm$, near the blue green spectrum.

The cone spectrum shown in the left of Figure 3.8 is the combined effect of the response of three different types of cones, with a peak at $560\ mm$. It has been found that there are three types of cones each having different visual pigments. The response of these three types of cones are shown separately in the right of Figure 3.8. The first, called the S cones, is sensitive to short wavelength and absorbs light best at $419\ nm$. The second, called the M cones, is sensitive to medium wavelength with best response at $531\ nm$. The third. called the L cones, is sensitive

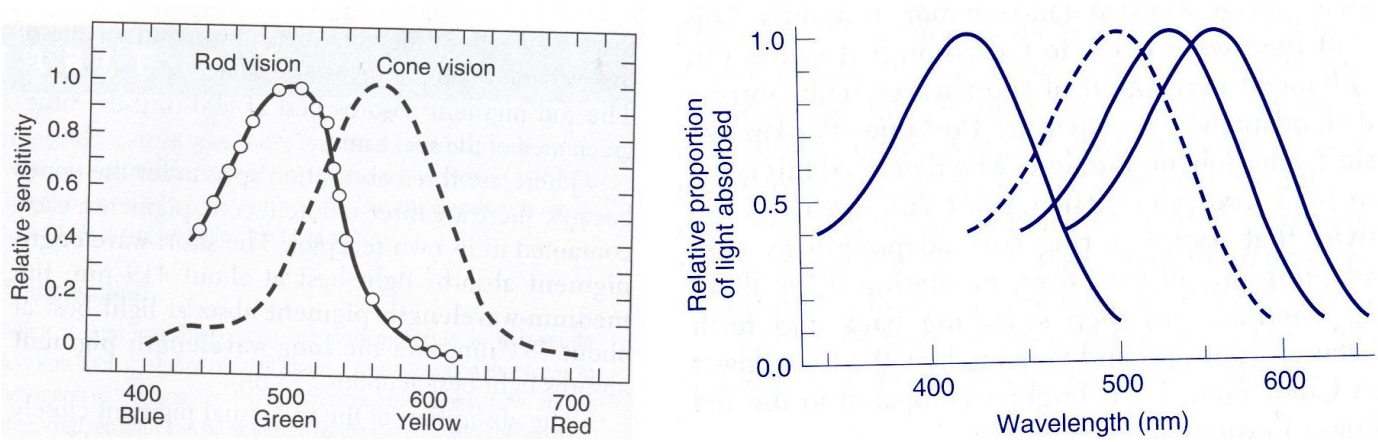



Figure 3.8: Left: The spectral sensitivity curve for rods and cones both. Right: The spectral sensitivity curve of the three different cones.

to long wavelength being most sensitive at $558\ nm$. The combined effect of this may seem inconsistent since it has very little sensitivity in the lower wavelengths while the S cones clearly have strong lower wavelength sensitivity. This the due to the relative numbers of the different cones. The L and M cones are much more in number than the S cones. The relative proportions of the S, M and L cones are approximately 1:6:12. The three sensitivities are weighted by their relative proportions to create the combined response for the cones. We will see in the later chapters that having three types of cones is the key to our color perception. In fact, since the rods are of unified type, they cannot lead to color perception. Hence, in dark when cones are not in their peak sensitivity and most of the vision is due to the rods, we find it difficult to perceive colors.

### 3.1.2 Why three types of cones?

At this point, a question that may cross your mind is why three types of cones? Can't we do with less? The goal of the human visual system is to detect the color of all the monochromatic wavelengths. As it turns out, the three cones that we have are optimal for our color vision.

Lets consider the red cone. Let us say that it produces unit response with 10 photons of 500nm wavelength light. Now, this cone is more sensitive to 560nm and lets say it produces a response of four with 10 photons of 560nm. Note that 40 photons of 500nm will produce the same response and there is no way to distinguish between a higher intensity 500nm wavelength and a lower intensity 560nm wavelength. In monochromatic light we need to distinguished between two parameters, wavelength and intensity. And it is not possible to detect both with just one sensor.

Of course, having two cones will resolve this situation. We can now detect both the wavelength and intensity. This is analogous to the situation where to find $n$ unknowns we need at least $n$ constraints. However, the kind of cones we have in our eye, two of them does not cover the entire visible spectrum. So, we would not be able to detect all the different wavelengths with just two cones. This is exactly the situation in a common kind of color blindness in which the person has only two kinds of cones: regardless of which one of the three pigments is missing there is always some wavelength of light that the person cannot distinguish from other wavelengths or white. (Such subjects are color defective, but certainly not color-blind.) Adding the third cone increases our ability to detect more wavelengths.

At this point, another related question that comes to mind is, can the rods be treated as a fourth type of cone? This is not true since rods and cones have entirely different spatial distribution. You will also see later that the way these connect to the rest of the nervous system through the ganglion cells are to provide complimentary visual

abilities. As a result, rods and cones are almost never stimulated at the same time. While rods are stimulated in low light vision, cones are usually used for normal light vision. For rare situations, even when they are triggered at the same time, the brain does not compare their responses to deduce any information. So, rods are treated as entirely different type of entity than cones.

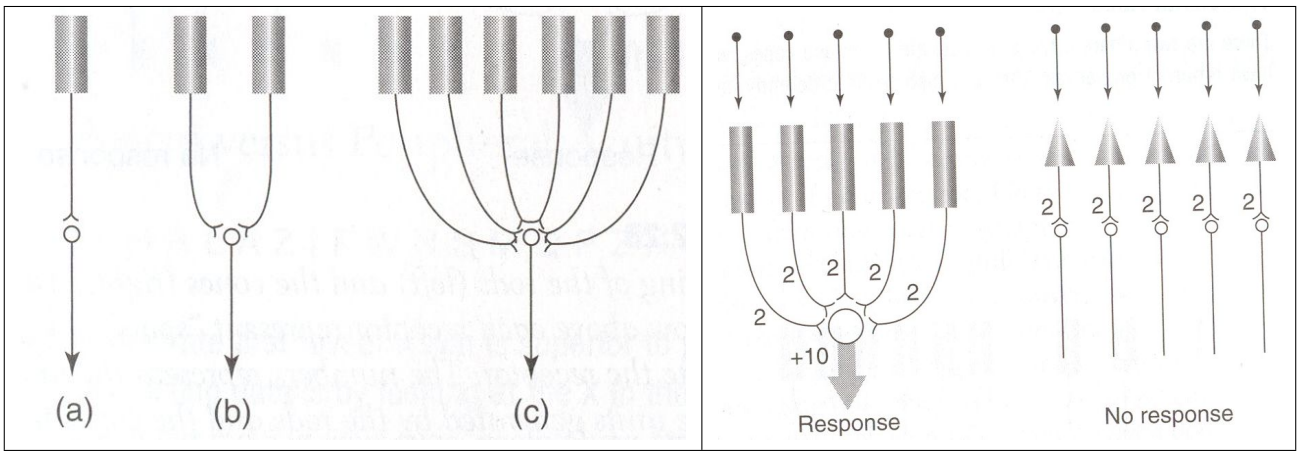



Figure 3.9: Left: Different degrees of convergence. (a) No convergence; (b) Convergence of degree 2. (c) Convergence of degree 6. Right: The wiring of the rods (left) and cones (right). The dot and arrow above each receptor signifies the spots of light stimulating the receptor. The number signifies number of response units generated by the receptors in response to light of intensity 2.0. Note that due to high convergence, the rod ganglion cells receive ten units of excitation and fires. However, the cone ganglion cells receive two units of response and hence does not fire.

**Effects of the Neural Processing on Perception**

We mentioned that some neural processing occurs in the cells of the retina itself. Now, we will see the effects of the neural processing on our perception. When the connectivity of the rods and cones to other neurons in the retina were studied, it was found that they differ in convergence. *Convergence* for a neuron is defined by the number of neurons synapsing it. This is illustrated in Figure 3.9. For the ganglion cells in the retina, there is more convergence of the rods than the cones. There are total of 126 millions of receptors (including cones and rods) in the retina that converge to just 1 million ganglion cells. Of these, 120 million are rods and 6 million are cones. Because of their sheer number, the convergence of rods is much higher. In an average, about 120 rods converge to a ganglion cell, as opposed to 6 cones. This convergence is the most important cause for two things.

1. *The rods are more sensitive in dark.* One reason for the rods to be instrumental for scotopic vision is of course their higher sensitivity. But in addition, it is also true that it takes less light for rods to evoke a response from the ganglion cell. This is illustrated in Figure 3.9. Lets assume that a ganglion cell fires when the stimulation is higher than nine (in hypothetical units). Let us assume that a light of intensity two, produces a response of two units in any receptor. Note that due to higher convergence, the responses from many rods with accumulate at the ganglion cell to create a response of ten units which is higher than the threshold. However, for cones, the stimulus is always less than the threshold of nine and the ganglion cell never fires.

2. *Cones result in a higher resolution vision (more detailed).* You may have often noticed that it is difficult to read in the dark. This reason for this is we do not have enough details or resolution to read in the dark. This ability to see details called *visual acuity* and only an all-cone vision have good visual acuity. One way

to measure visual acuity is to identify the minimum spacing that is required between two dots so that they can be perceived as two separate dots instead of one. Figure 3.10 illustrates how convergence is responsible for this. To understand this, we will assume that the receptors are stimulated by two spots of light, each of intensity ten. The question we ask is under what condition will they be visible as *two* separate spots. First let us consider the case of the two spots stimulating *adjacent* receptors. Note that for the rods, the stimulation to two adjacent receptors lead to firing of a single ganglion cell. This is no different than the case of being stimulated by a single large spot. Similarly, in case of cones, this leads to the firing of two adjacent ganglion cells but this is again no different than adjacent receptors being stimulated by a single large spot of light. However, the case is different when the two spots of light stimulate the first and the third receptor from the right. In case of the rods, still the same single ganglion cell is fired which is no different than stimulation from a single large spot. However, for cones, two non-adjacent ganglion cells fire. The silent ganglion cell in between them is the indicator that two different light spots have been presented. This shows that cones have higher acuity.

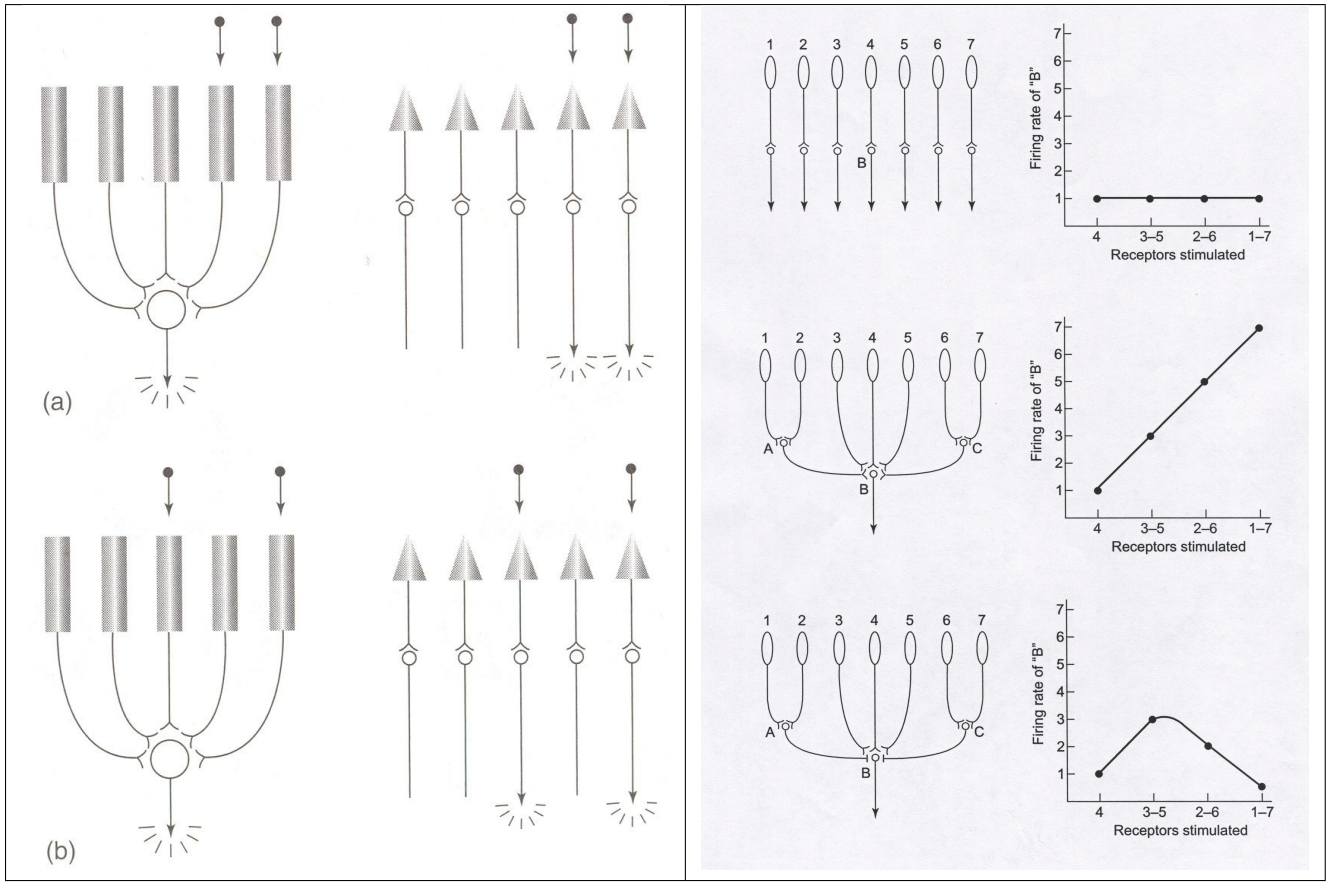



Figure 3.10: Left: Neural circuits for rods (left) and the cones (right). The receptors are being stimulated by two spots of light. Right:

### Neural Circuits and their Effects on Perception

Many neurons thus interconnected through convergence form what we call *neural circuits*. Neural circuits can be small comprising of a few neurons and can also be very large consisting of tens of thousands of neurons. Now,

we will see how the electrical signals are processed by the neural circuits in the retina and how that affects our perception. In the diagram, we represent receptors by ellipses, other neuron cells by circles, excitatory synapses by Y's and inhibitory synapses by T's.

Now, we will explain different neural circuits referring to the illustration in the right of Figure 3.10. We will study seven receptors (labelled $1$ to $7$) with different types of connection to different number of neurons. We will excite these circuits by first stimulating receptor $4$ with a small spot of light. Then we will increase the size of this spot so that other receptors are slowly stimulated one after another. We see the effect of this process on the response of the cell $B$. The first circuit shown has no convergence (each receptor is connected to a different cell) and has only excitatory synapse. In this, stimulating receptor $4$ generates a response in $B$. But with the increase in the size of the spot does not change the magnitude of the response, since the other receptors do not affect $B$ is any way. The corresponding graph illustrates this constant response of $B$. In the second circuit, we add convergence to the circuit so that $B$ now receives inputs from all the receptors. Again, these are all excitatory response. Now, with the increase in the spot size, other receptors connected to $B$ are affected and hence the response generated by $B$ increases linearly as shown in the corresponding graph. We now increase the complexity of the circuit further by adding two inhibitory synapses from $A$ and $C$ to $B$. Now first receptor $4$ is stimulated to create some response in $B$. As the spot of the light increases in size to stimulate $3$ and $5$, the response of $B$ increases due to excitatory synapses. But as the size of the light increases further, the inhibitory synapses reduces the response, generating the corresponding graph shown with the circuit. Thus, we see that the kind of synapse (inhibitory or excitatory) can have great effects on the kind of response generated. Now we will study how this kind of excitatory or inhibitory responses occur in our retina and affects our perception.

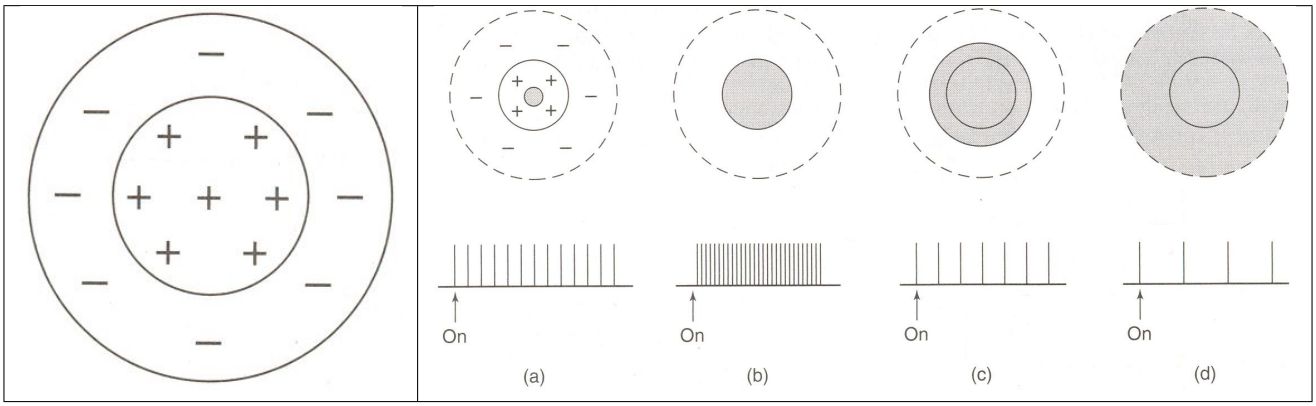



Figure 3.11: Left: The center-surround receptive fields of the ganglion cells in the retina with an excitatory center and inhibitory surround. Right: The response of such a ganglion cell to a spot of light increasing in diameter.

We have seen that a important property of the retina is that many receptors converge to a ganglion cell. These connections actually go through the horizontal and amacrine cells which often generate inhibitory responses. This two together gives the retina properties like the inhibitory-excitatory circuit we just studied above. The *receptive field* of a ganglion cell is defined by the region of receptors which can affect the ganglion cell. Physiological experiments have shown that the ganglions in the retina have a center-surround receptive field with excitatory center and inhibitory surround. This is illustrated in Figure 3.11. The response of such a ganglion to a spot of light increasing in diameter is also illustrated in Figure 3.11. Initially, we get excitatory response. When the diameter of the light is bigger than the excitatory center, inhibition starts and continues to get more and more. This is called the *center-surround antagonism*. This *lateral inhibition* from the inhibitory surround results in several perceptual effects as follows.

The effects of this lateral inhibition called the *hermann grid* is shown in Figure 3.12. Notice the ghost like gray images at the intersections of the "white" corridors. To convince yourself that they are not actually present,

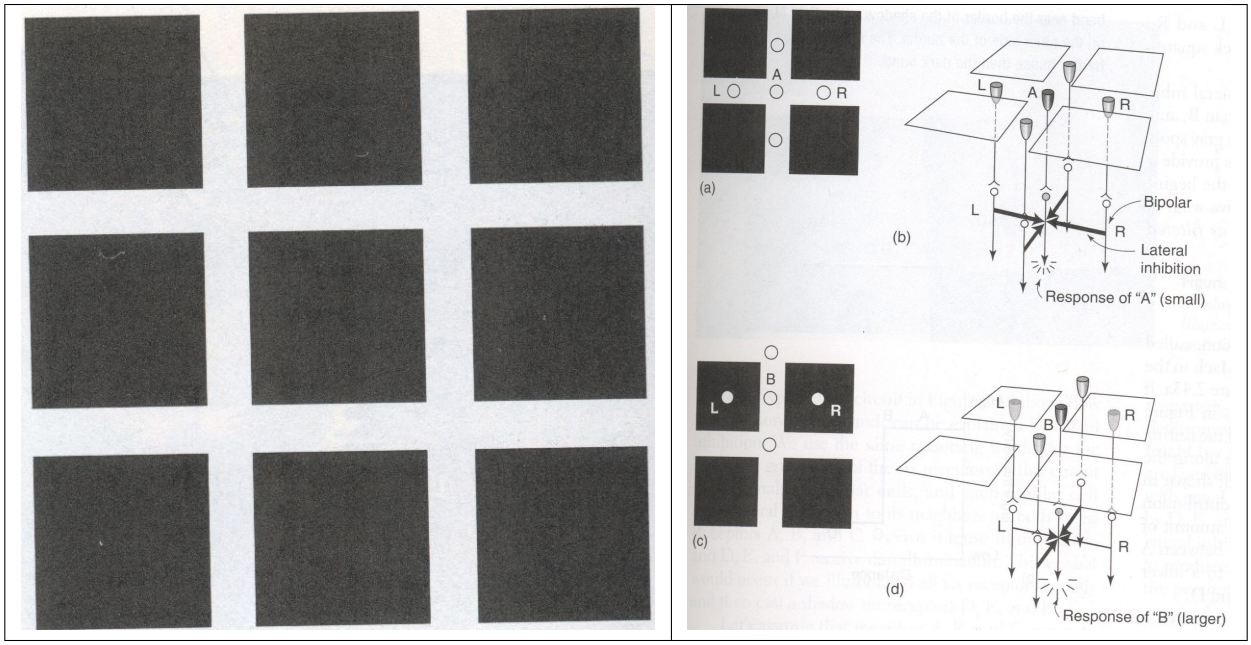



Figure 3.12: Left: Hermann Grid. Right: Explanation of Hermann Grid.

note that they vanish when you look directly at the intersection. This can be explained using lateral inhibition as follows. Let us study a receptor $A$ located at the intersection of the white corridors. Let the four neighbor receptors around it also lie on the white corridors. These receptors send signal to a bipolar cell which send inhibition to $A$. The thickness of the arrow from the bipolar cell to $A$ shows the strength of the inhibition. Note that since these neighboring receptors are stimulated by white, a strong inhibition is send to $A$. The scenario is different for a receptor which is on the white corridor but not in the intersection. Here the neighboring receptors connected to the bipolar cell get stimulated by black instead of white and hence send a 'weak' inhibitory signal. Thus, the intersections look darker than the corridors.

Another illusion called *mach bands* is also created due to this lateral inhibition. This is illustrated in Figure 3.13. If the intensity across the stripes in (a) of the left figure across $A$ to $D$ is measured using a light meter, the response generated is plotted in (b). This shows that the intensity distribution across each stripe is flat. However, when we look at the stripes, we perceive a small light band at $B$ and a small dark band at $C$. These are mach bands. The plot of what we perceive is illustrated in (c). The hypothetical circuit in the right of Figure 3.13 shows how these bands can be explained using lateral inhibition. Each of the six receptors send signal to bipolar cells, which sends lateral inhibition to its neighbor on both side. The strength of inhibition depends upon the strength of the stimulus. Thus, the inhibition caused by the receptors stimulated by the lower intensity is less than those stimulated by the higher intensity causing the mach bands.

The final perceptual effect caused by lateral inhibition that we will describe here is called *simultaneous contrast* and is illustrated in Figure 3.14. This happens when the perception of the brightness of one area is affected by the brightness of the surrounding area. For example, in the left picture of Figure 3.14, the center gray square will show exactly same measurement on a light meter. But, perceptually, the one surrounded by black looks brighter than the one surrounded by the white. The explanation is simple as shown in Figure 3.14. The lateral inhibition for the square with white surround is much larger than the one with black surround and hence it looks darker.

Till now we studied receptive fields with center excitation and surround inhibition. These are called on-center ganglion cells. Recent studies show that an off-center ganglion cells (center inhibition and surround excitation) is equally likely to occur, often even in the same spatial location fed by the same photoreceptors. These result in the

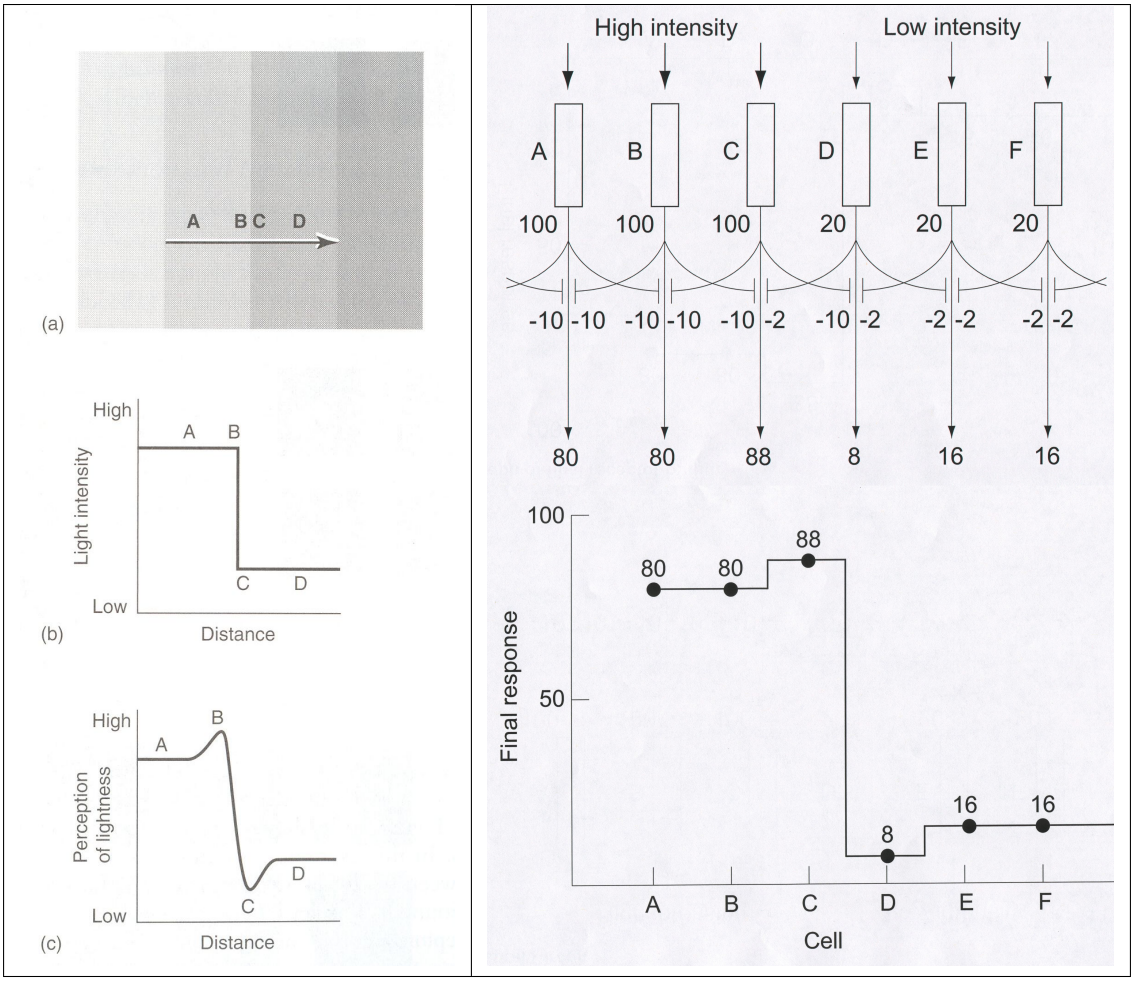



Figure 3.13: Left: Mach Bands. Right: Explanation of Mach Bands.

enhancement of the systems dynamic range.

Note also that if the excitatory and the inhibitory areas of the ganglion cells are balanced, then these cells will not respond to uniform light. This illustrates one aspect of image compression carried out in the retina. The brain is not bothered about the redundant visual information, only the information about changes is transmitted.

## 3.2 Application of Lateral Inhibition

The knowledge of lateral inhibition has helped us avoid mach bands artifacts in many graphics and image processing applications. First, it has lead to the theory of different kinds of *continuity*. If two curves, $A(t)$ and $B(t)$ join together at any point $p$ with the constraint that the value of both the curves at $p$ is identical, that means the combined curve will have a $C^0$ continuity. Mathematically, $A(t)$ and $B(t)$ are said to have $C^0$ continuity if

$$A(p) = B(p)$$

If two curves are joined at $p$ with the constraint that the gradient of the two curves are identical *both in direction and magnitude*, that means the combined curve will have $C^1$ continuity. Mathmatically, $A(t)$ and $B(t)$ are said to

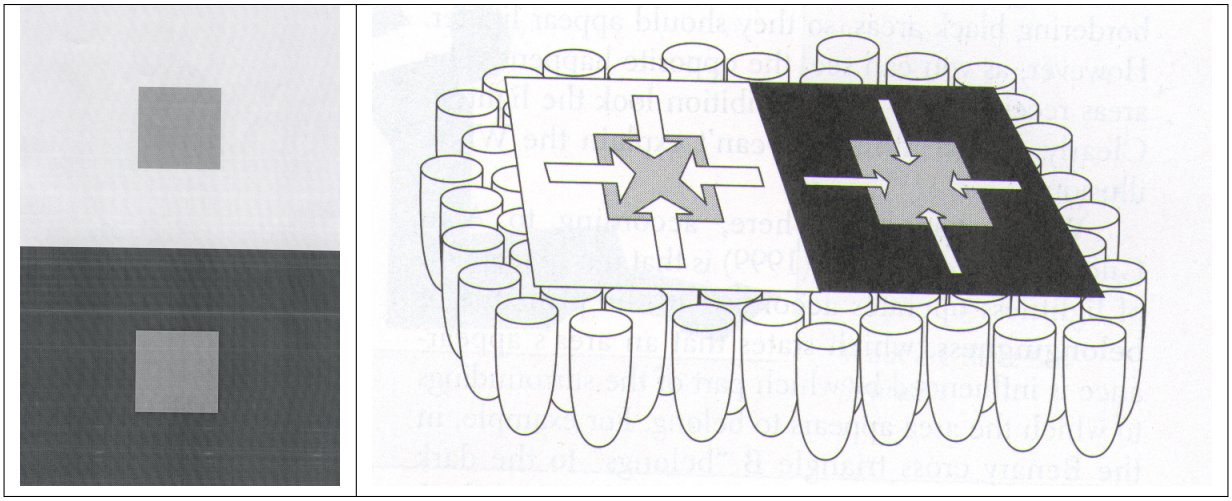


Figure 3.14: Left: Simultaneous Contrast. Right: Explanation of Simultaneous Contrast.

have $C^1$ continuity at $p$ if

$$\frac{dA}{dt}_{t=p} = \frac{dB}{dt}_{t=p}$$

A less restrictive goal of geometric continuty is described as, if the direction of the gradient of the curves $A$ and $B$ are identical at $p$ they have $G^1$ continuity. Note that in this case the magnitude of the gradient can be different for the two curves at $p$ and hence this is a less restrictive constraint than $C^1$. Similarly, the curves are said to have $C^n$ continuity at $p$ if the $n^{th}$ derivative is identical at $p$. Mathematically,

$$\frac{d^nA}{dt}_{t=p} = \frac{d^nB}{dt}_{t=p}$$

Correspondingly, $G^n$ is defined as the direction of the $n^{th}$ derivative being identical at $P$. For example, $C^2$ means that the curvature of curves are similar at $P$. This is illustrated in Figure 3.15.

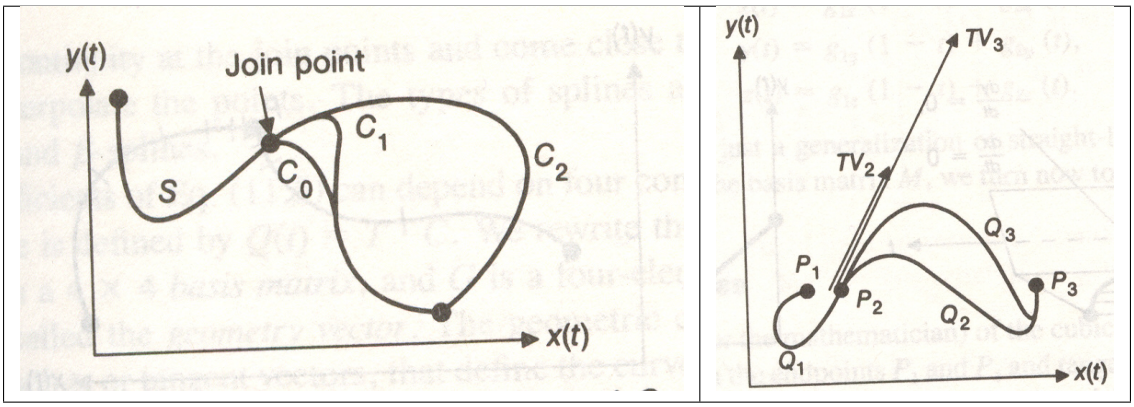



Figure 3.15: Left: This shows the $C^0$, $C^1$ and $C^2$ continuity. Right: The demonstrates the different between the $G^1$ and $C^1$ continuity. For curves $Q_1$ and $Q_2$, the gradient (given by the tangent $TV_2$) at $P_2$ is identical (both in magnitude and direction). Hence they are $C^1$ continuous at $P_2$. But the magnitude of the gradients of $Q_1$ ($TV_2$) and $Q_3$ ($TV_3$) are different at $P_2$, but their directions are same. So $Q_1$ and $Q_3$ are $G^1$ continuous and not $C^1$ continuous.

Mach bands are formed when the slope of the curve is discontinuous. Or in other words, piecewise curves are joined at a point where their gradients are not identical. So, any curve that does not have $C^1$ continuity will show mach bands artifacts. The same concept can be extended to 3D. The same concept can be extended to 2D. That is why in many applications, as we will see next, $C^1$ continuity is assured to avoid mach bands artifacts.

### 3.2.1 Shading

Both diffused and specular lighting has been modeled in computer graphics.

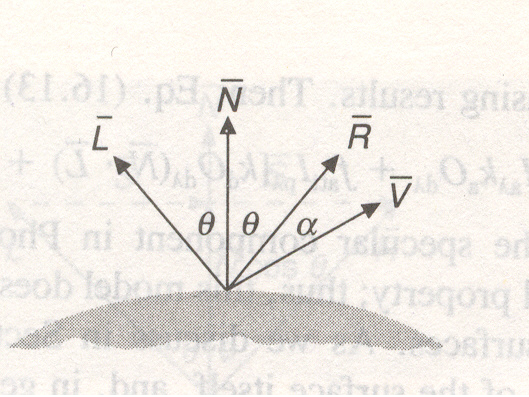



Figure 3.16: The light vector, normal to the surface and the view vector at any point on the model.

The diagram on the left shows the 3D normal at the surface $\bar{N}$, the light vector $\bar{L}$ and the view vector $\bar{V}$ at any point, say $p$ of the model. The diffused lighting at $P$ is given by

$$I_d = k_d(\bar{N} \cdot \bar{L})$$

Note that since this is diffused lighting, this does not depend on the view vector $\bar{V}$.
The specular lighting at any point $p$ is given by

$$I_s = k_s(2\bar{N}(\bar{N} \cdot \bar{L}) - \bar{L}) \cdot \bar{V}$$

Note that as expected, the specular lighting depends both on the light vector $\bar{L}$ and the view vector $\bar{V}$. Ideally, when rendering a triangular mesh, the lighting should be computed at every pixel of the triangle. But this becomes too expensive computationally due to two reasons: (a) evaluating the normal at every pixel, (b) evaluating the illumination model at every pixel. Thus interactive rendering rates are lost. So, the first approximation to this was *flat shading*, where the illumination was calculated at one location on the triangle, usually the average of the color at the three vertices, and then this color was used to paint the whole triangle. This does not have $C_1$ continuity at the edge of the triangles and hence shows Mach bands. To remove these artifacts, the illumination was evaluated at the three vertices and the linearly interpolated from these in the interior of the triangle. This is called Gouraud shading. This follows $C^1$ continuity and hence removes the artifacts. This is illustrated in Figure 3.17.

### 3.2.2 Intensity Blending

In several applications like panoramic image generation and multi-projector displays, intensity across image boundaries needs to be smoothed out. This is called *intensity blending*. This concept is explained in Figure 3.18. The blending quality depends on two factors: (1) the blending function used, (2) the width of blending region. As the width is increased, blending quality becomes better as shown in Figure 3.19. This is because, with increasing width, the $C^1$ discontinuity of the image decreases and the mach band effect decreases. Also, as smoother and smoother functions are used, the $C^1$ discontinuity reduces again giving better perceptual effects.

In fact, for a plethora of applications like image editing, contrast compression, tone matching etc, wherever there is the need of smoothing intensities, functions that maintain $C^1$ or $C^2$ continuity are used. Often $C^2$ continuity is preferred for better appearance. But, $C^1$ continuity is the minimum that you need for artifact free smoothing.

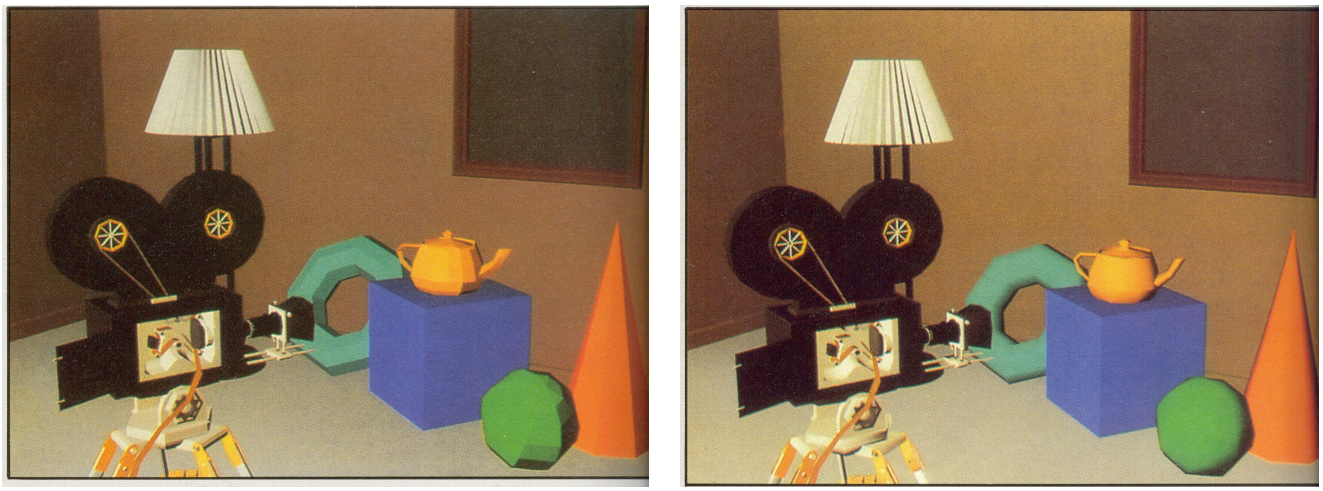


Figure 3.17: Left: Mach Bands for diffuse Flat Shading. Right: The artifacts dissappear for diffure Interpolated Shading (Gouraud Shading)

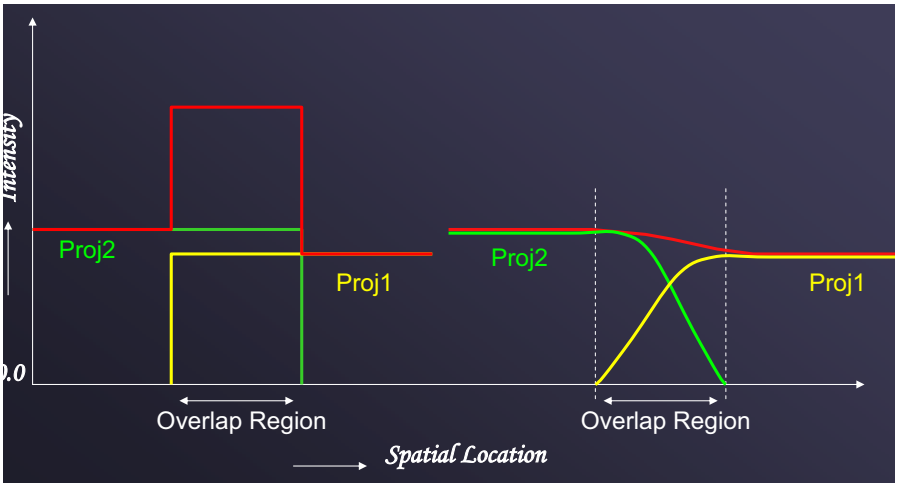



Figure 3.18: Blending in the overlap region between two projectors.

### 3.2.3 Geometric Modeling Primitives

The goal of modeling is to represent continuous 3D geometry using discrete 2D/3D approximations. Since most of the time we are interested with the surfaces of the objects rather than the volume, mostly we are interested in modeling the continuous surface of 3D world with 2D discrete patches. To aid smooth artifact-free modeling, patches like Bezier and Hermite patches, are used that follow the three following properties.

(a) The patches are at least $C^1$ continuous.

(b) When subdivided, the children patches generated are also at least $C^1$ continuous.

(c) Any operation on the patches would maintain $C^1$ continuity.

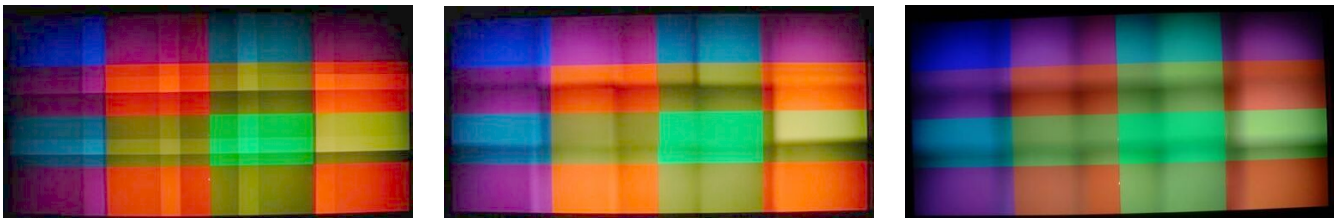


Figure 3.19: Multi-projector display before blending (left), after blending with $50$ pixels overlap (middle), after blending with $125$ pixels overlap (right).

## 3.3 Lateral Geniculate Nucleus (LGN)

The receptors in the retina get excited with the light entering the eye. They transduce the environmental energy into electrical energy which is then transmitted towards the lateral geniculate nucleus and then the brain. Now we will study what happens in the lateral geniculate nucleus (LGN).

As shown in Figure 3.1, the optical nerves from the two eyes first travel to the LGN. As we found in case of the ganglion cells in the eye, the neurons of the LGN also have a center-surround organization of receptive field. This means that the signal from the eye does not merely flow through the LGN but the LGN plays an important role to *regulate* the flow of information to brain. This flow is shown in Figure 3.20. $90\%$ of the optical nerves go to the LGN while another $10\%$ go to the superior colliculus. The superior colliculus is responsible for controlling eye movements and needs feedback information. Thus, this is an example how action is influenced by the perceptual process. As shown in Figure 3.20, the LGN receives signals not only from the eye, but it also receives *feedback* signals from areas of the brain like the visual cortex, from the thalamus (T) and also from the LGN itself. In fact, the flow of information summarized in the right image of Figure 3.20 shows that (1) the amount of information coming back to the LGN from the brain is significant, (2) in fact, the smallest of all the flows is the flow from the LGN to the cortex. So LGN is said to regulate the information flow to the brain significantly.

### 3.3.1 Organization of Information

As we will see throughout this course that the organization of information in the brain is a very important factor. In fact, this organization is largely responsible for our efficiency in any search operations like locating a book in the library or searching for a friend in the crowd. This is analogous to good database storage system. The efficiency organization of any storage system is almost always evaluated by how fast the data can be retrieved. As we will find that organization in the human system is extremely efficient.

The organization of information starts from the eye itself. This is the reason that the regions that are adjacent in the scene are also adjacent in the retinal image. That is, the image formed is coherent and different parts of the scene does not get images at different random location in the retina. This may seem obvious, but this is where the first step of organization starts.

#### Organization by Left and Right Eye

The LGN is a bilateral structure, i.e. there is one LGN is each of the right and left hemispheres of the head. Viewing the LGN in cross section reveals six layers, as shown in Figure 3.21. Each layer receives signal only from a single eye. The *ipsilateral* eye (the eye in the same side of the head as the LGN) sends signals to layers $2$, $3$ and $5$. The *contralateral* eye (the eye is the opposite side of the head as the LGN) sends signals to the layers $1$, $4$ and $6$. Thus, each eye sends half of the signals to the left hemisphere LGN and the other half to the right hemisphere LGN.

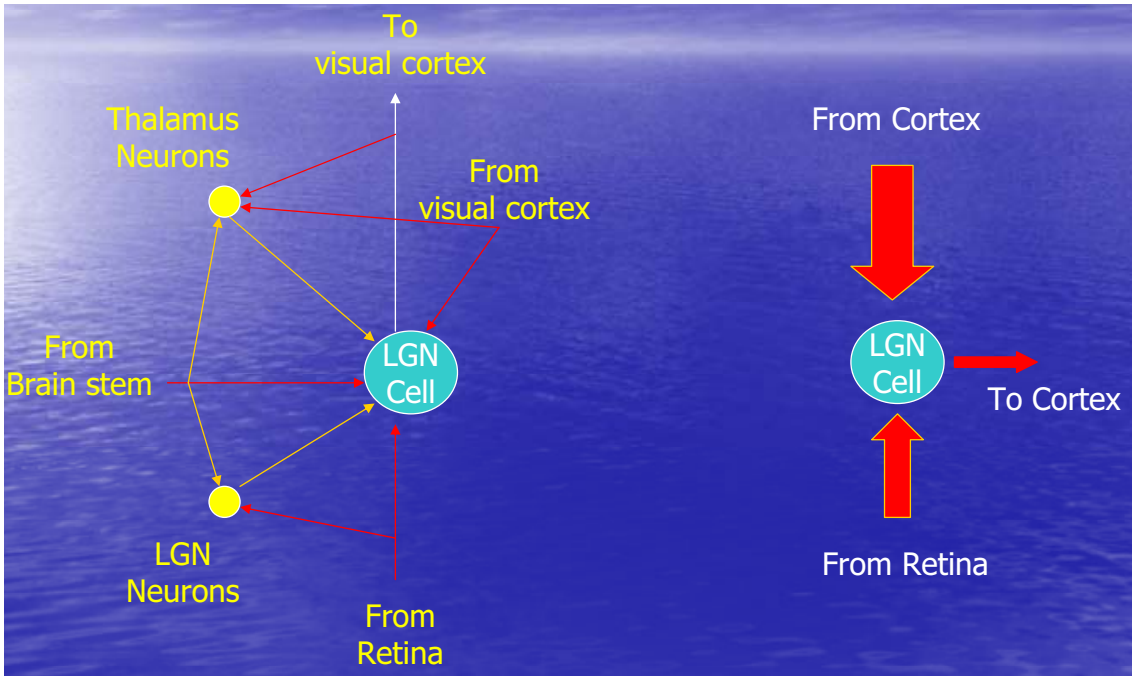



Figure 3.20: The regulation of information flow in the lateral geniculate nucleus.

**Organization as a Retinotopic Map**

The neurons entering the LGN are arranged in such a fashion that the same area in the retina end up in the same area of the LGN. Retinotopic means that each location on the LGN correspond to a location in the retina and the adjacent locations in the LGN are also adjacent in the retina. Thus, this creates a map of the retina in the LGN called the retinotopic map. For example, adjacent neurons $A$, $B$, $C$, and $D$ in the LGN are signals from the neurons $A'$, $B'$, $C'$ and $D'$ in the retina, as shown in Figure 3.21. This happens for all the six layers in the LGN. Thus, if we draw a line through the layers as shown in Figure 3.21, the areas in the LGN have signals from exactly the same region of the retina.

**Organization by Types of Ganglion Cells Arriving at LGN**

Three kinds of ganglion cells send information from the retina to the LGN.

1. *P-cells or parvocellular cells:* These cells have small or medium sized cell bodies. They respond to sustained stimuli with sustained firing. These cells send signals to the layers 3, 4, 5 and 6 of the LGN.

2. *M-cells or magnocellular cells:* These cells have larger cell bodies and respond with brief bursts of firing. These cells send signals to the layers 1 and 2 of the LGN.

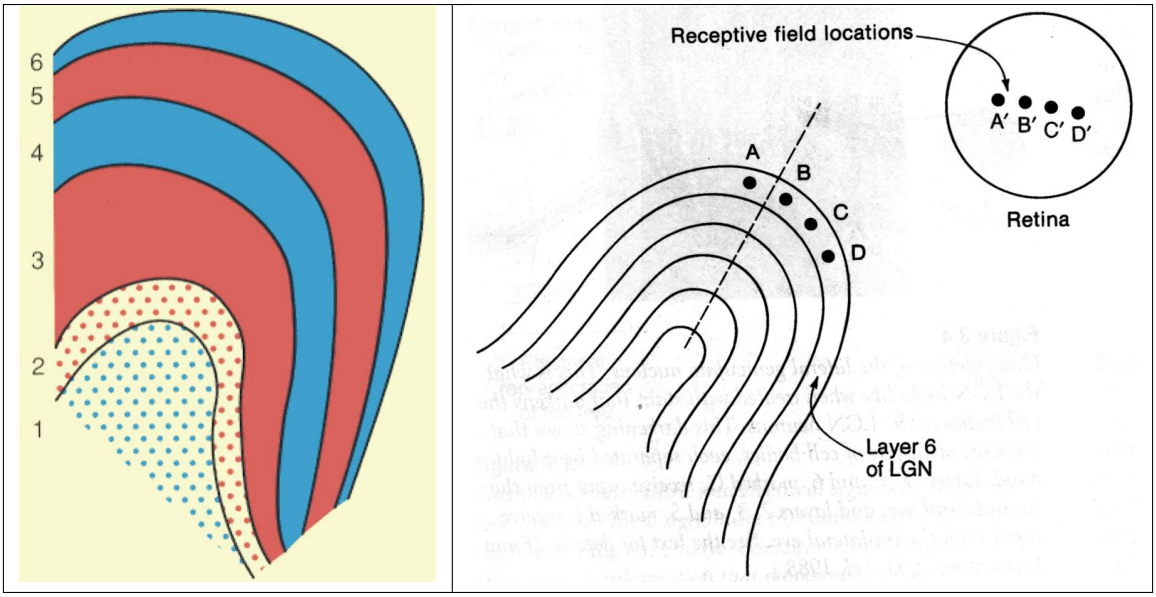



Figure 3.21: Left: Cross section of the LGN showing the layers arranged by the left and right eye. Right: Retinotopic Mapping of the neurons in the LGN.

3. *K-cells or koniocellular cell:* These cells very large cell bodies. These cells send signals to all the six layers of the LGN. However their functions are yet to be deciphered.

Since the layers 1 and 2 of the LGN receive signals from the M-cells, these are called the *magno* layer. For similar reasons, layers 3, 4, 5 and 6 are called the *parvo* layers. It has been found that the magno layer is responsible for the ability to detect motion and the parvo layers are responsible for the perception of color, textures, patterns and depth of small or finely detailed objects. As we will see, that these two different kinds of layers sends two different kinds of signals to the visual cortex. The final organization of the LGN is summarized in Figure 3.22.

## 3.4 Striate Cortex

One and a half million axons travel from the LGN to the visual striate cortex. The visual cortex is far more complex than LGN containing about 250 million neurons. The functions of the cells of the visual cortex were first discovered by David Hubel and Torsten Wiesel in 1950. They received Noble prize for their contribution in this direction. They found that the cells in the striate cortex respond to specific aspects of stimuli like orientation, motion and edge. Because of this ability to respond to very specific features, the cells in the striate cortex are often called *feature detectors*.

There are three types of cells found in the striate cortex.

1. **Simple Cortical Cells:** These cells respond to specific orientation. Different sets of simple cells respond to different direction. For example, the response of one kind of simple cell to different orientation is shown in Figure 3.23. Note how the peak sensitivity if for vertical orientation. This shows that this particular kind of cells are sensitive to vertical orientation. Different kinds of such simple cells together respond to all different kinds of orientations. These cells have a receptive field similar to that of center-surround, but it is not circular in shape. So they act more like edge detectors or line detectors. Detector of each orientation can again be of four types depending on their receptive fields, as shown in Figure 3.24. In fact, more recent studies show that they may not have a simple center surround receptive field as believed by and Weisl. Many

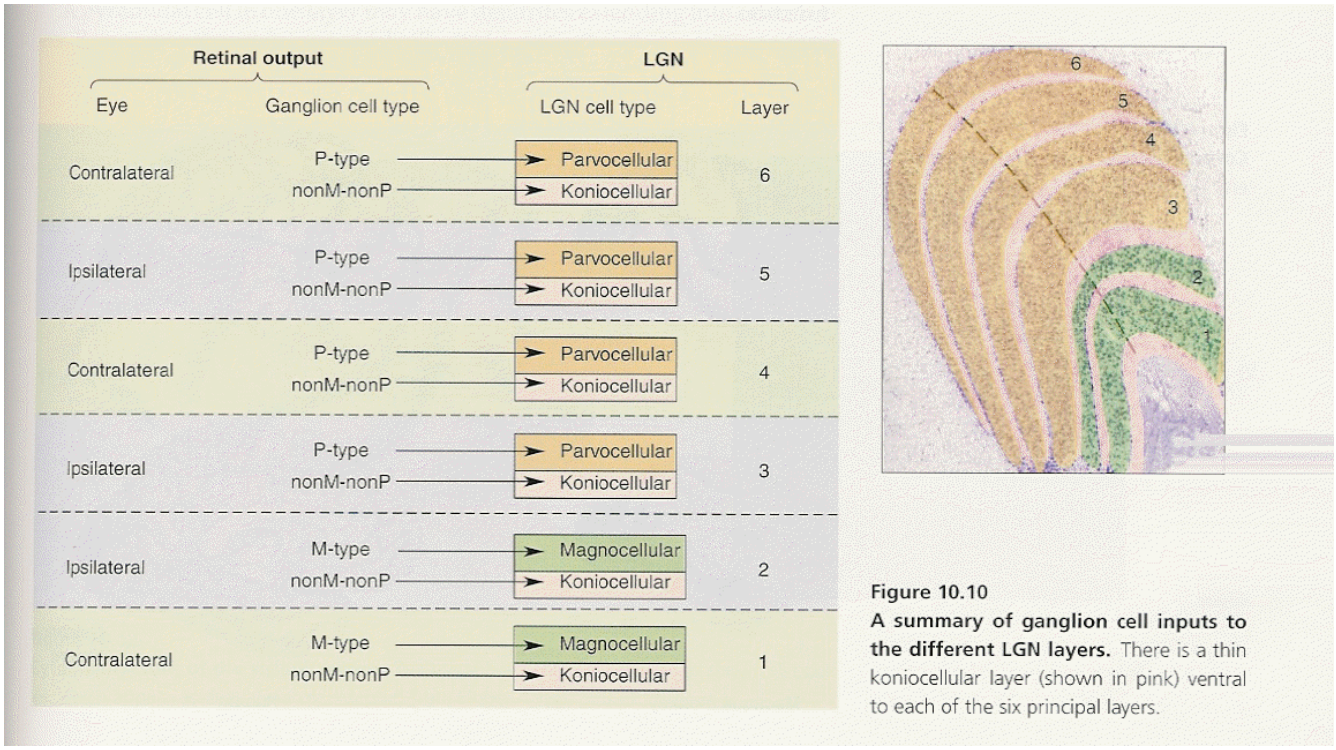



Figure 3.22: Summary of Organization in the LGN.

of them have smaller lobes of excitatory and inhibitory receptive fields by the side of their primary lobes as shown in Figure 3.24.

Also, these cells undergo what is called *selective adaptation*. This shows that if the simple cells are adapted to a particular orientation for a long time, they adapt to that orientation and become less sensitive to that orientation.

2. **Complex Cortical Cells:** These cells like simple cells also respond to various orientation but in a very different manner.

   - *Nonlinearity:* These cells are highly non-linear and hardly ever responds to small stationary spots. So it is difficult to find their receptive field by the standard stimulus.
   - *Motion Sensitivity:* These cells are highly responsive to *moving* lines and edges anywhere within their receptive field. Often, they are also sensitive to the direction of motion.
   - *Position Insensitivity:* They are not very sensitive to the position of the stimuli in their receptive field. Thus, small difference in the position of the stimuli will not affect their response.
   - *Spatial Extension:* Complex cells usually have larger receptive fields than the simple cells.

3. **End Stopped Cells:** The end stopped cells are even more specific in their receptive field. They respond to moving lines and corners of specific sizes. That is the reason they are named as end stopped cells. Recently, scientists are more inclined to believe that these are more hypercomplex simple or complex cells rather than forming a class of their own.

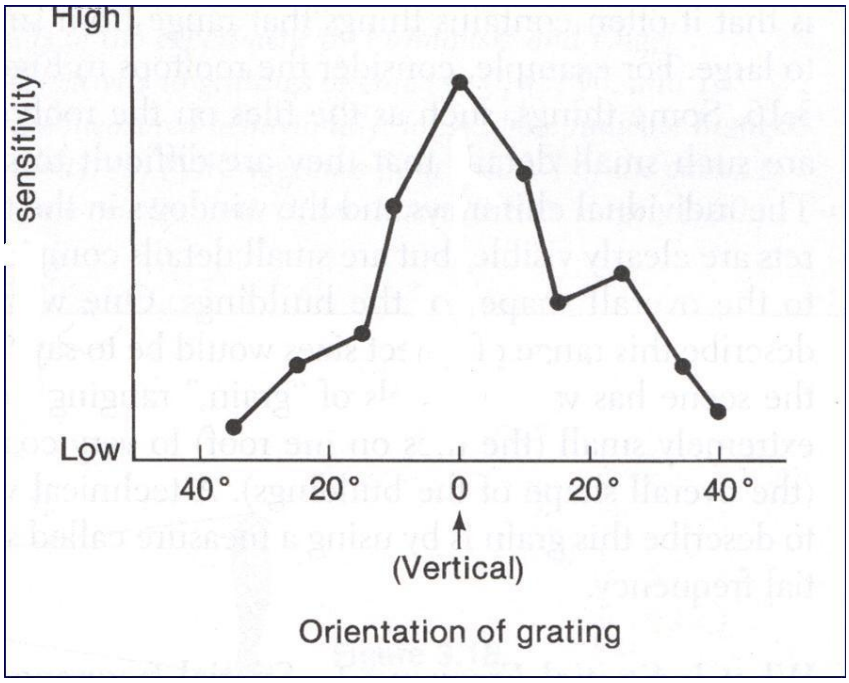



Figure 3.23: The sensitivity of a simple cell sensitive to vertical orientation.

### 3.4.1 Organization

Just like the LGN, an important question here is what is the organization of the striate cortex. In this section we will see that the cells are organized in the striate cortex in the form of a hypercolumn.

Just as in LGN, there is a retinotopic map of the retina on the cortex. So each point in the retina corresponds to a point on the cortex. However, the most important feature of the map is that the area representing the fovea is much larger on the cortex. In the retina, the fovea occupies just $0.01\%$. But its retinotopic map on the cortex occupies $8 - 10\%$. This is called the cortical magnification factor.

The cortical magnification factor is due to two reasons.

1. First, the density of the receptors are much higher in the retina than in the fovea. This is also true for the ganglion cells. For example, in the fovea there is about $50,000$ ganglion cells per square millimeter and only about $1000$ per square millimeter in the periphery. This mismatch in density is much less in the cortex. In fact, the density of the neurons receiving signals from the fovea and the periphery in the cortex is close to uniform.

2. Second, each foveal input is allotted extra cortical neurons. So, a ganglion cell from or near the fovea is allotted three to six time more cortical cells than the ones from the periphery. In addition to the lower convergence of the cones to the ganglion cells, this is also another reason for higher acuity vision in the fovea.

As a result of the retinotopic map, a columnar organization is found on the cortex. This shows that the cells on the same column of the cortex are stimulated by the same area of the retina and the adjacent columns are stimulated by adjacent areas on the retina.

It was also found that the cells in each column of this columnar organization are stimulated by different orientation. Thus orientation columns are formed perpendicular to the retinotopic columns.

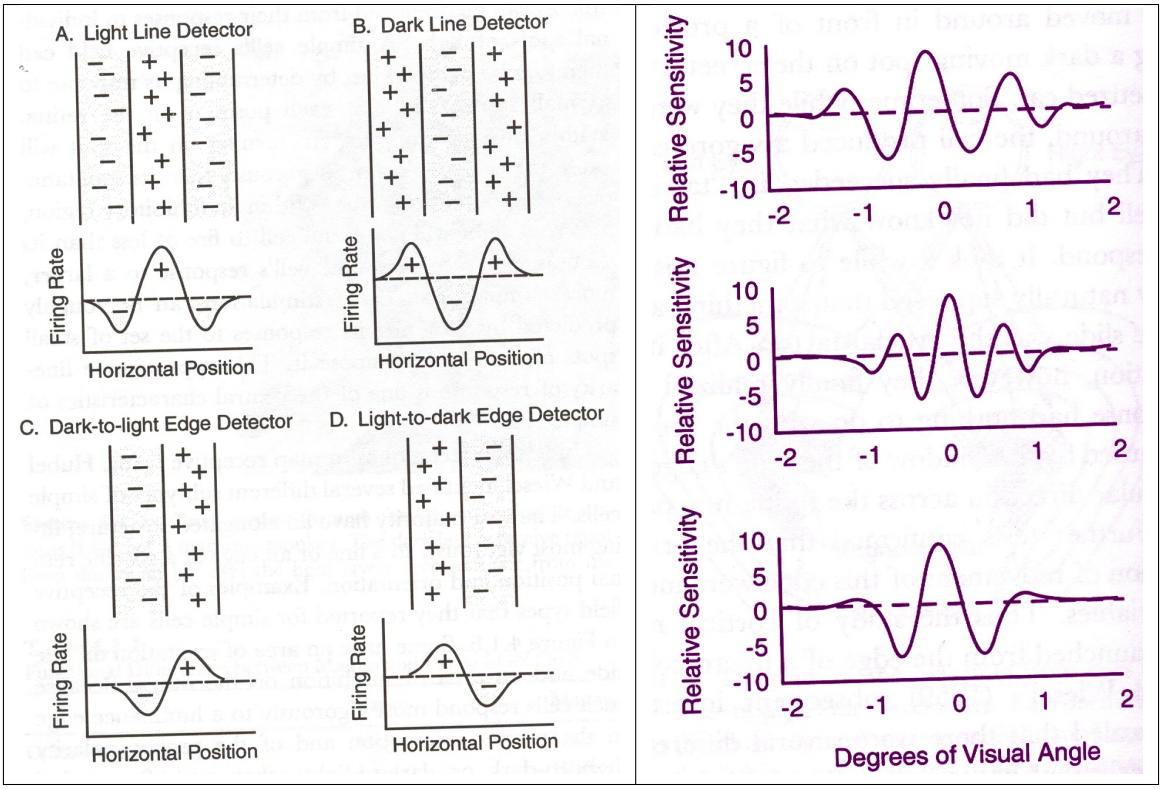



Figure 3.24: Left: Four different types of simple edge detectors for vertical orientation. Right: Recent studies show a more complex receptive field with secondary lobes by the side of primary ones.

Third, columns of cells which have a ocular dominance are also found. This means cells along a column are more sensitive to signals from one eye than from the other.

All these suggest a hypercolumn structure of the cortex as shown in Figure 3.25. Figure 3.25 also shows how the retina information is mapped on this hypercortex when viewing an object like a tree. The three areas of the trunk from the retina maps to three columns of the cortex, $A$, $B$ and $C$. Since this comes from one of the eye, it maps to one of the ocular dominance column within this column. Next, since the trunk of the tree is of vertical orientation, it only stimulates a column of cells within this ocular dominance slab.

### 3.4.2 Development of Receptive Fields

The presence of this highly developed cell in the cortex of the brain immediately raise the question, 'Are these cells present from birth or do they develop with age? If they are present from birth, do they get better with learning?' Visual deprivation experiments performed on kittens to test this. A set of kittens were reared with their eyes shut from their childhood. This is called selective rearing. Their sensitivity to orientation were then compared with that of normal kittens. It was found that these visually deprived kittens still had orientation sensitivity, but this was not nearly as good as the normal ones. This shows that these cells do exist from birth but probably develop a lot with learning.

### 3.4.3 Application to Edge Detection Algorithms

From the above studies it is evident that detecting edges in an image is probably one of the most basic function that the human visual system performs. So, efforts are being made for a long time to find computational approaches

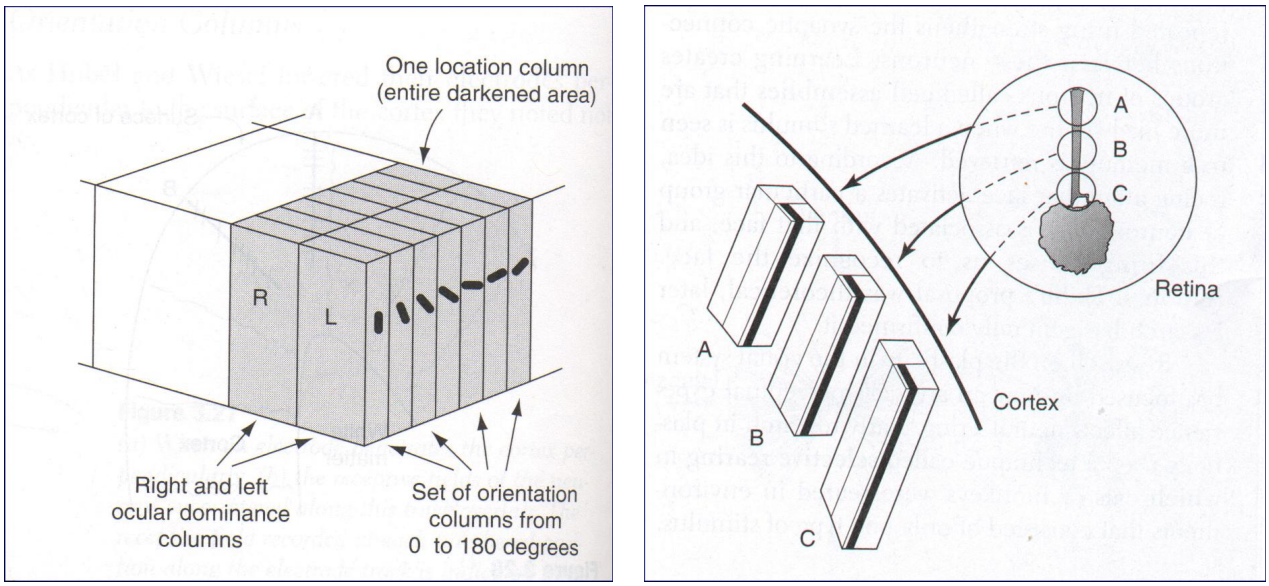



Figure 3.25: Left: Hypercolumn view of the cortex. Right: Mapping of the information from the retina on to this hypercolumn.

to detect edges. Early attempts to detect edges were purely unsystematic. They ended up trying different kinds of *edge operators* on images and check which one works the best. Edge operator is an computational scheme for detecting the likelihood of locating an edge in gray-scale images over a certain neighborhood. The operation is called *convolution* of an edge operator with the image. The edge operator is moved on every pixel and the weighted sum of the image using the weights in the operator is found. This is illustrated in Figure 3.26. This shows a simple vertical and horizontal edge operator acting on a gray scale image. After this operator is applied, certain features of the image are detected which signify the edges. For example, in this case, highly positive or negative value in the image signifies a horizontal or vertical edge. Also note that convolution is an operation that is absolutely parallelizible.

Studying the horizontal and vertical edge detector, one would realize that all it is doing is finding the gradient or slope of the intensity in $x$ and $y$ direction at every pixel. This is equivalent to finding the first derivative in the continuous domain. Since they are simple operators that find the slope in pre-defined directions, they cannot detect the edges in all direction. In fact, there are several methods that try to find the derivative of the image in different directions to detect the edges. These are called *first derivative operators*. However, note that to design operators that would detect edges of all orientations *comprehensively* would be very difficult to design. So, more complicated *second order differential operators* were designed to detect edges. The most popular of these is the Marr-Hildreth's zero crossing algorithm that has strong perceptual basis. Several of these operators are shown in Figure 3.26.

### Marr-Hildreth's Zero Crossing Method

They first tried to charactarize the intensity profile that would lead to an edge that the algorithm would try to find. Such a luminance profile is graphed in Fig 3.27. Next, they studied the slope of this luminance profile is computed using first derivative. Finding the maximum of this slope would give us the location of the edge. But note that these operators are direction dependent i.e. an operator for vertical edge cannot detect a oblique or horizontal edge and vice versa. However, if the derivative of the slope is taken, which is basically the second derivative of the intensity profile defining the curvature of the surface, it crosses zero at the edge. Note that, whichever direction the edge

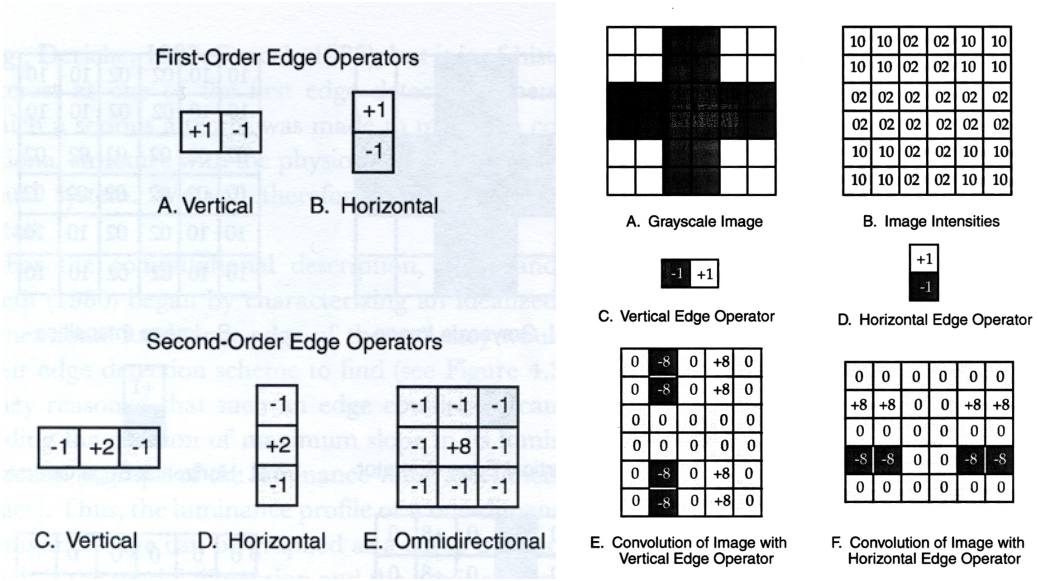



Figure 3.26: Left: Different kinds of edge detectors. Right: Convolution of an image with an edge detector.

is from, having a zero surrounded by a high value and a low value would mean an edge. Also, since the second derivative is symmetrical about the midpoint, it was easy to construct an omni-directional second derivative edge operator that would produce zero crossing at edges of all directions. Also, since all the edges operators can be collapsed into one, this indicates a computational efficiency not offered by the other edge detectors. Note another important fact about this marr-hildreth edge operator. In the continuous domain this edge detector would be similar to the receptive field of the eye and the different other cells we see in the visual pathway as shown in Figure 3.27. Figure 3.28 shows the result of this edge detection algorithm.

### Neural Implementation

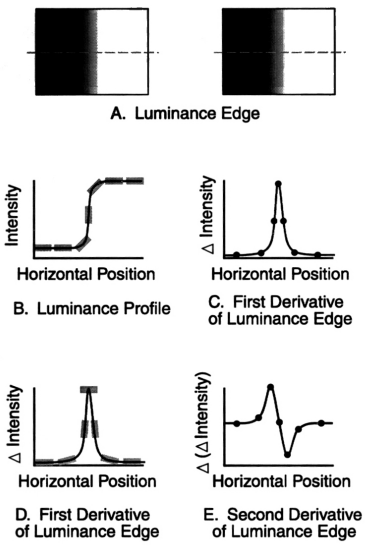
Marr-Heldrith not only gave an algorithm for edge detection, but also tried to develop a model for how the visual cells can achieve this kind of processing. They predicted the presence of three types of cells.

- First, there should be cells that do the convolution.
- Second, there should be cells like those of $P$ and $Q$ in Figure 3.29 that fire with the maxima and the minima. Such adjacent cells will have receptive fields that are complement of each other.
- Finally, they predicted that their should be a third kind of cell that has an AND response to the firing of adjacent $P$ and $Q$ kind of cells. These are called the zero detectors. Edge detectors sharply tuned to certain orientations can then be formed by combining many aligned zero detectors.

Interestingly, Hubel and Weisl found the existance of all these kinds of cells in our visual pathways. The lateral inhibition in the ganglion cells perform the convolution. The simple cells in the cortex respond to maxima and minima of the signals send from ganglion cells to them. Finally, the complex cells have been found to act like zero detectors. Thus, there are evidence that this probably the algorithm our visual system applies to detect edges.

### Scale Integration

However, one complication of the computation theories of edge detection methods is that these edges occur at different scales and size. Some are slow changes over broader regions, while others are sharper changes over



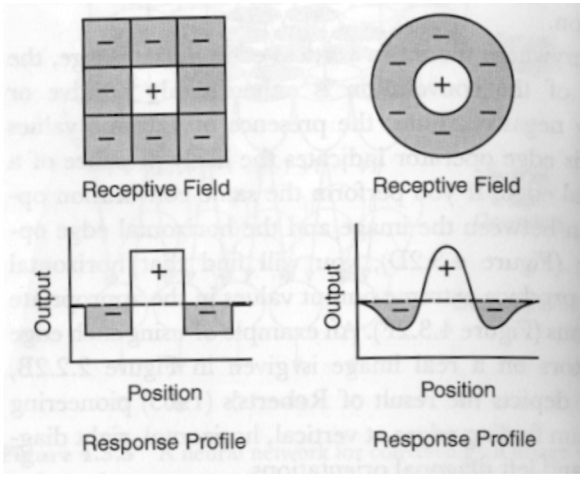




Figure 3.27: Left: Edge detection by detecting zero crossing of the second derivitive of the luminance profile. Right: Discrete and continuous version of the omnidirectional edge operator.

smaller areas. They are all edges nevertheless and has to be detected. An example of this is shown in Figure 3.30. Here the edges are detected at different scales. This is done by applying the edge operator to different resolutions of the image from high to low. At the lowest resolutions, only the gross edges will be detected while for the higher ones, finer edges will be detected. Also note that the edges detected in the lower resolution do not dissappear in the higher one. Only finer edges get added. Often, the edges which are detected in the lowest are the most important.

An important problem that should be mentioned here is called *scale integration*. The size of image changes with different resolution. So how do we match the gross edges with the same edges found in the detection of finer edges in the higher resolution detection. This has been addressed by Witkin's scale space algorithm. He showed that there is a virtual continuum of the scale sizes through which the edge information can be followed. As as we go from coarser (top) to finer (bottom) edges in this space, new edges and continue through this continuum. This is illustrated in Figure 3.31.

Since humans can do this scale integration pretty easily, this would mean that the human visual system would show a comprehensive representation of sizes. This is supported by experiments which show that the visual receptive fields of some cortical cells show dense respresentation of sizes.

### 3.4.4 Higher Level Processing

Hubel and Wiesel won the Noble prize for their discovery of the simple, complex and the end-stopped cells in the visual cells in 1960. But, several studies being carried out near this time started to show that far more complex processing is done in the visual cortex. In fact, by late 1970s a new era started with research showing that a large part of extra striate cortex also involved in visual perception. Though complete information is still not available, we will study here all that has been found since then.

#### Higher Level Processing in the Straite Cortex

So far, we have seen cells that respond to lines and edges. But, recent research shows that there are cells in the visual cortex ($V1$) that have a property for *contextaul modulation*. This means that the stimulation response of

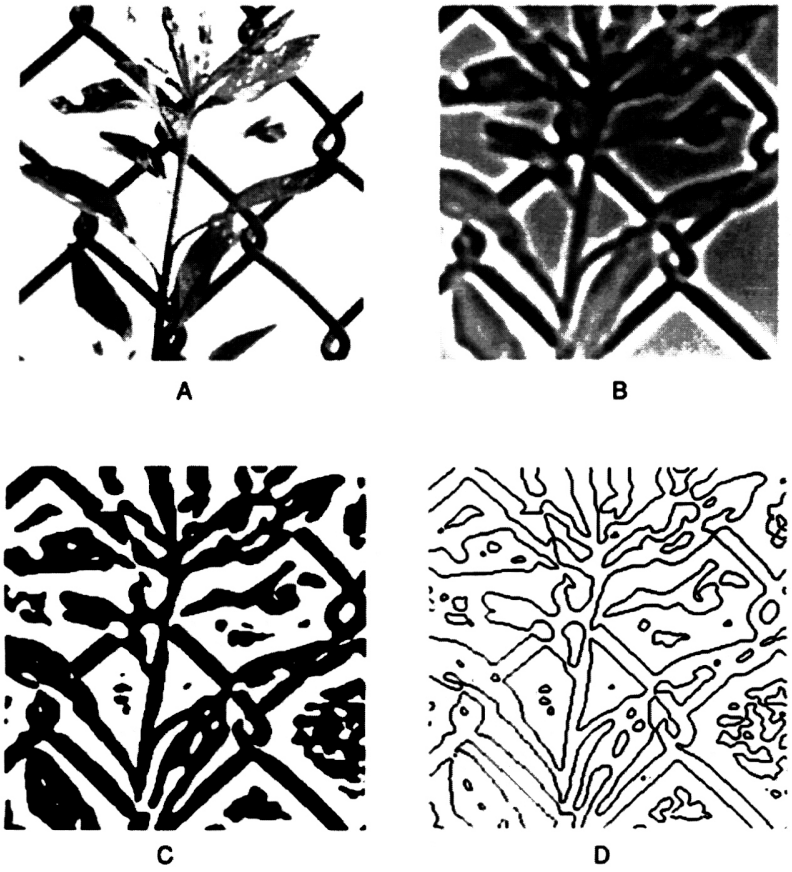



Figure 3.28: Finding edges by zero crossings in curvatue. (A) The image, (B) The image convolved with the edge operator with the values represented in gray scale. So negative values have darker grays and positive have lighter grays with zero being somewhere in the middle. (C) The convolved image with positive values in white and negative values in black. (D) Zero crossings detected from (C) to show the location of edges.

these cells can be modified by changing the context of their surroundings. Here is an example. Figure 3.33 shows that when a single bar is presented in the receptive field of a monkey, these neurons fire. But this firing reduces as this bar is surrounded by bars of many orientation. Basically, in this case, the bar does not 'stand out' and gets submerged in its context or surrounding. Finally, when this bar to surrounded by some similarly oriented bars to create a group that stands out, the firing of this neuron again increases. The degree by which an object stands out from its surrounding is called *salience*. Thus, these cells respond to salience.

Another experiment is shown in Figure 3.33. Here the money was presented a pattern of slanted line in its receptive field. Note that, initially there is about 80 ms of no activity in the neuron after which is starts firing at the rate of 300 ms. However, as the monkey keeps staring at the pattern, the neuron firing reduces. Now, notice how this response changes when we present the second pattern which is a slanted one surrounded by oppositely oriented pattern. Here also response is seen after 80 ms but this does not reduce with time. This shows that salience have some effect on selective adaptation. Also, the initial 80ms is probably required to process the information. Since salience is a much more difficult property to detect, the additional neural processing takes some extra time.

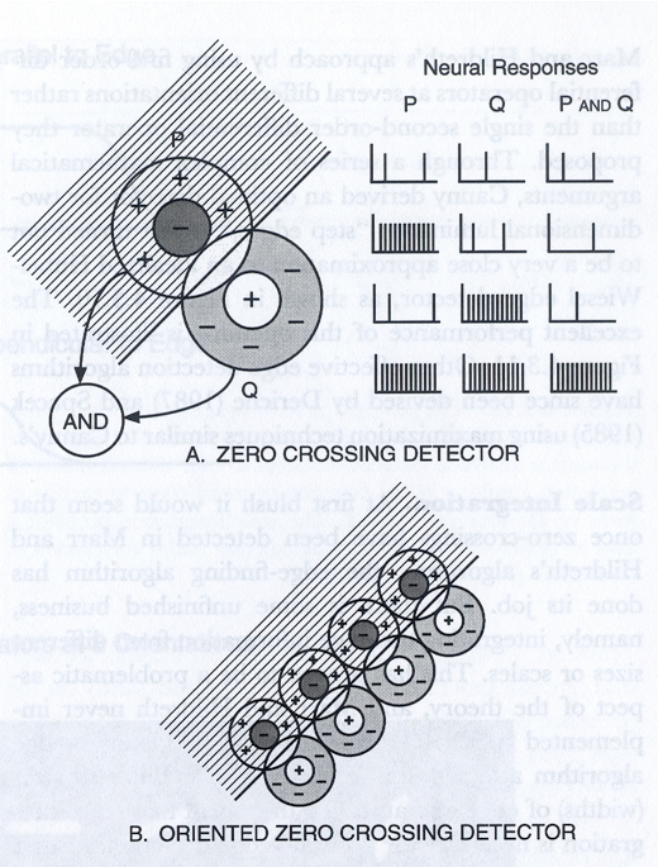



Figure 3.29: Model for edge detection in the visual pathways.

**Higher Level Processing in the Extra Striate Cortex**

In 1980s, it was found that actual there are visual processing neural streams going from straite cortex to extra straiate cortex where some of the imporant higher level processing takes place. It has been found that there are two streams of visual processing, one going to the ventral part of the brain to the temporal lobe and another to the dorsal part of the brain in the pareital lobe as shown in Figure 3.32. The ventral pathway is called the *what* pathway and deals with object discrimination. The dorsal pathway is called the *where* pathway and usually deals with object or landmark location. In fact, most scientists like to call this path the *how* path since this area is also involved in deciding how to perform actions.

An experiment was performed on the primates to discover this. Three sets of monkeys were taken. On one set a surgery was performed to remove its parietal lobe. The ventral lobe was removed for the second set and the third set was kept untouched. Then they were give two tasks. It was given a known food along with an unknown object. If the monkey identified the known food from these two, he was rewarded. The second task was to find the known food from under a cylindrical cover. It was found that the monkeys with their ventral lobe removed struggled to distinguish between the known food and the unknown object. On the other hand, monkeys whose parietal lobe was removed found it difficult to find the covered food. These gave rise to the concepts of what and where pathway.

Research in dorsal and ventral pathway revealed that they are parallel pathways serving different functions. However, they are not entirely independent of each other. There are some cross talk across the two paths. Figure 3.34 shows a sketch of how the neural pathways are connected from the ganglion cells through the LGN to the visual cortex and the extra-striate cortex. $V1$ is an area in the striate cortex. $V2$, $V3$ and $V4$ are various areas in

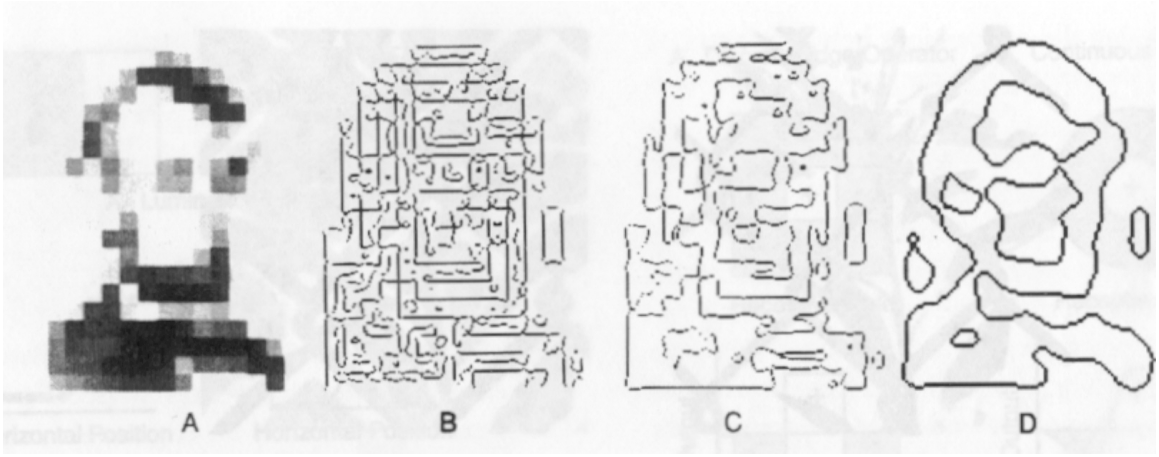



Figure 3.30: This image shows the edges detected when the algorithm is run at different scales. (A) for fine edges, (B) for coarser edges (C) for coarsest edges.

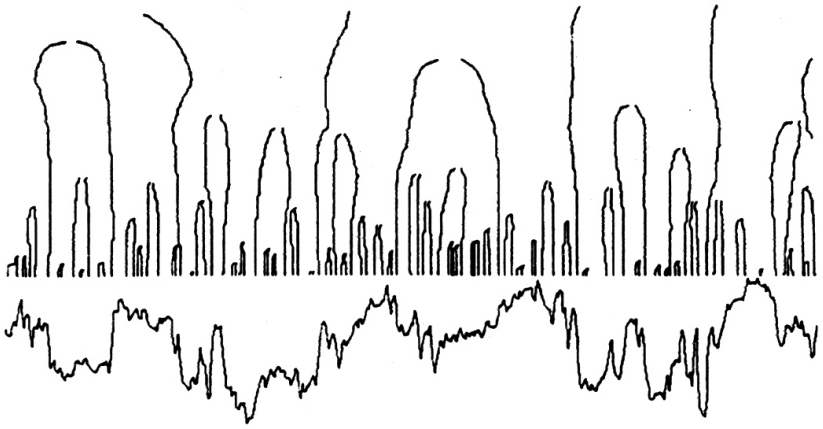


Figure 3.31: Scale integration model.

the extra straite cortex which are relatively independent of each other. Note that the connections from the P-cells in the ganglion which are responsible for details, color and texture percpetion end up in the ventral pathway. Whereas the connections from the M-cells in the ganglions which are responsible for motion end up in the parietal pathway. The IT and MT are the infotemporal and middle temporal part of the brain, both in the extra striate cortex.

The next big break through was the discovery of *modular* neurons in the extra straiate cortex. Research has shown that certain areas in the extra striate cortex process information of very specific shapes or objects. This specification is called modularity, and the structure that contains neurons responding to these kind of specifications is called modular neurons for that particular specification. These type of neurons are mostly found in the infotemporal and middle temporal region of the brain.

An experiment was performed with two sets of monkeys, one with parts of their MT removed by surgery and others normal. The monkeys were shown a bunch of moving dots. The movements of these dots can be completely un-correlated (random) or can be partially or completely correlated to move in a particular direction. If $x\%$ of the dots move in one direction and the rest have a rando, direction, then it is said to be $x\%$ correlation in direction of motion. Figure 3.35 shows dots with various degrees of correlation. The monkeys were trained to tell the

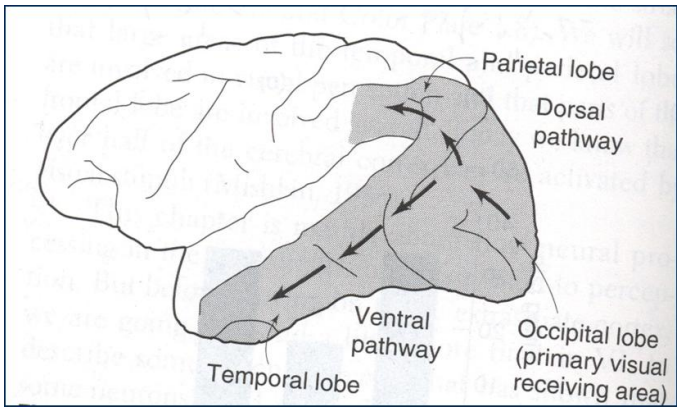



Figure 3.32: The *what* and *where* pathways in the ventral and parietal lobes.

*predominant* direction of the movements of the dots from these kinds of moving dots. The normal monkeys were able to detect the predominant direction even with $1-2\%$ of correlation. However, the monkeys with parts of their MT removed could not detect the predominant direction for less than $10-20\%$ of correlation. This and other studies confirmed the importance of MT in motion perception.

In the infotemporal area, two types of cells were found.

- *Primary Cells:* This cells respond vigorously to fairly simple shapes like slits, spots, ellipses and squares.
- *Elaborate Cells:* These cells respond to complex stimuli of specific shapes combined with specific color and texture. For example, an elaborate cell may respond to a shape like the table tennis racket (a disc with a thin bar), but would not respond to the bar or the disc alone, or even to a disc with a shorter bar. Similarly, an elaborate cell that responds to the star will not respong to a sphere.

Also it was found that these cells also had a columnar organization as the cells in the visual cortex. So, a hypercolum structure, as shown in Figure 3.36, was suggested.

In fact, the most important discovery in this direction were neurons that responded to faces. Neurons were discovered in a particular area of IT which would only fire when viewing a face. This neurons respond when viewing a whole human being (along with the face) but stops responding when the face is covered with a paper. Hence, this area was accordingly named as *fusiform face area (FFA)* or the *fusiform gyrus*. This was further supported by neurophysiological study of people with brain damage. These show that a damage caused to the temporal lobe can lead to a condition called *prosopagnosia*, in which people fail to regognize familiar faces. They may not be even able to recognize people very close to them like family, friends and relatives.

### Sensory Coding

At this point, what generates our interest is then how exactly do we identify different objects? Sensory codes define how different sets of neurons respond to the same stimulus. So, it helps us to find a pattern across different neurons. One question that comes to mind after the discovery of FFA, i.e. is the sensory code a *specificity code* or a *distributed code*. When every neuron is specialized to respond to a very specific stimulus, then it is called specificity coding. But when different levels of response across different neurons help us to identify an object, it is called distributed coding. This is kind of like the number system. The more number of levels we have for each digit, the lesser number of digits we need for representing a number. For example, in the unary system where each digit can take only one number, the number $10$ will be represented by $10$ tally marks, i.e. $10$ digits. But, in the binary system, where each digit can now take $2$ values, less number of digits ($4$ digits) can be used to represent

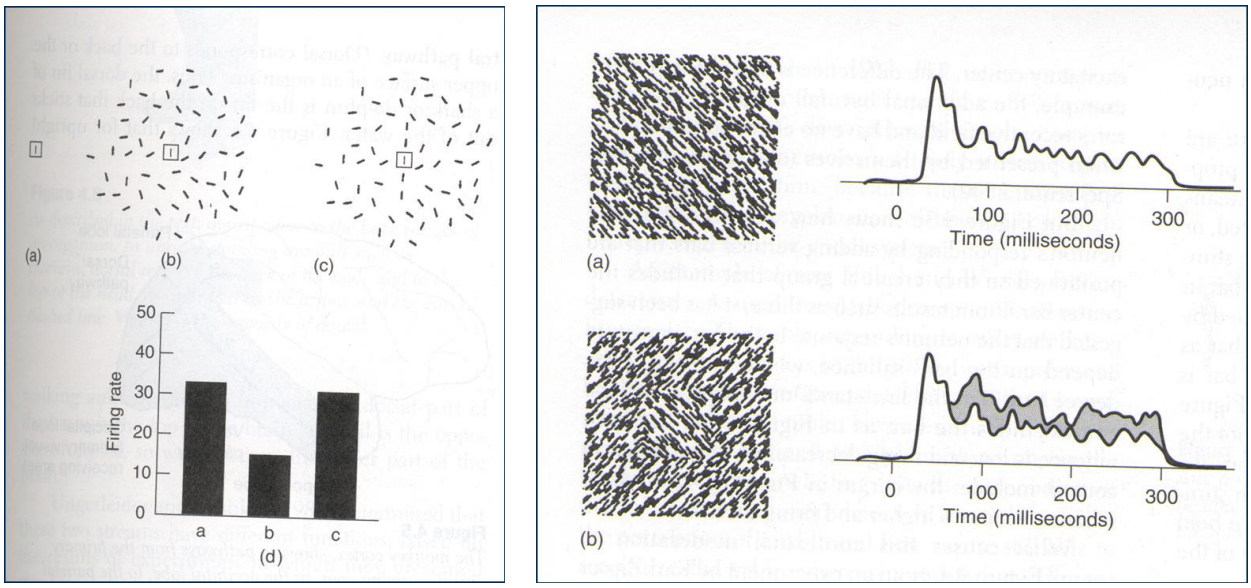



Figure 3.33: Left: How context affects a cell in the visual cortex $V1$: (a) a single vertical bar by itself in the receptive field, (b) the same bar surrounded by many bars of different orientation, (c) the same bar with other bars lined up with it to create a group, (d) response of a neuron to conditions (a),(b) and (c). Right: Effect of a context on a monkey's $V1$ neuron.

10. The number of digits becomes still smaller when we come to decimal system (2 digits) and reduces to 1 digit in the hexadecimal system. The unary system is the case of specificity coding while the others are the example of distributed coding.

Note that if brain followed specificity coding, it would mean that we would need one different neuron to perceive every different object. Needless to say, given the number of objects that we see and identify, this seems highly impossible. So, it is true that distributed coding is used and the difference in relative amount of response in different neurons caused by different stimulus, helps us to distinguish between them. If a large number of neurons are involved, then the levels of different response that it can achieve is low (analogous to more digit and hence less number of levels per digit). If there are small number of neurons involved, they respond in a large number of levels. For example, the motion sensitive cells in the MT are a few but can detect various kinds of motions.

Another thing that we need to ask in this context is how do we respond to the same object in a similar fashion even if it is at different distances, is oriented differently and has different size. In fact, researchers found that there are three special types of neurons in the IT: (1) *size invariant* neurons that continue to respond to an object similarly even if it is at different distances, (2) *location invariant* neurons that respond to objects similarly irrespective of where they are in the visual field and (3) *view invariant* neurons that respond to objects similarly even if the view is changed.

## 3.5 The Role of Attention

So far, we have assumed that a stimulus is presented to the eye and several neurons are responding to it to let us percieve the stimulus. But this is not the whole story. We seldom sit passively as stimuli is presented to us. In fact, we take an active part in the perception by seeking out what is important or interesting to us in the stimuli. This process of seeking out stimuli and then focussing on it is called *attention* and the stimulus thus attended to is called the *attended stimulus*. In this section, we will see how attention affects our stimulus.

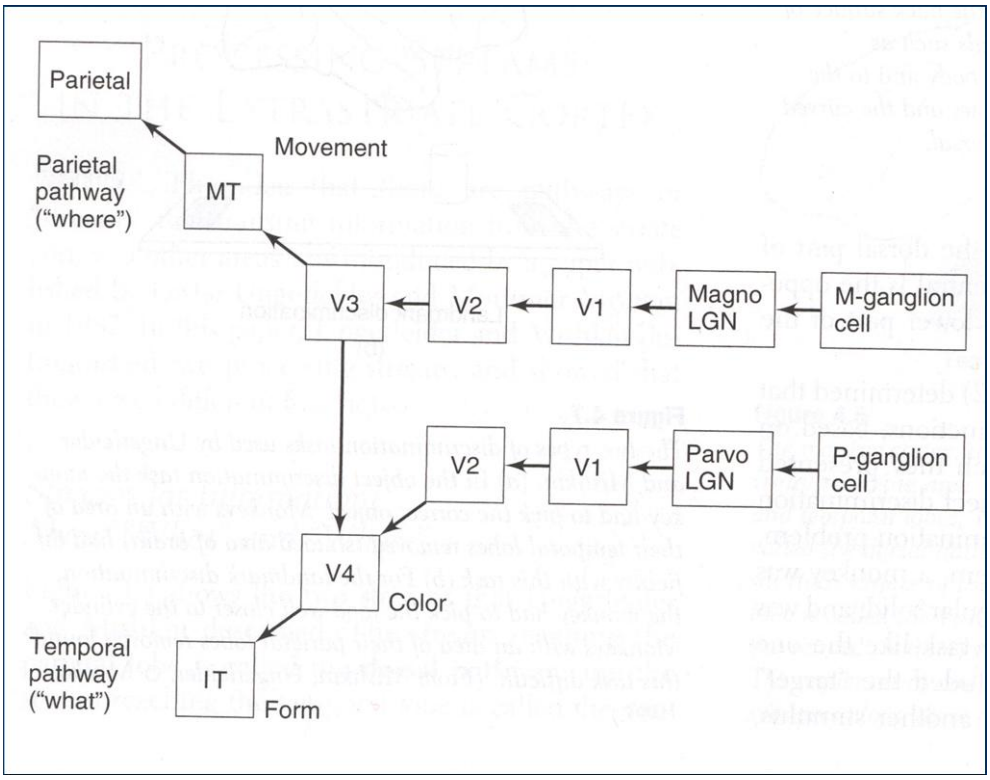



Figure 3.34: The connections from the retinal ganglion cells to the extra striate cortex.

### 3.5.1 Inattentional Blindness

Can we see without attention? Inattentional blindness is exactly the situation where the stimulus that is not attended to is not percieved. The following experiment shows this convincingly. A plus sign made of lines of two different lengths were presented to subjects. The subjects were asked to find out which line is of greater length. While they were doing this job a small triangle was also presented right in their receptive field. Once the job was accomplished, the subjects were given a bunch of shapes (like circle, square, ellipse etc) and asked to identify which shape was presented to them when they were accomplishing the job of finding the shorter line. Almost all of the subjects were unable to pick the right shape since they were not attending to it. This is called inattentional blindness.

### 3.5.2 Attentional Blink

The attentional blink is the inability to see a second stimulus shortly after the first stimulus is presented. This is demonstrated by a procedure called the *rapid serial visual perception*. Subjects are presented with a series of pictures that are presented one after another, each picture for about $100$ ms ($\frac{1}{10}$ of a second). A series of letters are presented and two numbers are embedded in these series. These numbers are in succession. The subjects are asked to identify the two numbers. The subjects were able to identify the first number but not the second one. However, when the numbers were presented for $500$ ms each, they were able to identify the number.

The reason for this is when the subject is focussing attention on the first target number, it also uses up the subject's attention for $500$ ms afterwards. If anything is presented within this time, the subject cannot attend to that stimuli. This is called *attentional blink*. Just as we do not see anything when we blink our eye, similarly we cannot attend to a stimuli during the period of attentional blink.

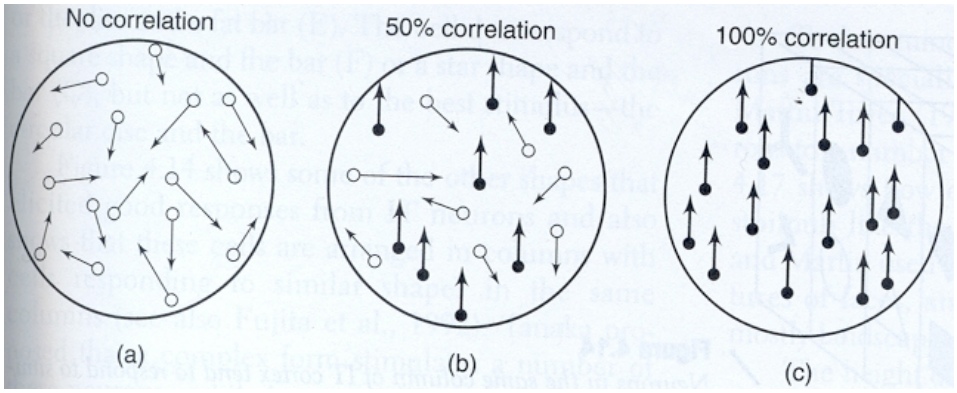



Figure 3.35: The moving dots pattern shown to the monkeys to find the function of MT neurons.

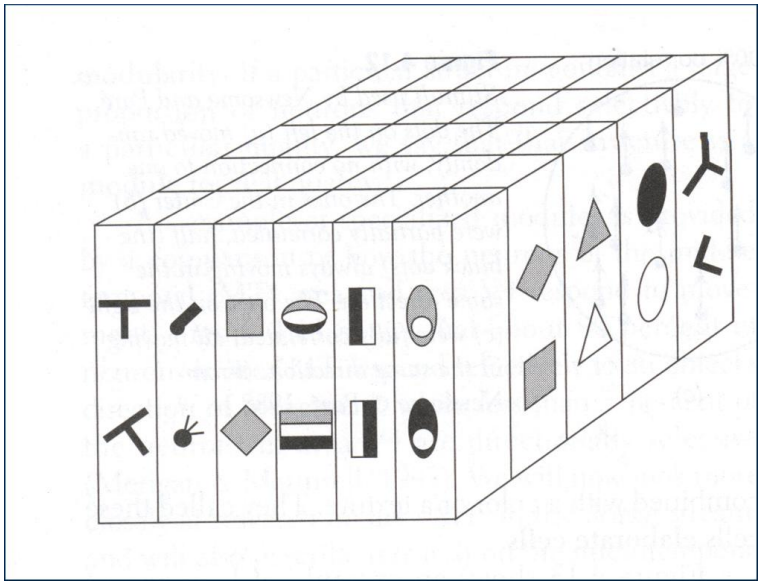


Figure 3.36: The hypercolumn structure of arrangement for the primary and the elaborate cells.

### 3.5.3 Change Blindness

For attentional blink or inattentional blindness, the stimuli is presented very fast. There are evidences that people are not very good at seeing unattended objects even when the stimuli is presented very slowly so that everything in the scene can be perceived easily. This was demonstrated by making the subjects view a short video sequence of a conversation between two women. No flashy action scenes but just simple conversation between two women. Figure 3.37 shows four scenes from this video. These scenes show a few of the changes. In shot B, the scarf of the woman dissappears. It again reappears in shot C. In shot C, the color of the plates have changed to red. They have again changed to white in scene D. All these changes went unnoticed by the audience when they were not told that the scene will have some changes that they have to detect. Even when they were told that there will be changes in the scene, they were able to detect only $25\%$ of the changes. Though this may seem counterintuitive, this experiment has been repeated many times. Basically, we usually have a general awareness of our surrounding, but most of the time we are much less aware of most of the details than what we would like to think.

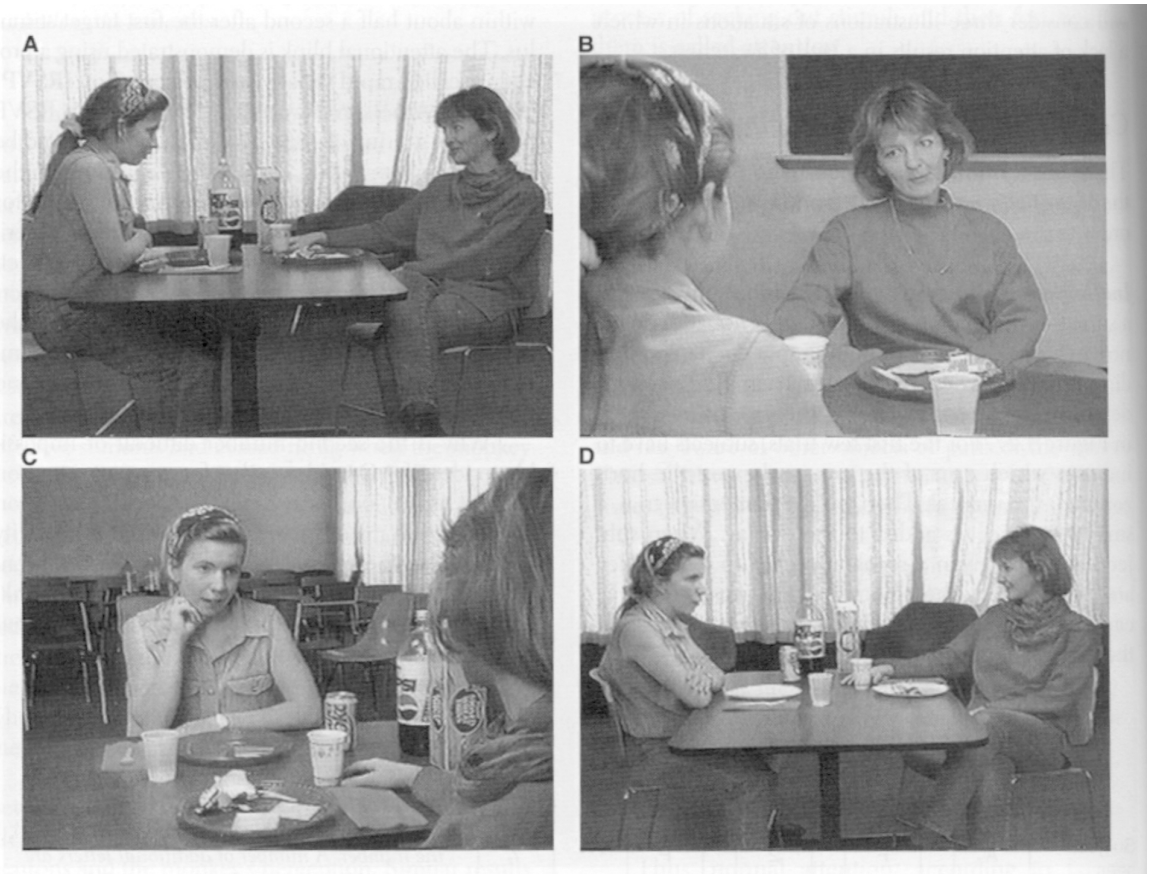



Figure 3.37: Some of the shots of the video for change blindness experiment.

## 3.6 The Binding Problem

Till now we have been saying that when we percieve something, several changes happen in different areas of the brain. For example, if a red sports car passes by us, what happens in our visual pathways. Cell sensitive to complex form fire in IT; cells sensitive to movement fire in MT; and other cells sensitive to color fire in different other areas of the visual pathway. But, even though all these firing happens in different areas of the brain, we do not perceive the car as separated movement, color and shape. All these are integrated as a perception of the car. The raises the important question: how are these qualities integrated? This is called the *binding problem*.

Of course, every part of the nervous system is connected with every other part. So, the neurons from different parts can share information. This is true. But what kind of information is shared is the second question to be asked. A possible hypothesis has been proposed by a group of German researchers in early 1990s. They suggested that when the different attributes of the *same* object is represented in different groups of neurons in different parts of the visual pathway, the firing of all these neurons happen in a synchronized fashion.

As an evidence, the following experiment was performed. The firing of two different neurons in the visual cortex (7mm apart) were measured. Both of them were sensitive to vertical orientation. Figure 3.38 shows the plots of how these two neurons responded to different kinds of stimuli. These plots are called *cross correlograms*. First, a single long bar was swept across both's receptive field. The neurons were seen to fire synchronized with each other as shown by the plot. Next, two different bars were swept across their receptive fields, but in the same direction and as is evident from the plot, the synchronization reduced. Next, two different bars were swept across their receptive fields in different directions and as you notice there is barely any synchronization.

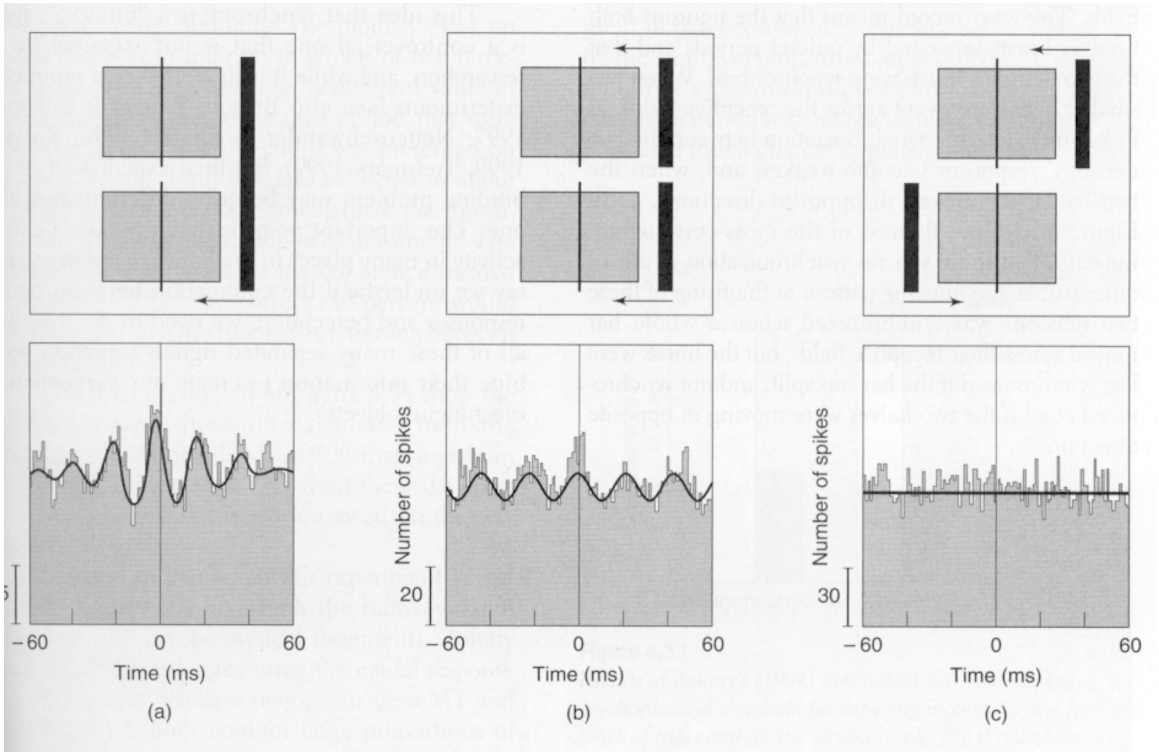



Figure 3.38: Correlograms showing the synchronization amongst neurons to explain the binding problem.