Chapter 2

Receptive Fields and Visual Responses

2.1 Introduction

The methods discussed in chapter 1, particularly spike-triggered averages and reverse correlation techniques, are used extensively to study properties of neurons at early stages of the visual system, in particular retina (retinal ganglion cells), lateral geniculate nucleus (LGN), and primary visual cortex (V1, or area 17 in the cat). The response properties of neurons at successive stages of the visual pathway get more complex and nonlinear. At early stages, however, some neurons (X cells in the cat or P cells in the monkey retina and LGN, and simple cells in primary visual cortex) can be described using linear filters with relatively simple static nonlinearities such as thresholds. Other neurons (complex cells in primary viusal cortex) can be modeled with slight extensions of this formalism. These models are approximations, not exact descriptions of neural responses, and they are purely descriptive not mechanistic. Nevertheless, they play an important role in our understanding of visual processing for a number of reasons. The spike-triggered averages upon which these models are based provide a framework for characterizing the selectivity of visually responsive neurons. In addition, models based on linear filters and static nonlinearities provide a reference point for identifying and characterizing novel effects in experimental data. Finally, by providing a first-order description of neural responses, these models form a basis for developing mechanistic models that attempt to explain how visual responses arise from the synaptic, cellular and network properties of retinal, LGN, and cortical circuits. Models based on linear filters and static nonlinearities have also been used to describe responses in V2, area 18 and area MT, but they fail to capture the response characteristics at later stages of the visual system.

The conversion of a light stimulus into an electrical signal and ultimately an action potential sequence occurs in the retina. Figure 2.1A is an anatomical diagram showing the five principal cell types of the retina and figure 2.1B is a rough circuit diagram. In the retina, light is first converted into an electrical signal by a phototransduction cascade within rod and

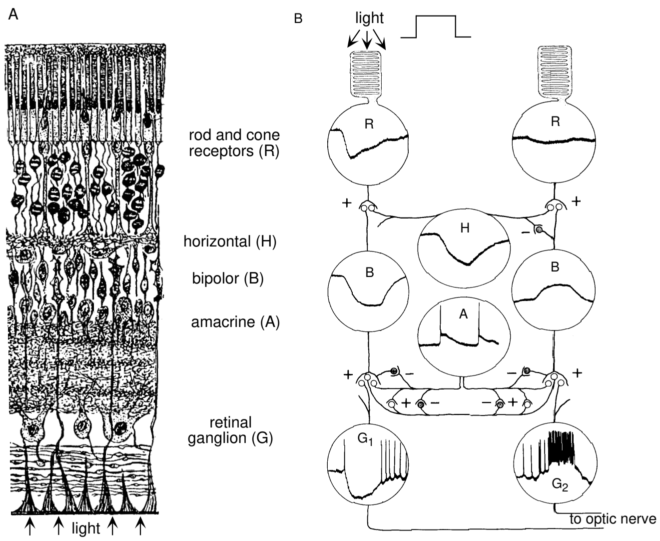

Figure 2.1: A) An anatomical diagram of the circuitry of the retina of a dog. Cell types are identified at right. B) Intracellular recordings from retinal neurons of the mudpuppy responding to a flash of light. In this experiment illumination was provided from the top of the figure, but in the intact eye, illumination is, counter-intuitively, from the bottom of the figure. Pluses and minuses represent excitatory and inhibitory synapses respectively. (A figure adapted from Nicholls *et al*, 1992, drawing from Cajal, 1911; B adapted from Dowling, 1992)

cone photoreceptor cells. Figure 2.1B shows intracellular recordings made in neurons of the retina of a mudpuppy (an amphibian) by Werblin and Dowling (1969). The stimulus used for these recordings was a flash of light falling primarily in the region of the left-hand photoreceptor in figure 2.1B. The rod cells, especially the directly illuminated cell on the left side of figure 2.1B, are hyperpolarized by the light flash. This electrical signal is passed along to bipolar and horizontal cells. Note that in one of the bipolar cells, the signal has been inverted leading to depolarization. These smoothly changing membrane potentials provide a graded representation of the light intensity during the flash. This form of coding is adequate for signaling within the retina where distances are small. However, it is inadequate for the task of conveying information from the retina to the brain.

The retinal output is carried along the optic nerve formed by the axons of retinal ganglion cells, the output cells of the retina. As seen in figure 2.1B, the subthreshold potentials of the two retinal ganglion cells shown are similar to those of the bipolar cells immediately above them in the figure, but now with superimposed action potentials. The two retinal ganglion cells shown in the figure have different responses and transmit different sequences of action potentials to the brain. G_2 fires while the light is on, G_1 fires when it turns off. These are called ON and OFF responses, respectively. The optic nerve conducts the output spike trains of retinal ganglion cells to the lateral geniculate nucleus of the thalamus, which acts as a relay station between the retina and primary visual cortex (figure 2.2). Prior to arriving at the thalamus, some retinal ganglion cell axons cross the midline a the optical chiasm. This allow the left and right sides of the visual fields from both eyes to be represented on the left and right sides of the brain respectively (figure 2.2).

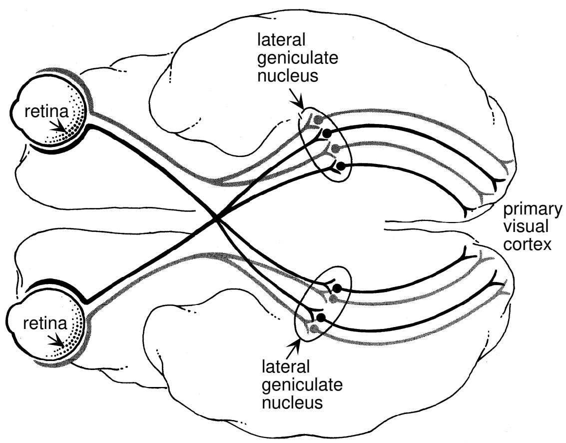

Figure 2.2: Pathway from the retina through the lateral geniculate nucleus (LGN) of the thalamus to the primary visual cortex in the human brain. (Adapted from Nicholls, Martin, R & Wallace, 1992)

Neurons in the retina, LGN and primary visual cortex respond to light stimuli in restricted regions of the visual field called their receptive fields. Patterns of illumination outside the receptive field of a given neuron cannot generate a response directly, but can significantly affect responses to stimuli within the receptive field. We will not consider such effects in this chapter, but will return to them in chapter 7. Within the receptive fields, there are regions where illuminations higher than the background light intensity enhance firing, and other regions where lower than average light intensities enhance firing. The spatial arrangement of these regions determines the selectivity to different patterns of illumination. The term receptive field is often generalized to refer not only to the overall region where light affects neuronal firing, but also to the spatial and temporal structure within this region. To streamline the discussion in this chapter, we consider only greyscale images, although the methods presented can be extended to include color. Much of the data we show comes from cats which have very little color vision. We also restrict the discussion to two-dimensional visual images, ignoring how visual responses depend on viewing distance and encode depth (although see exercise 8).

A first step in studying the selectivity of any neuron is to identify the types of stimuli that make it respond maximally. Retinal ganglion cells and LGN neurons have similar selectivities and respond best to circular spots of light surrounded by darkness or dark spots surrounded by light. Neurons in primary visual cortex respond best to elongated light or dark bars with a certain width and to boundaries between areas of light and dark. Gratings with alternating light and dark regions are effective and frequently used stimuli for these neurons. Visually responsive neurons also have interesting temporal selectivities. In primary visual cortex, neurons respond best to transitions in time between light and dark, a temporal analog of their optimal responses to light-dark boundaries. Static images, even those with spatial boundaries, are not very effective at evoking responses. Moving light bars and gratings, or gratings undergoing periodic light-dark reversals (counterphase gratings) are more effective stimuli. In addition, many neurons in primary visual cortex are directionally selective; they respond more strongly to stimuli moving in one direction than in the opposite direction.

The Retinotopic Map

A striking feature of most visual areas in the brain, including primary visual cortex, is that the visual world is mapped onto the cortical surface in a topographic manner. This means that neighboring points in a visual image evoke activity in neighboring regions of visual cortex. Objects located a fixed distance from the eye lie on a sphere. Locations on this sphere can be represented by the same longitude and latitude angles used for the surface of the earth Typically, 'north pole' for this spherical coordinate system is located at the fixation point, the image point that focuses onto the fovea or center of the retina. In this system of coordinates, the latitude coordinate is called the eccentricity and longitude coordinate the azimuth.

In most experiments, images are displayed on a flat screen (called a tangent screen) that does not coincide exactly with this sphere, but in most cases the difference is negligable. On the screen, image points can be represented either using polar coordinates (the analog of longitude and latitude coordinates) or ordinary Cartesian coordinates (figure 2.3A). Both systems are used. The eccentricity ϵ and the x and y coordinates of the Cartesian system are based on measuring distances on the screen. However, it is customary to divide these measured distances by the viewing distance from the eye to the screen and to multiply the result by $180^\circ/\pi$ so that these coordinates are ultimately expressed in units of degrees. This makes sense because it is the angular not the absolute size and location of an image that is typically relevant for studies of the visual system.

Figure 2.4A shows a dramatic illustration of the retinotopic map in primary visual cortex. This is part of the left-side primary visual cortex of a macaque monkey that was exposed to a visual image consisting of con-

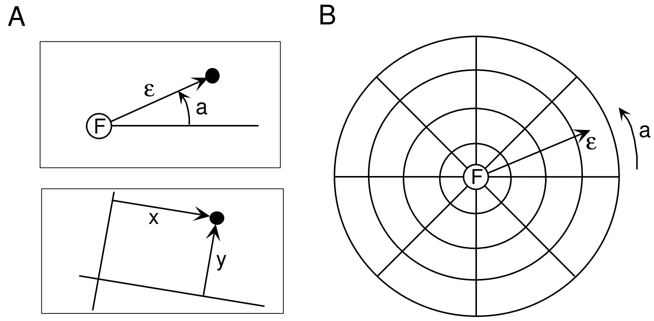

Figure 2.3: A) Two coordinate systems used to parameterize image location. Each rectangle represents a tangent screen, and the filled circle is the location of a particular image point on the screen. The upper panel shows polar coordinates. The origin of the coordinate system is the fixation point F, the eccentricity ϵ is proportional to the radial distance from the fixation point to the image point, and the coordinate a is the angle between the radial line from F to the image point and the horizontal axis. The lower panel shows Cartesian coordinates. The location of the origin for these coordinates and the orientation of the axes are arbitrary. They are usual chosen to center and align the coordinate system with respect to a particular receptive field being studied. B) A bullseye pattern of radial lines of constant azimuth, and circles of constant eccentricity. The center of this pattern at zero eccentricity is the fixation point F. Such a pattern was used to generated the image in figure 2.4A.

centric circles and radial lines, like the pattern in figure 2.3A. In primary visual cortex, the visual world is split in half, with the the region $-90^\circ \leq a \leq +90^\circ$ for ϵ from zero to about 70° (for both eyes) represented on the left side of the brain, as shown in figure 2.3A, and the reflection of this region about the vertical meridian represented on the right side of the brain. The pattern on the cortex seen in figure 2.4A, made by imaging a radioactive analog of glucose, indicates regions that were activated by this visual image. The vertical lines correspond to the circles in the image, and the roughly horizontal lines are due to the activity evoked by the radial lines in the image. The fovea is represented at the left-most pole of this piece of cortex and eccentricity increases toward the right. Azimuthal angles are positive in the lower half of the cortex shown, and negative in the upper half. Figure 2.4B is an approximate mathematical description of the map illustrated in figure 2.4A. To construct this map we assume that eccentricity is mapped onto the horizontal coordinate X of the cortical sheet, and a is mapped onto its Y coordinate. The equations for X and Y as functions of ϵ and a can be obtained through knowledge of a quantity called the cortical magnification factor, $M(\epsilon)$. This determines the distance across a flattened sheet of cortex separating the activity evoked by two nearby image points. Suppose that the two image points in question have eccentricities ϵ and $\epsilon + \Delta\epsilon$ but the same value of the azimuthal coordinate a. The angular distance between these two points is $\Delta\epsilon$. The distance separating the activity evoked by these two image points on the cortex is ΔX.

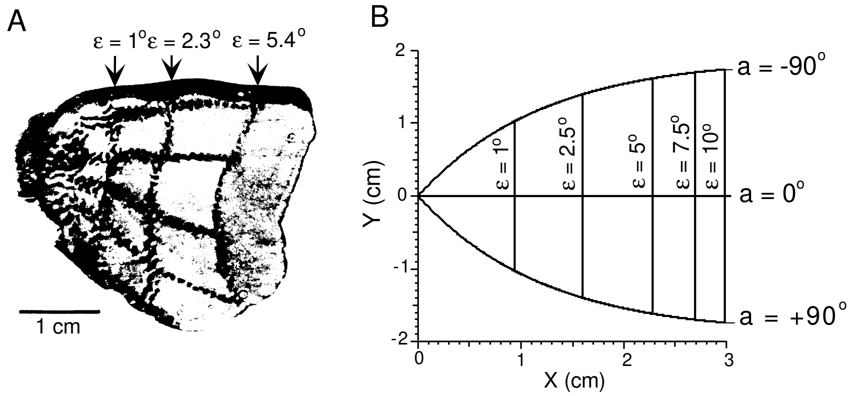

Figure 2.4: A) An autoradiograph of the posterior region of the primary visual cortex from the left side of a macaque monkey brain. The pattern is a radioactive trace of the activity evoked by an image like that in figure 2.3A. The vertical lines correspond to circles at eccentricities of 1°, 2.3°, and 5.4°, and the horizontal lines (from top to bottom) represent radial lines in the visual image at a values of $-90°$, $-45°$, $0°$, $45°$, and $90°$. Only the part of cortex corresponding to the central region of the visual field is shown. B) The mathematical map from the visual coordinates ϵ and a to the cortical coordinates X and Y described by equations 2.3 and 2.5. (A from Tootell *et al*, 1982)

By the definition of the cortical magnification factor, these two quantities must satisfy $\Delta X = M(\epsilon)\Delta\epsilon$ or, taking the limit as ΔX and $\Delta\epsilon$ go to zero,

$$\frac{dX}{d\epsilon} = M(\epsilon)\,. \tag{2.1}$$

The cortical magnification factor for the macaque monkey, obtained from results such as figure 2.4A is approximately (Van Essen *et al*, 1984)

$$M(\epsilon) = \frac{\lambda}{\epsilon_0 + \epsilon}\,. \tag{2.2}$$

with $\lambda \approx 12$ mm and $\epsilon_0 \approx 1°$. Integrating equation 2.1 and defining $X = 0$ to be the point representing $\epsilon = 0$, we find

$$X = \lambda \ln(1 + \epsilon/\epsilon_0)\,. \tag{2.3}$$

If the same cortical amplification factor is applied to points with the same eccentricity but different a values on a flat tangent screen, we find that

$$\frac{dY}{da} = -\frac{\epsilon M(\epsilon)\pi}{180°}\,. \tag{2.4}$$

The minus sign accounts for the fact that the visual field is inverted on the cortex, the factor of ϵ corrects for the increase of arc length as a function of eccentricity, and the factor of $\pi/180°$ converts from degrees to radians. Solving equation 2.4 gives

$$Y = -\frac{\lambda\epsilon a\pi}{(\epsilon_0 + \epsilon)180°}\,. \tag{2.5}$$

Figure 2.4B shows that these coordinates agree fairly well with the map in figure 2.4A.

For eccentricities appreciably bigger than 1°, $X \approx \lambda \ln(\epsilon/\epsilon_0)$ while $Y \approx -\lambda\pi a/180^\circ$. These two formulae can be combined by defining the complex numbers $Z = X + iY$ and $z = (\epsilon/\epsilon_0)\exp(-i\pi a/180^\circ)$ (with i equal to the square root of -1) and writing $Z = \lambda \ln(z)$. For this reason, the cortical map is sometimes called a complex logarithmic map (see Schwartz, 1977). The logarithmic transformation causes images that are scaled radially outward on the retina to be represented at locations on the cortex translated in the X direction. For an image scaled radially by a factor γ, eccentricities change according to $\epsilon \to \gamma\epsilon$ while a is unaffected. Scaling of the eccentricity produces a shift $X \to X + \lambda \ln(\gamma)$ over the range of values where the simple logaritmic form of the map is valid. For smaller eccentricities, the map we have derived is only approximate even in the complete form given by equations 2.3 and 2.5. This is because the cortical magnification is not really isotropic as we have assumed in this derivation, and a complete description requires accounting for the curvature of the cortical surface (see Rovamo and Virsu, 1984). An interesting application of the cortical map to account for the spatial structure of visual hallucinations, due to Ermentrout and Cowan (1979), is given in exercise 1.

Visual Stimuli

In chapter 1, we characterized a time-dependent stimulus by a single function $S(t)$. For the visual system, the characterization of the stimulus is more complex. We consider images appearing on a two-dimensional surface, such as a video monitor, with pixel locations parameterized by Cartesian coordinates x and y as in the lower panel of figure 2.3A. The light intensity or luminance at the point (x, y) and at time t is the stimulus $S(x, y, t)$. This can be written as $S(x, y, t) = \overline{S} + s(x, y, t)$ where $s(x, y, t)$ represents the luminance relative to the average or background level of illumination on the screen, $\overline{S}$. By considering how responses depend on $s(x, y, t)$, we eliminate complications associated with adaptation to the overall level of illumination.

During recordings, visual neurons are usually stimulated by images that vary in both space and time. In the following discussion, we make use of a stimulus called a counterphase sinusoidal grating described by

$$s(x, y, t) = A\cos\big(Kx\cos(\Theta) + Ky\sin(\Theta) - \Phi\big)\cos(\omega t) \qquad (2.6)$$

and shown in figure 2.5. For this stimulus, ω is the temporal frequency, K the spatial frequency (these are angular frequencies), Θ the orientation, Φ the spatial phase, and A the contrast of the grating. This stimulus oscillates in time with a period $2\pi/\omega$ and, at any fixed time, it also oscillates as a function of position with wavelength $2\pi/K$. At any fixed time, the stimulus does not vary in the direction given by the orientation angle Θ.

For convenience, Θ is measured relative to the y axis rather than the x axis (figure 2.5). The spatial phase Φ determines the location of the grating, and the contrast A controls the maximum degree of difference between light and dark areas. Figure 2.5 shows a cartoon of a grating (a spatial square-wave is drawn rather than a sinusoid) and illustrates the significance of the parameters K, Θ, Φ, and ω. Because x and y are measured in degrees, K has the rather unusual units of radians per degree and $K/2\pi$ is typically reported in units of cycles per degree. Φ has units of radians, ω is in radians per s, and $\omega/2\pi$ is typically reported in Hz.

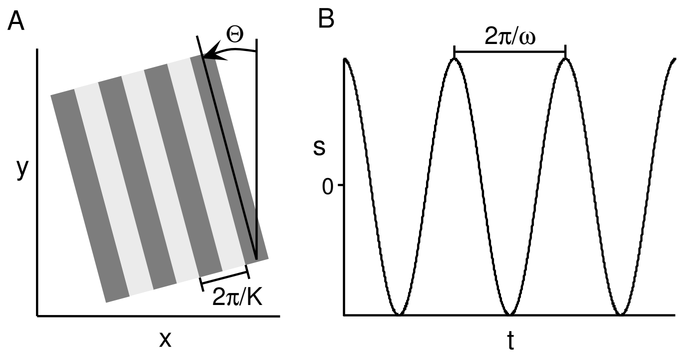

Figure 2.5: A counterphase grating. A) A portion of a square-wave grating analogous to the sinusoidal grating of equation 2.6. The lighter stripes are regions where $s > 0$, and $s < 0$ within the darker stripes. K determines the wavelength of the grating and Θ its angular orientation. Changing the spatial phase Φ shifts the entire light-dark pattern in a direction perpendicular to the stripes. B) The light-dark intensity at any point of the spatial grating oscillates as a cosine function of time with period $2\pi/\omega$.

Many factors limit the maximal spatial frequency that can be resolved by the visual system, but one interesting effect arises from the size and spacing of individual photoreceptors on the retina. The region of the retina with the highest resolution is the fovea at the center of the visual field. Within the macaque or human fovea, cone photoreceptors are densely packed in a regular array. Along any direction in the visual field, a regular array of tightly packed photoreceptors of size Δx samples points at locations $m\Delta x$ for $m = 1, 2, \ldots$. The (angular) Nyquist frequency that defines the resolution of such an array is given by

$$K_{\text{nyq}} = \frac{\pi}{\Delta x}\,. \tag{2.7}$$

There are 120 cones per degree at the fovea of the macaque retina which makes $K_{\text{nyq}}/(2\pi) = 1/(2\Delta x) = 60$ cycles per degree. To understand the significance of the Nyquist frequency, consider sampling two cosine gratings with spatial frequencies of K and $2K_{\text{nyq}} - K$, with $K > K_{\text{nyq}}$. These are described by $s = \cos(Kx)$ and $s = \cos((2K_{\text{nyq}} - K)x)$. At the sampled points, these functions are identical because $\cos((2K_{\text{nyq}} - K)m\Delta x) =$

$\cos(2\pi m - Km\Delta x) = \cos(-Km\Delta x) = \cos(Km\Delta x)$ by the periodicity and evenness of the cosine function (see figure 2.6). As a result, oscillations at these two spatial frequencies cannot be distinguished. Any spatial frequency $K > K_{\rm nyq}$ will be interpreted as a frequency $2K_{\rm nyq} - K$, a phenomenon known as aliasing.

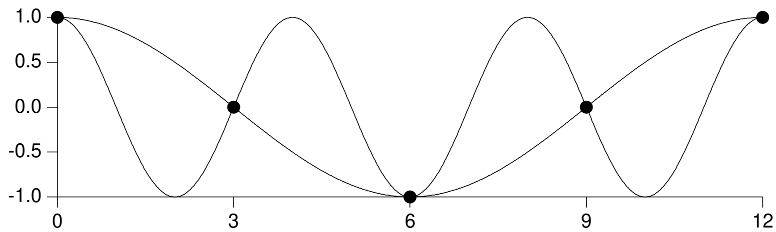

Figure 2.6: Alisasing and the Nyquist frequency. The two curves are the functions $\cos(\pi x/6)$ and $\cos(\pi x/2)$ plotted against x, and the dots show points sampled with a spacing of $\Delta x = 3$. The Nyquist frequency in this case is $\pi/3$, and the two cosine curves match at the sampled points because their spatial frequencies satisfy the Nyquist condition $2\pi/3 - \pi/6 = \pi/2$.

Experiments that consider reverse correlation and spike-triggered averages use various types of random and white-noise stimuli in addition to bars and gratings. A white-noise stimulus, in this case, is one that is uncorrelated in both space and time so that

$$\frac{1}{T}\int_0^T dt\, s(x,y,t)s(x',y',t+\tau) = \sigma_s^2\delta(\tau)\delta(x-x')\delta(y-y')\,. \tag{2.8}$$

Of course, in practice a discrete approximation of such a stimulus must be used by dividing the image into pixels and time into small bins. In addition, more structured random sets of images (randomly oriented bars, for example) are sometime used to enhance the average response obtained during stimulation.

Reverse Correlation Methods

The spike-triggered average for visual stimuli is defined, as in chapter 1, as the average over trials of stimuli at times $t_i - \tau$, where t_i for $i = 1, 2, \ldots, n$ are the times when spikes were fired during a trial of duration T. Because the light intensity of a visual image depends on location as well as time, the spike-triggered average stimulus is a function of three variables,

$$C(x,y,\tau) = \frac{1}{\langle n\rangle}\left\langle \sum_{i=1}^{n} s(x,y,t_i-\tau)\right\rangle \tag{2.9}$$

where, as in chapter 1, the brackets denote trial averaging. $C(x,y,\tau)$ is thus the average value of the visual stimulus at the point (x,y) a time τ

before a spike was fired. Similarly, we can define the correlation function between the trial-averaged firing rate $\langle r(t)\rangle$ at time t and the stimulus at time $t+\tau$ as

$$Q_{rs}(x, y, \tau) = \frac{1}{T}\int_0^T dt\, \langle r(t)\rangle s(x, y, t+\tau)\,. \tag{2.10}$$

The spike-triggered average is related to the reverse correlation function (see chapter 1) by the equation $C(x, y, \tau) = Q_{rs}(x, y, -\tau)/\bar{r}$ where $\bar{r}$ is the average firing rate over the entire trial, $\bar{r} = \langle n\rangle/T$.

To estimate the firing rate of a neuron in response to a particular image, we add a static function F of the output of a linear filter of the stimulus to the background firing rate r_0, as was done in chapter 1,

$$r_{\text{est}}(t) = r_0 + F\left(L(t)\right)\,. \tag{2.11}$$

As in chapter 1, the linear estimate $L(t)$ is obtained by integrating over the past history of the stimulus with a linear kernel weighting function. Because the stimulus also depends on spatial location, we must decide how contributions from different image locations are to be combined to determine $L(t)$. The simplest assumption is that the contributions from different spatial points add linearly so that $L(t)$ is obtained by integrating over all x and y values,

$$L(t) = \int_0^\infty d\tau \int dx dy\, D(x, y, \tau) s(x, y, t-\tau)\,. \tag{2.12}$$

The kernel $D(x, y, \tau)$ that acts as the weighting factor in this integral determines how strongly, and with what sign, the visual stimulus at the point (x, y) and at time $t-\tau$ affects the firing rate of the neuron at time t. Note that the time and space variables play different roles in D; x and y represent *absolute* spatial positions in the visual field while τ represent a time *relative* to the time of stimulus estimation. As in chapter 1, the optimal kernel is given in terms of the response stimulus correlation function for a Gaussian white-noise stimulus with variance parameter σ_s^2 by

$$D(x, y, \tau) = \frac{Q_{rs}(x, y, -\tau)}{\sigma_s^2}\,. \tag{2.13}$$

Visually responsive neurons in the retina, LGN, and primary visual cortex are divided into two classes depending on whether the contributions from different locations within the visual field sum linearly, as assumed in equation 2.12, or not. As mentioned previously, X-cells in the cat retina and LGN, P-cells in the monkey retina and LGN, and many simple cells in primary visual cortex appear to satisfy this assumption. Other neurons, Y cells in the cat retina and LGN, M cells in the monkey retina and LGN, and complex cells in primary visual cortex, do not show linear summation across the spatial receptive field and nonlinearities must be included in models of their responses (as is done for complex cells later in this chapter).

The structure of the linear kernel $D(x, y, \tau)$ defines the full space-time receptive field of a neuron. Because $D(x, y, \tau)$ is a function of three variables, it can be difficult to measure and visualize. For some neurons, the kernel can be written as a product of two functions, one that describes the spatial receptive field and the other the temporal receptive field,

$$D(x, y, \tau) = D_s(x, y) D_t(\tau) . \tag{2.14}$$

Such neurons are said to have separable space-time receptive fields. Separability requires that the spatial structure of the receptive field does not change over time except by an overall multiplicative factor. When $D(x, y, \tau)$ cannot be written as the product of two terms as in equation 2.14, the neuron is said to have a nonseparable space-time receptive field.

The function F in equation 2.11 that determines the estimated firing rate typically contains constants used to set the firing rate to realistic values. We can take advantage of this and normalize $D(x, y, \tau)$ as we choose, later correcting for the arbitrary normalization by adjusting the parameters in F. We typically normalize D_s so that its integral is one, and use a similar rule for the components from which D_t is constructed. If x and y are measured in degrees and τ in seconds, this means that D_s has units of 1/degrees^2, D_t has units of 1/s or Hz, and D has units of Hz/degrees^2. With this convention, L has the same units as the stimulus, and it must be multiplied by an appropriate constant to be converted into a firing rate with units of Hz.

In discussing the response properties of retinal, LGN, and V1 neurons, we will not follow the path of the visual signal, nor the historical order of experimentation, but, instead, begin with primary visual cortex and then move back to the LGN and retina.

2.2 Modeling Simple Cells

2.2.1 Spatial Receptive Fields

Reverse correlation functions, and the corresponding linear kernels, of simple cells can be either separable or nonseparable. We start with separable kernels because they are simpler to describe. Figures 2.7A and C show the spatial structure of spike-triggered average stimuli (proportional to $D_s(x, y)$) for two simple cells in the primary visual cortex of a cat (area 17). The area over which the neurons are responsive to light is limited to a few degrees in each direction. The extent of this region depends on where the receptive field is located on the retina; receptive field size increases with eccentricity. In the monkey, receptive fields range in

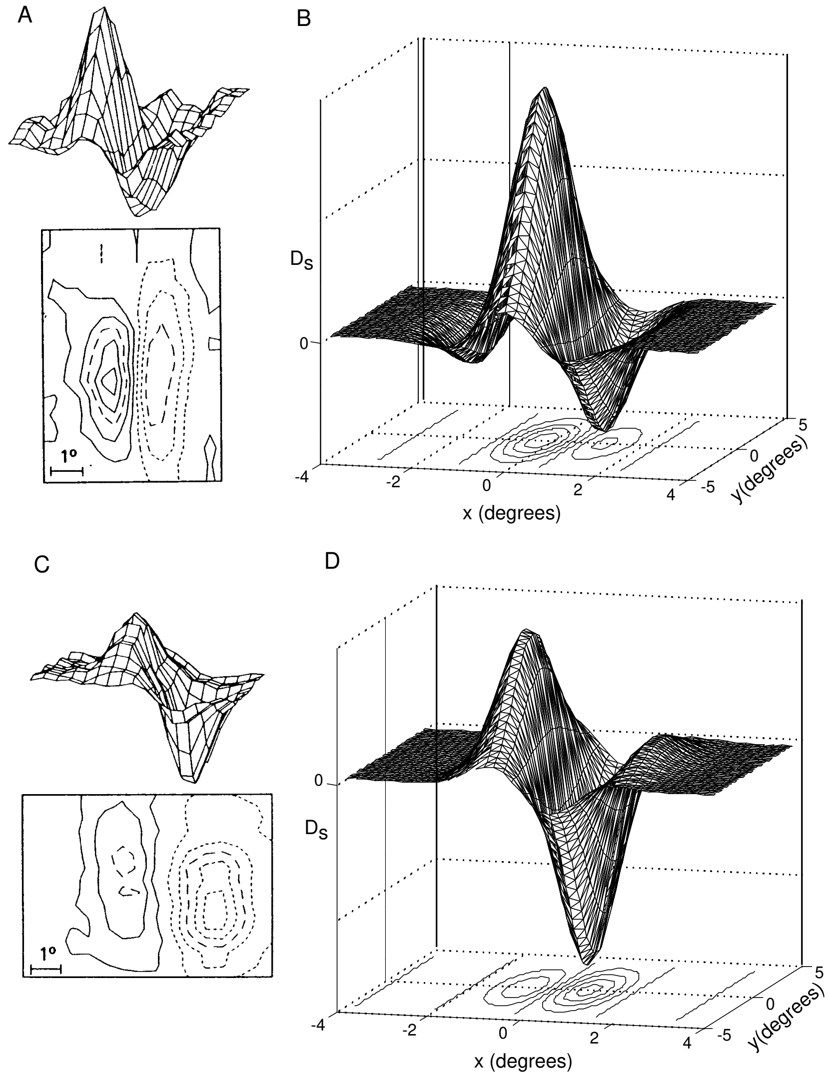

Figure 2.7: Spatial receptive field structure of simple cells. A) and C) Spatial structure of the receptive fields of two neurons in cat primary visual cortex determined by averaging stimuli between 50 ms and 100 ms prior to an action potential. The upper plots are three-dimensional representations with the horizontal dimensions acting as the x-y plane and the vertical dimension indicating the magnitude and sign of $D_s(x, y)$. The lower contour plots represent the x-y plane. Regions with solid contour curves are ON areas where $D_s(x, t) > 0$ and regions with dashed contours show OFF areas where $D_s(x, y) < 0$. B) and D) Gabor functions of the form 2.15 with $\sigma_x = 1^\circ$, $\sigma_x = 2^\circ$, $k = 1.8/^\circ$, and $\phi = 1 - \pi/2$ (B) or $\phi = 1 - \pi$ (D) chosen to match the receptive fields in A and C. (A and C from Jones and Palmer, 1987a)

size from around 0.1° at the fovea to a few degrees at the periphery. Receptive fields are often elongated in one direction, as they are in figures 2.7A and C. Within the receptive field, there are some regions where D_s is positive, called ON regions, and others where it is negative, called OFF regions. The neuron will respond vigorously if ON regions where $D_s > 0$ are highly illuminated ($s > 0$) while OFF regions where $D_s < 0$ are darkened ($s < 0$) relative to the background level of illumination. Because of these features, the neurons of figures 2.7A and C respond most vigorously to light-dark edges positioned along the border between the ON and OFF regions and oriented parallel to this border along the elongated direction of the receptive fields (figure 2.8). Figures 2.7 and 2.8 show receptive fields with two major subregions. Simple cells are found with from one to five subregions. Along with the ON-OFF patterns we have seen, another typical arrangement is a three-lobed receptive field with an OFF-ON-OFF or ON-OFF-ON arrangement as seen in figure 2.14B.

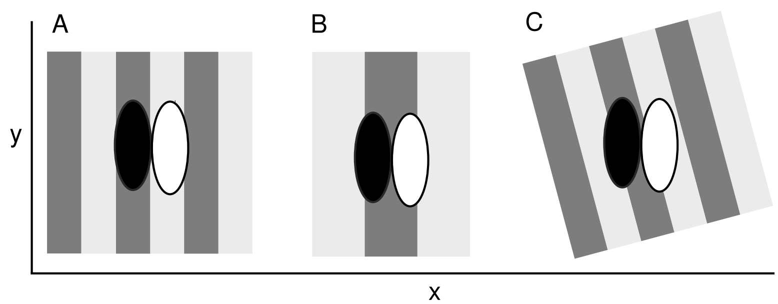

Figure 2.8: Grating stimuli superimposed on spatial receptive fields like those shown in figure 2.7. The receptive field is shown as two oval regions, one dark to represent an OFF area where $D_s < 0$ and one white to denote an ON region where $D_s > 0$. A) A grating with the spatial wavelength, orientation, and spatial phase (left-right position) shown produces a high firing rate because a dark band completely overlaps the OFF area of the receptive field and a light band overlaps the ON area. B) The grating shown is a non-optimal stimulus due to a mismatch in both the spatial phase and frequency, so that the ON and OFF regions of the receptive field each overlap with both light and dark stripes. C) The grating shown is at a non-optimal orientation so that each region of the receptive field overlaps both light and dark stripes.

A reasonable mathematical description of the spatial receptive field of a simple cell is provided by a Gabor function, which is a product of a Gaussian function and a sinusoidal function (see Daugman, 1985; Jones and Palmer, 1987b). To keep things manageable, we choose the coordinates x and y so that the borders between the ON and OFF regions are parallel to the y axis. We also choose the origin of the coordinates x and y to be at the center of the receptive field. With this choice, we can approximate the observed receptive field structures using the expression (Gabor function)

$$D_s(x,y) = \frac{1}{2\pi\sigma_x\sigma_y}\exp\left(-\frac{x^2}{2\sigma_x^2} - \frac{y^2}{2\sigma_y^2}\right)\cos(kx-\phi)\,. \qquad (2.15)$$

The parameters in this function determine the properties of the spatial receptive field: σ_x and σ_y determine its extent in the x and y directions respectively; k, the preferred spatial frequency, determines the spacing of light and dark bars that produce the maximum response (the preferred spatial wavelength is $2\pi/k$); and ϕ is the preferred spatial phase which determines where the ON-OFF boundaries fall within the receptive field. For this spatial receptive field, the sinusoidal grating of the form 2.6 that produces the maximum response for a fixed value of A has $K = k$, $\Phi = \phi$, and $\Theta = 0$.

Figure 2.9 shows x and y plots of a variety of Gabor functions with different parameter values, and figures 2.7B and D, show Gabor functions chosen specifically to match the data in figures 2.7A and C. As seen in figure 2.9, Gabor functions can have various types of symmetry, and variable numbers of significant oscillations (or subregions) within the Gaussian envelope. The number of subregions within the receptive field is determined by the value of $k\sigma_x$ and is typically expressed in terms of a quantity known as the bandwidth b. The bandwidth is defined as $b = \log_2(K_+/K_-)$ where $K_+ > k$ and $K_- < k$ are the spatial frequencies of gratings that produce one half the response amplitude of a grating with $K = k$. High bandwidths correspond to low values of $k\sigma_x$, meaning that the receptive field has few subregions and poor spatial frequency selectivity. Neurons with more subfields are more selective to spatial frequency, and they have smaller bandwidths and larger values of $k\sigma_x$. The base two logarithm used in its definition means that the bandwidth is the width of the spatial frequency tuning curve measured in octaves. For V1 neurons, bandwidths range from about 0.5 to 2.5.

The spatial frequency tuning curve as a function of K for a Gabor receptic field with preferred spatial frequency k and receptive field width σ_x is proportional to $\exp(-\sigma_x(k-K)^2/2)$ (see equation 2.28 below). The values of K_+ and K_- needed to compute the bandwidth are thus determined by the condition $\exp(-\sigma_x(k-K_\pm)^2/2) = 1/2$. Solving this equation gives $K_\pm = k \pm (2\ln(2))^{1/2}/\sigma_x$ from which we obtain

$$b = \log_2\left(\frac{k\sigma_x + \sqrt{2\ln(2)}}{k\sigma_x - \sqrt{2\ln(2)}}\right) \qquad (2.16)$$

or

$$k\sigma_x = \sqrt{2\ln(2)}\,\frac{2^b+1}{2^b-1}\,. \qquad (2.17)$$

Bandwidth is only defined if $k\sigma_x > \sqrt{2\ln(2)}$, but this is usually the case for V1 neurons.

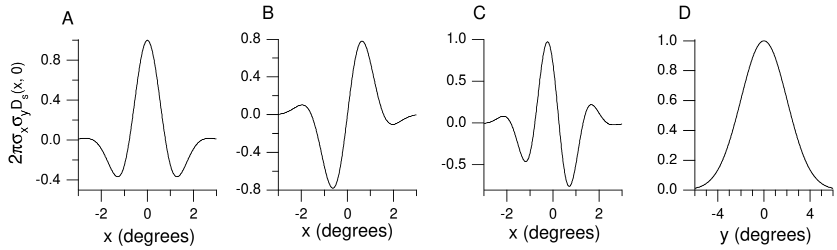

Figure 2.9: Gabor functions of the form given by equation 2.15. For convenience we plot the dimensionless function $2\pi\sigma_x\sigma_y D_s$. A) A Gabor function with $\sigma_x = 1^\circ$, $k = 2/^\circ$, and $\phi = 0$ plotted as a function of x for $y = 0$. This function is symmetric about $x = 0$. B) A Gabor function with $\sigma_x = 1^\circ$, $k = 2/^\circ$, and $\phi = \pi/2$ plotted as a function of x for $y = 0$. This function is antisymmetric about $x = 0$ and corresponds to using a sine instead of a cosine function in equation 2.15. C) A Gabor function with $\sigma_x = 1^\circ$, $k = 3/^\circ$, and $\phi = \pi/4$ plotted as a function of x for $y = 0$. This function has no particular symmetry properties with respect to $x = 0$. D) A Gabor function with $\sigma_y = 2^\circ$ plotted as a function of y for $x = 0$. This function is simply a Gaussian.

The response characterized by equation 2.15 is maximal if light-dark edges are parallel to the y axis, so the preferred orientation angle is zero. An arbitrary preferred orientation θ can be generated by rotating the coordinates, making the substitutions in equation 2.15

$$x \rightarrow x\cos(\theta) + y\sin(\theta) \quad y \rightarrow y\cos(\theta) - x\sin(\theta)\,. \tag{2.18}$$

This produces a spatial receptive field that is maximally responsive to a grating with $\Theta = \theta$. Similarly, a receptive field centered at the point (x_0, y_0) rather than at the origin can be constructed by making the substitutions

$$x \rightarrow x - x_0 \quad y \rightarrow y - y_0\,. \tag{2.19}$$

Gabor functions are by no means the only functions that can be used to fit spatial receptive fields of simple cells, nor do they necessarily provide the best fits. For example, gradients of Gaussians can be used rather than cosines times Gaussians.

Temporal Receptive Fields

Figure 2.10 shows a series of snap shots of a spike-triggered average stimulus that reveals the temporal development of the spatial receptive field of a neuron in the cat primary visual cortex. More than 300 ms prior to a spike, there is little correlation between the visual stimulus and the upcoming spike. Around 210 ms before the spike ($\tau = 210$ ms), a two-lobed OFF-ON receptive field, similar to the ones in figures 2.7A and C, is evident. As τ decreases (recall that τ measures time in a reversed sense), this structure first fades away and then reverses, so that the receptive field 75 ms before a spike has the opposite sign from what appeared at $\tau = 210$ ms. Due to latency effects, the spatial structure of the receptive field is less significant for $\tau < 75$ ms. The stimulus preferred by this cell is an appropriately aligned dark-light boundary that reverses to a light-dark boundary over time.

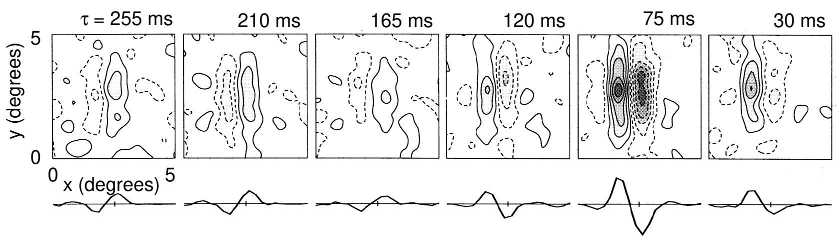

Figure 2.10: Temporal evolution of a spatial receptive field. Each panel is a plot of $D(x, y, \tau)$ for a different value of τ. As in figure 2.7, regions with solid contour curves are areas where $D(x, y, \tau) > 0$ and regions with dashed contours have $D(x, y, \tau) < 0$. The curves below the contour diagrams are one-dimension plots of the receptive field as a function of x alone. The receptive field is maximally different from zero for $\tau = 75$ ms with the spatial receptive field reversed from what it was at $\tau = 210$ ms. (From DeAngelis *et al*, 1995)

Reversal effect like those seen in figure 2.10 are a common feature of space-time receptive fields. If the space-time receptive field is separable, the reversal can be described by a function $D_t(\tau)$ that rises from zero, becomes positive, then negative, and ultimately goes to zero as τ increases. Adelsen and Bergen (1985) proposed the function shown in Figure 2.11,

$$D_t(\tau) = \alpha \exp(-\alpha\tau)\left(\frac{(\alpha\tau)^5}{5!} - \frac{(\alpha\tau)^7}{7!}\right) \tag{2.20}$$

for $\tau \geq 0$, and $D_s(\tau) = 0$ for $\tau < 0$. Here, α is a constant that sets the scale for the temporal development of the function. Single phase responses are also seen for V1 neurons and these can be modeled by eliminating the second term in equation 2.20. Three-phase responses, which are sometimes seen, must be described by a more complicated function.

Response of a Simple Cell to an Oscillating Grating

The function $L(t)$ that determines the linear response of a simple cell with a separable space-time receptive field can be computed in closed form for an oscillating grating stimulus (equation 2.6). This reveals the selectivity of the modeled response. The separable receptive field is given by the product of the spatial factor in equation 2.15 and the temporal factor in 2.20. Because the spatial and temporal aspects of the response separate,

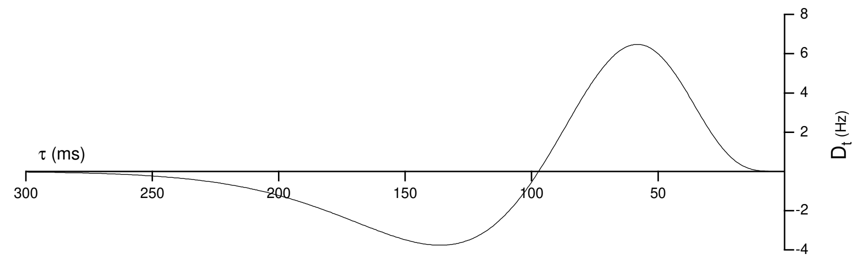

Figure 2.11: Temporal structure of a receptive field. The function $D_t(\tau)$ of equation 2.20 for $\alpha = 1/(15\text{ ms})$.

we can write the linear response as a product of two terms,

$$L(t) = L_s L_t(t) \tag{2.21}$$

where

$$L_s = \int dxdy\, D_s(x,y) A \cos\left(Kx\cos(\Theta) + Ky\sin(\Theta) - \Phi\right) . \tag{2.22}$$

and

$$L_t(t) = \int_0^\infty d\tau\, D_t(\tau) cos\left(\omega(t-\tau)\right) . \tag{2.23}$$

For simplicity, we consider a spatial receptive field with $\sigma_x = \sigma_y = \sigma$ and $\theta = 0$. Performing the integrals in equations 2.22 and 2.23, we find

$$\begin{aligned} L_s \quad = \quad & \frac{A}{2}\exp\left(-\frac{\sigma^2(k^2+K^2)}{2}\right)\left(\cos(\phi-\Phi)\exp\left(\sigma^2 kK\cos(\Theta)\right)\right. \\ & \left. + \cos(\phi+\Phi)\exp\left(-\sigma^2 kK\cos(\Theta)\right)\right) . \end{aligned} \tag{2.24}$$

and

$$L_t(t) = \frac{\alpha^6|\omega|\sqrt{\omega^2+4\alpha^2}}{(\omega^2+\alpha^2)^4}\cos(\omega t - \delta) . \tag{2.25}$$

with

$$\delta = \arctan\left(\frac{\omega}{\alpha}\right) + 8\arctan\left(\frac{2\alpha}{\omega}\right) - \pi . \tag{2.26}$$

To show the selectivity of this response, we plot (in figure 2.12) L_s as functions of the parameters Θ, K, and Φ that determine the orientation, spatial frequency, and spatial phase of the stimulus. It is also useful to write out the spatial part of the linear response L_s for various special parameter values. If the spatial phase of the stimulus and the preferred spatial phase of the receptive field are zero ($\Phi = \phi = 0$), we find that

$$L_s = A\exp\left(-\frac{\sigma^2(k^2+K^2)}{2}\right)\cosh\left(\sigma^2 kK\cos(\Theta)\right) \tag{2.27}$$

which determines the orientation and spatial frequency tuning for an optimal spatial phase. For a grating with the preferred orientation $\Theta = 0$ and a spatial frequency that is not too small, equation 2.24 can be simplified by noting that $\exp(-\sigma^2 kK) \approx 0$ for the values of $k\sigma$ normally encountered. Using this approximation we find, for this case,

$$L_s = \frac{A}{2} \exp\left(-\frac{\sigma^2(k-K)^2}{2}\right) \cos(\phi - \Phi) \qquad (2.28)$$

which reveals how the response depends on spatial phase for the optimal orientation and spatial frequencies in the appropriate range.

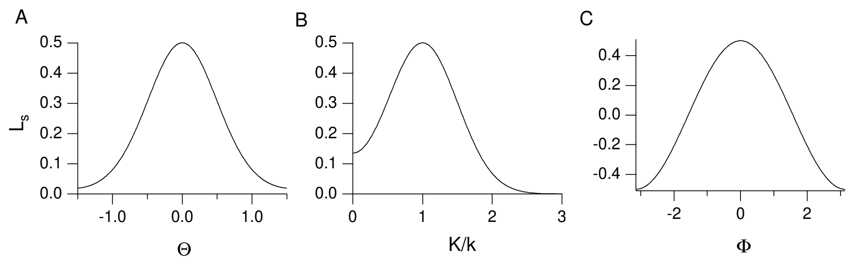

Figure 2.12: Selectivity of a Gabor filter $\theta = \phi = 0$, $\sigma_x = \sigma_y = \sigma$ and $k\sigma = 2$ acting on a cosine grating with $A = 1$. A) L_s as a function of stimulus orientation Θ for a grating with the preferred spatial frequency and phase, $K = k$ and $\Phi = 0$. B) L_s as a function of the ratio of the stimulus spatial frequency to its preferred value, K/k, for a grating oriented in the preferred direction $\Theta = 0$ and with the preferred phase $\Phi = 0$. C) L_s as a function of stimulus spatial phase Φ for a grating with the preferred spatial frequency and orientation, $K = k$ and $\Theta = 0$.

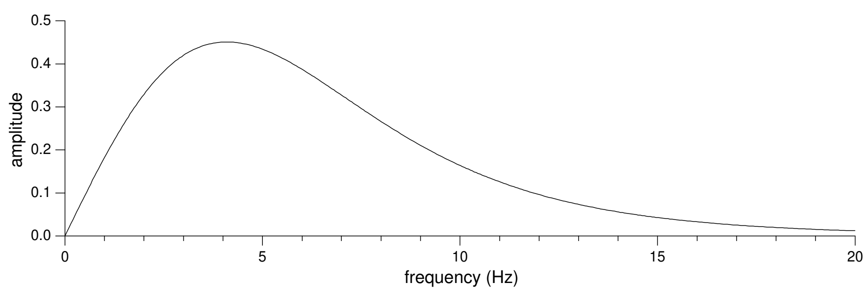

Figure 2.13: Frequency response of the simple cell model. The amplitude of the sinusoidal oscillations of $L_t(t)$ produced by a counterphase grating is plotted as a function of the temporal oscillation frequency $\omega/2\pi$.

The temporal frequency dependence of the linear response amplitude is determined by the factor multiplying the cosine function in equation 2.25. This is plotted as a function of temporal frequency of the stimulus ($\omega/2\pi$ rather than the angular frequency ω) in figure 2.13. The peak value around 4 Hz and roll off above 10 Hz are typical for V1 neurons and cortical neurons in a number of other primary sensory areas.

Space-Time Receptive Fields

It is instructive to display the function $D(x, y; \tau)$ in a space-time plot rather than as a sequence of spatial plots (as in figure 2.10). To do this, we suppress the y dependence and plot x-τ projections of the space-time kernel. Space-time plots of receptive fields from two simple cells of the cat primary visual cortex are shown in figure 2.14. The receptive field in figure 2.14A is approximately separable while that of figure 2.14B is nonseparable. The receptive field in figure 2.14A has OFF and ON subregions that reverse to ON and OFF subregions as a function of τ similar to the reversal seen in figure 2.10. Figure 2.15A shows an x-τ contour plot of a separable

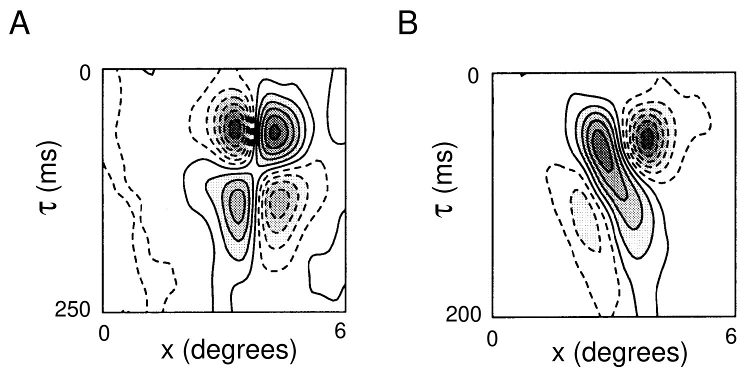

Figure 2.14: Space-time receptive fields of two simple cells from cat primary visual cortex. OFF regions are shown with dashed contour lines and ON regions with solid contour lines. A) An x-τ plot of an approximately separable space-time receptive field. The receptive field has side-by-side OFF and ON regions that reverse as a function of τ. B) An x-τ plot of a nonseparable space-time receptive field. The receptive field has a central ON region and two flanking OFF regions that shift to the left over time. (Adapted from DeAngelis *et al*, 1995)

space-time kernel, similar to the one shown in figure 2.14A, generated by multiplying a Gabor function by the temporal kernel of equation 2.20.

We can also plot the visual stimulus in a space-time diagram, suppressing the y coordinate by assuming that the image does not vary as a function of y. For example, figure 2.16A shows a grating of vertically oriented stripes moving to the left on an x-y plot. In the x-t plot of figure 2.16B, this image appears as a series of sloped dark and light bands. These represent the projection of the image in figure 2.16A onto the x axis evolving as a function of time. The leftward slope of the bands corresponds to the leftward movement of the image.

Figures 2.17 and 2.18 show how estimates of the responses produced by various types of images can be constructed from knowledge of the receptive field structure. The image in figures 2.17A-C is a dark bar that is flashed on for a brief period of time. To estimate the linear response at different times we show a cartoon of the space-time receptive field struc-

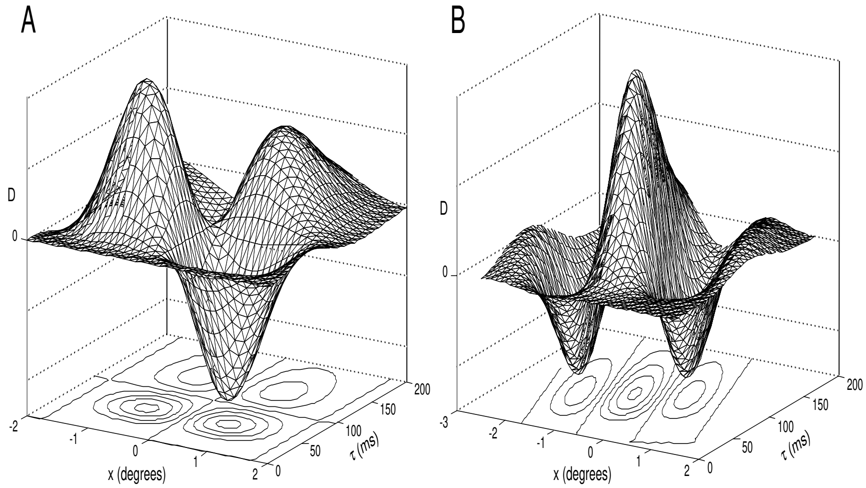

Figure 2.15: Mathematical descriptions of space-time receptive fields. A) A separable space-time receptive field constructed by multiplying a Gabor function (evaluated at $y = 0$) with $\sigma_x = 1^\circ$, $k = 1.8/^\circ$, and $\phi = \pi/2$ by the temporal kernel of equation 2.20 with $\alpha = 15$ ms. B) A nonseparable space-time receptive field constructed from equations 2.29 - 2.31. The Gabor function used (evaluated at $y = 0$) had $\sigma_x = 1^\circ$, $k = 2/^\circ$, and $\phi = 0$. D_t is given by the expression in equation 2.20 with $\alpha = 20$ ms except that the second term, with the seventh power function, was omitted because the receptive field does not reverse sign in this example. The x-τ rotation angle used was $\psi = \pi/9$ and the conversion factor was $c = 0.02\ ^\circ/\text{ms}$.

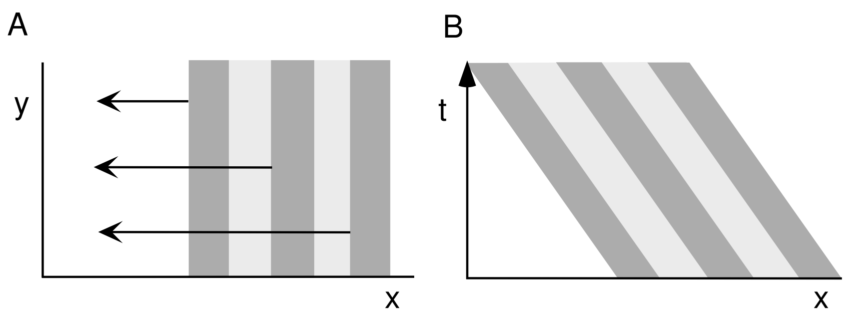

Figure 2.16: Space and space-time diagrams of a moving grating. A) A vertically oriented grating moves to the left on a two-dimensional screen. B) The space-time diagram of the image in A. The x location of the dark and light bands moves to the left as time progresses upward, representing the motion of the grating.

ture positioned at three different times in figures 2.17A, B, and C. The time corresponding to the height of the horizontal axis of the receptive field diagram is the time when the estimation is being made. The integral of the linear kernel times the stimulus can be visualized by noting how the OFF (black) and ON (white) regions overlap the image. Figure 2.17A corresponds to an estimate of $L(t)$ at the moment when the image first

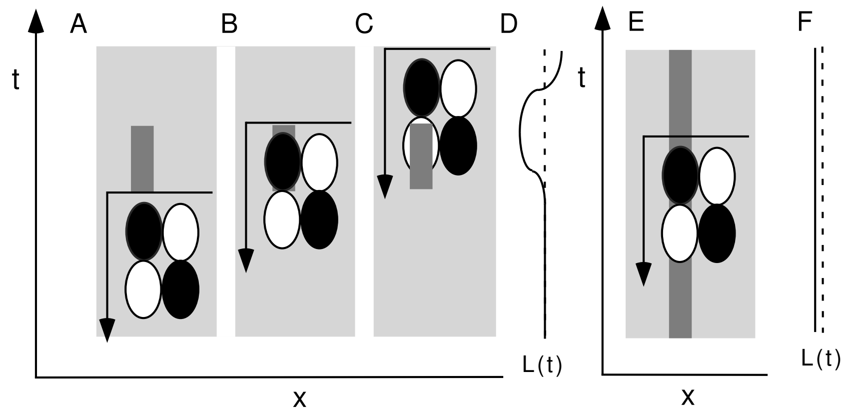

Figure 2.17: Responses to dark bars estimated from a separable space-time receptive field. Dark ovals in the receptive field diagrams are OFF regions and light circles are ON regions. The linear response at any time is determined by positioning the receptive field diagram so that its horizontal axis matches the time of response estimation and noting how the OFF and ON regions overlap with the image. A-C) The image is a dark bar that is flashed on for a short interval of time. This inhibits the response (A) until the dark image overlaps the OFF region (B) when $L(t) > 0$. The response is inhibited later when the dark bar overlaps the ON region (C) and $L(t) < 0$. D) A plot of $L(t)$ versus time corresponding to the responses generated in A-C. Time runs vertically in this plot, and $L(t)$ is plotted horizontally with the dashed line indicating the zero axis. E) The image is a static dark bar neither flashed nor moving. The bar overlaps both an OFF and an ON region generating opposing positive and negative contributions to $L(t)$. F) The weak response corresponding to E, plotted as in D.

appears. At this time, $L(t) = 0$. As time progresses, the receptive field diagram moves upward to later times. Figure 2.17B generates an estimate at the moment of maximum response when the dark image overlaps the OFF area of the space-time receptive field, producing a positive contribution to $L(t)$. Figure 2.17C shows a later time when the dark image overlaps an ON region, generating a negative $L(t)$. The response for this flashed image is thus transient firing followed by an inhibition of response as shown in Figure 2.17D.

Figures 2.17E and F show that a static dark bar is not an effective image for this receptive field structure. The static bar overlaps both the OFF region for small τ and the reversed ON region for large τ, generating opposing positive and negative contributions to $L(t)$. The flashed dark bar of figures 2.17A-C is a more effective stimulus because there is a time when it overlaps only the OFF region. Figure 2.18 shows an even more effective stimulus, a moving grating. The grating moves to the left in 2.18A-C. At the time corresponding to the positioning of the receptive field diagram in 2.18A, a dark band stimulus overlaps both OFF regions and light bands overlap both ON regions. Thus, all four regions contribute positive amounts to $L(t)$. As time progresses and the receptive field moves upward in the fig-

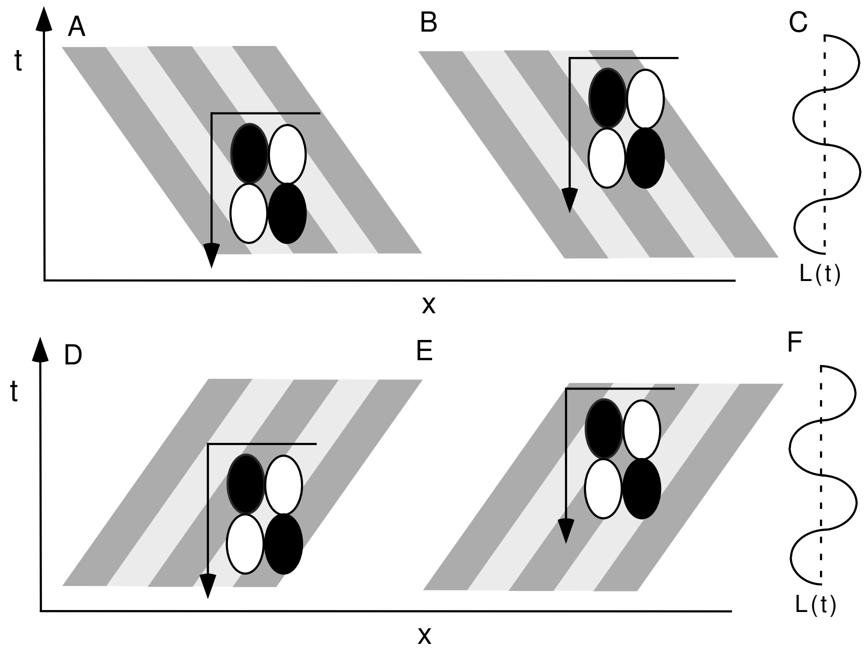

Figure 2.18: Responses to moving gratings estimated from a separable space-time receptive field. The receptive field is the same as in figure 2.17. A-C) The stimulus is a grating moving to the left. At the time corresponding to A, OFF regions overlap with dark bands and ON regions with light bands generating a strong response. At the time of the estimate in B, the alignment is reversed, and $L(t)$ is negative. C) A plot of $L(t)$ versus time corresponding to the responses generated in A-B. Time runs vertically in this plot and $L(t)$ is plotted horizontally with the dashed line indicating the zero axis. D-F) The stimulus is a grating moving to the right. The responses are identical to those in A-C.

ure, the alignment will sometimes be optimal, as in 2.18A, and sometimes be the opposite with light bands overlapping OFF regions and dark bands overlapping ON regions, as in 2.18B. This produces an $L(t)$ that oscillates as a function of time between positive and negative values (2.18C). Figures 2.18D-F show that a neuron with this receptive field responds equally to a grating moving to the right. Like the left-moving grating in figures 2.18A-C, the right-moving grating can overlap the receptive field in an optimal manner (2.18D) producing a strong response, or in a maximally negative manner (2.18E) producing strong suppression of response, again resulting in an oscillating response (2.18F). Separable space-time receptive fields can produce responses that are maximal for certain speeds of grating motion, but they are not sensitive to the direction of motion.

Nonseparable Receptive Fields

The reverse correlation functions for some neurons are not separable. An example is shown in figure 2.14B. This neuron has a three-lobed OFF-ON-OFF spatial receptive field, and these subregions shift to the left as time moves forward (and τ decreases). This means that the optimal stimulus for this neuron has light and dark areas that move toward the left. One way to model a nonseparable receptive field structure is to use a separable function constructed from a product of a Gabor function for D_s and equation 2.20 for D_t, but express these as functions of a mixture of the x and τ variables. The rotation of the space-time receptive field, as seen in figure 2.15B for example, is acheived by rotating the space and time coordinates using the transformation

$$D(x, y, \tau) = D_s(x', y)D_t(\tau') \tag{2.29}$$

with

$$x' = x\cos(\psi) + c\tau\sin(\psi) \tag{2.30}$$

and

$$\tau' = \tau\cos(\psi) - \frac{x}{c}\sin(\psi)\,. \tag{2.31}$$

The factor c converts between the units of time (ms) and space (degrees) and ψ is the space-time rotation angle. A reverse correlation function generated by this procedure similar to the one shown in figure 2.14B is plotted in figure 2.15B. The rotation operation is not the only way to generate nonseparable space-time receptive fields. They are often constructed by adding together two or more separable space-time receptive fields with different spatial and temporal characteristics.

Figure 2.19 shows how a nonseparable space-time receptive field can produce a response that is sensitive to the direction of motion of a grating. Figures 2.19A-C show a left-moving grating and, in 2.19A, the cartoon of the receptive field is positioned at a time when a light area of the image overlaps the central ON region and dark areas overlap the flanking OFF regions. This produces a large positive $L(t)$. At other times, the alignment is maximally non-optimal (2.19B), and over time, $L(t)$ oscillates between fairly large positive and negative values (2.19C). The nonseparable space-time receptive field does not overlap optimally with the right-moving grating of figures 2.19D-F at any time and the response is correspondingly weaker (2.19F). Thus, a neuron with a nonseparable space-time receptive field can be selective for the direction of motion of a grating.

2.3 Static Nonlinearities

Once the linear response estimate $L(t)$ has been computed, the firing rate of a visually responsive neuron can be approximated by using equation

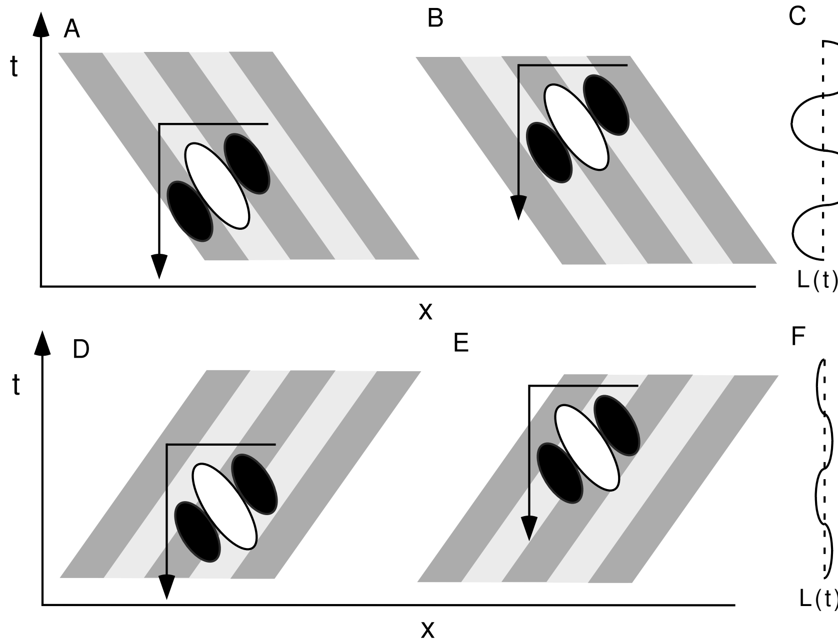

Figure 2.19: Responses to moving gratings estimated from a nonseparable space-time receptive field. Dark areas in the receptive field diagrams represent OFF regions and light areas ON regions. A-C) The stimulus is a grating moving to the left. At the time corresponding to A, OFF regions overlap with dark bands and the ON region overlaps a light band generating a strong response. At the time of the estimate in B, the alignment is reversed, and $L(t)$ is negative. C) A plot of $L(t)$ versus time corresponding to the responses generated in A-B. Time runs vertically in this plot and $L(t)$ is plotted horizontally with the dashed line indicating the zero axis. D-F) The stimulus is a grating moving to the right. Because of the tilt of the space-time receptive field, the alignment with the right-moving grating is never optimal and the response is weak (F).

2.11. For simplicity, we ignore the background firing rate in this section by setting $r_0 = 0$. In this case, the firing rate is given by $r_{\mathrm{est}}(t) = F(L(t))$ where F is an appropriately chosen function, the static nonlinearity. The simplest choice for F, consistent with the positive nature of firing rates, is rectification, $F = g[L]_+$ with g a scaling factor introduced to fit the magnitude of measured firing rates. This choice makes the firing rate a linear function of contrast, which does not match the data on contrast dependence of visual responses. Neural responses saturate as the contrast of the image increases and are more accurately described by $r \propto A^n/(A_{1/2}^n + A^n)$ where $A_{1/2}$ is a parameter equal to the contrast that produces a half-maximal response, and n is near two. This led Heeger (1992) to propose that an appropriate static nonlinearity to use is

$$F = \frac{g[L]_+^2}{A_{1/2}^2 + g[L]_+^2} \tag{2.32}$$

which produces the correct contrast dependence. Heeger (1992 and 1993) has considered a number of variants and extensions of this idea suggesting, for example, that the denominator of this expression should include L factors for additional neurons with nearby receptive fields. This can account for the effects of visual stimuli outside the 'classical' receptive field. Discussion of these effects is outside the scope of this chapter.

Modeling Complex Cells

Neurons in primary visual cortex are divided into two broad classes, simple cells and complex cells. Linear methods, such as spike-triggered averages, are useful for revealing the properties of simple cells, at least to a first approximation. Complex cells, on the other hand, display features that are fundamentally incompatible with a linear analysis (see Movshon *et al*, 1978a & b, for example). The spatial receptive fields of complex cells cannot be divided into separate ON and OFF regions that sum linearly to generate the response. Areas where light and dark images excite the neuron overlap, and the resulting cancellation between ON and OFF responses makes it difficult to measure and interpret spike-triggered average stimuli. The responses of complex cells to either moving or counterphase gratings show little dependence on spatial phase. Nevertheless, complex cells are as selective to the spatial frequency and orientation of a grating as simple cells. Thus, a complex cell is selective for a particular type of grating independent of its spatial position. This form of translationally invariant selectivity may represent an early stage in the visual processing that ultimately leads to recognition of objects independent of their position in the visual field. Finally, the firing rate of a complex cell responding to a stimulus oscillating with frequency ω, like the counterphase grating of equation 2.6, oscillates with frequency 2ω, a phenomenon known as frequency doubling.

Even though spike-triggered average stimuli and reverse correlation functions fail to capture the response properties of complex cells, complex-cell responses can be modeled, to a first approximation, by a relatively straightforward extension of the reverse correlation approach. The key observation (Adelson and Bergen, 1985; Pollen and Ronner, 1982) comes from equation 2.28, which shows how the linear response of a simple cell depends on spatial phase for an optimally oriented grating (and for K not too small). Consider two such responses with preferred spatial phases ϕ and $\phi - \pi/2$. Including both the spatial and temporal response factors 2.28 and 2.25, we find

$$L_1 = \frac{A\alpha^6|\omega|\sqrt{\omega^2 + 4\alpha^2}}{2(\omega^2+\alpha^2)^4} \exp\left(-\frac{\sigma^2(k-K)^2}{2}\right)\cos(\phi - \Phi)\cos(\omega t - \delta) \tag{2.33}$$

for preferred spatial phase ϕ, and

$$L_2 = \frac{A\alpha^6|\omega|\sqrt{\omega^2+4\alpha^2}}{2(\omega^2+\alpha^2)^4}\exp\left(-\frac{\sigma^2(k-K)^2}{2}\right)\sin(\phi-\Phi)\cos(\omega t-\delta) \tag{2.34}$$

for preferred spatial phase $\phi - \pi/2$, since $\cos(\phi - \pi/2 - \Phi) = \sin(\phi - \Phi)$. If we square and add these two terms, we obtain a result that does not depend on Φ,

$$L_1^2 + L_2^2 = \frac{A^2\alpha^{12}\omega^2(\omega^2+4\alpha^2)}{4(\omega^2+\alpha^2)^8}\exp\left(-\sigma^2(k-K)^2\right)\cos^2(\omega t-\delta)\,, \tag{2.35}$$

because $\cos^2(\phi-\Phi)+\sin^2(\phi-\Phi) = 1$. Thus, we can model the response

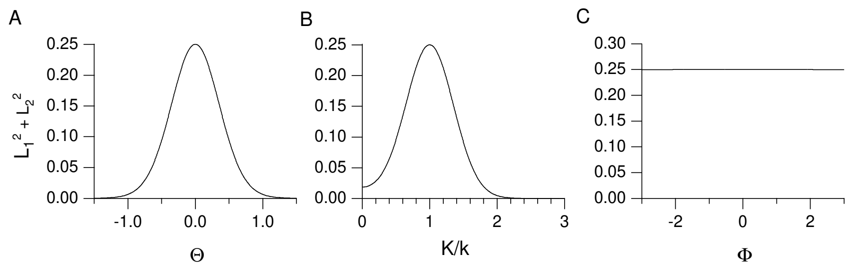

Figure 2.20: Selectivity of a complex cell model responding to a sinusoidal grating. The width and preferred spatial frequency of the Gabor functions underlying the model satisfy $k\sigma = 2$. The response is plotted for a spatial frequency and at a time for which $A\cos^2(\omega t-\delta)\alpha^{12}\omega^2(\omega^2+4\alpha^2)/(4(\omega^2+\alpha^2)^8) = 1$. A) The complex cell response $L_1^2 + L_2^2$ as a function of stimulus orientation Θ for a grating with the preferred spatial frequency $K = k$. B) $L_1^2 + L_2^2$ as a function of the ratio of the stimulus spatial frequency to its preferred value, K/k, for a grating oriented in the preferred direction $\Theta = 0$. C) $L_1^2 + L_2^2$ as a function of stimulus spatial phase Φ for a grating with the preferred spatial frequency and orientation $K = k$ and $\Theta = 0$.

of a complex cell by writing

$$r(t) = r_0 + g\left(L_1^2 + L_2^2\right)\,. \tag{2.36}$$

The selectivities of such a model to grating orientation, spatial frequency, and spatial phase are shown in figure 2.20. The response of the model complex cell is tuned to orientation and spatial frequency, but the spatial phase dependence is absent. In computing the curve for figure 2.20C, we used the exact expressions for L_1 and L_2, not the approximation 2.28 used to simplify the discussion above. Although it is not visible in the figure, there is a weak dependence on Φ when the exact expressions are used (see exercise 6).

The complex cell response given by equations 2.36 and 2.35 reproduces the frequency doubling effect seen in complex cell responses because

$$\cos^2(\omega t-\delta) = \frac{1}{2}\cos\left(2(\omega t-\delta)\right) + \frac{1}{2}\,. \tag{2.37}$$

Figure 2.21 shows a comparison of model simple and complex cell responses to an oscillating grating and illustrates this phenomenon.

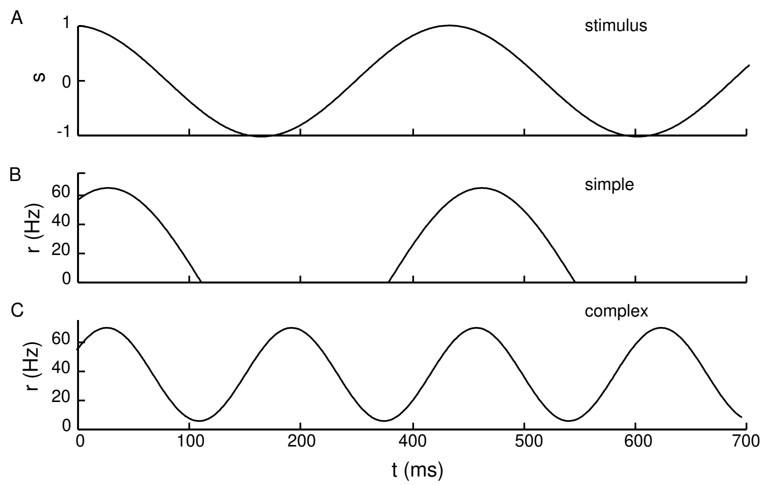

Figure 2.21: Temporal responses of model simple and complex cells to an oscillating grating. A) The stimulus s(x, y, t) at a given point (x, y) plotted as a function of time. B) The rectified linear response of a model simple cell to this grating with a temporal kernel given by equation 2.20 with $\alpha = 1/(15\text{ ms})$. C) The frequency doubled response of a model complex cell with the same temporal kernel but with the estimated rate given by a squaring operation rather than rectification. The background firing rate is $r_0 = 5$ Hz. Note the phase shift of both B and C relative to A, indicating a nonzero value of δ.

The complex cell model we have presented is called an 'energy' model because of its resemblance to the equation for the energy of a simple harmonic oscillator. The pair of linear filters used, with spatial phases separated by $\pi/2$ is called a quadrature pair. Because of rectification, the terms L_1^2 and L_2^2 cannot be constructed by squaring the outputs of single simple cells. However, they can each be constructed by summing the squares of rectified outputs from two simple cells with spatial phases separated by π. Thus, we can write the complex cell response as the sum of the squares of four rectified simple cell responses

$$r(t) = r_0 + g\left([L_1]_+^2 + [L_2]_+^2 + [L_3]_+^2 + [L_4]_+^2\right) , \tag{2.38}$$

where the four simple cells have spatial phases ϕ, $\phi + \pi/2$, $\phi + \pi$, and $\phi + 3\pi/2$. While such a construction is possible, it should not be interpreted too literally because complex cells receive input from many sources including the LGN and other complex cells. Rather, this model should be viewed as purely descriptive. The energy model is not the only way to generate responses resembling those of complex cells, but it is a clever and compact description.

2.4 Receptive Fields in the Retina and LGN

We end this discussion of the visual system by returning to the initial stages of the visual pathway and briefly describing the receptive field properties of neurons in the retina and LGN. The receptive fields of retinal ganglion cells and LGN relay neurons are extremely similar, so we do not treat them separately. The spatial structure of the receptive fields of retinal ganglion and LGN neurons has a center-surround structure consisting either of a circular central ON region surrounded by an annular OFF region, or the opposite arrangement of a central OFF region surrounded by an ON region. Such receptive fields are called ON-center or OFF-center respectively. Figure 2.22A shows the spatial receptive fields of an ON-center cat LGN neuron.

The spatial structure of retinal ganglion and LGN receptive fields is well-captured by the difference-of-Gaussians model (Rodieck, 1965) in which the spatial receptive field (the space-time receptive field at a given time or integrated over a time interval) is expressed as

$$D_s(x, y) = \pm\left(\frac{1}{2\pi\sigma_{\text{cen}}^2}\exp\left(-\frac{x^2+y^2}{2\sigma_{\text{cen}}^2}\right) - \frac{B}{2\pi\sigma_{\text{sur}}^2}\exp\left(-\frac{x^2+y^2}{2\sigma_{\text{sur}}^2}\right)\right). \tag{2.39}$$

Here the center of the receptive field has been placed at $x = y = 0$. The first Gaussian function in equation 2.39 describes the center and the second the surround. The size of the central region is determined by the parameter σ_{cen}, while σ_{sur}, which is greater than σ_{cen}, determines the size of the surround. B controls the magnitude of the response to constant illumination. The $\pm$ sign allows both ON-center ($+$) and OFF-center ($-$) cases to be represented. Figure 2.22B shows a spatial receptive field formed from the difference of two Gaussians that matches the receptive field structure in figure 2.22A.

Figure 2.22C indicates that, as in the case of V1 neurons, the spatial structure of the receptive field reverses over time with, in this case, a central ON region reversing to an OFF region as τ increases. Similarly, the OFF surround region changes to an ON region with increasing τ, although the reversal and the onset are slower for the surround than for the central region. Because of the difference between the time course of the center and surround regions, the space-time receptive field of an LGN cell is not separable, although the center and surround components are individually separable. The basic features of LGN neuron space-time receptive fields are captured by the mathematical caricature

$$D(x, y, \tau) = \pm\left(\frac{D_t^{\text{cen}}(\tau)}{2\pi\sigma_{\text{cen}}^2}\exp\left(-\frac{x^2+y^2}{2\sigma_{\text{cen}}^2}\right) - \frac{BD_t^{\text{sur}}(\tau)}{2\pi\sigma_{\text{sur}}^2}\exp\left(-\frac{x^2+y^2}{2\sigma_{\text{sur}}^2}\right)\right). \tag{2.40}$$

For a given neuron, the parameter B in this equation may take a different value from the B parameter in equation 2.39. Separate functions of time

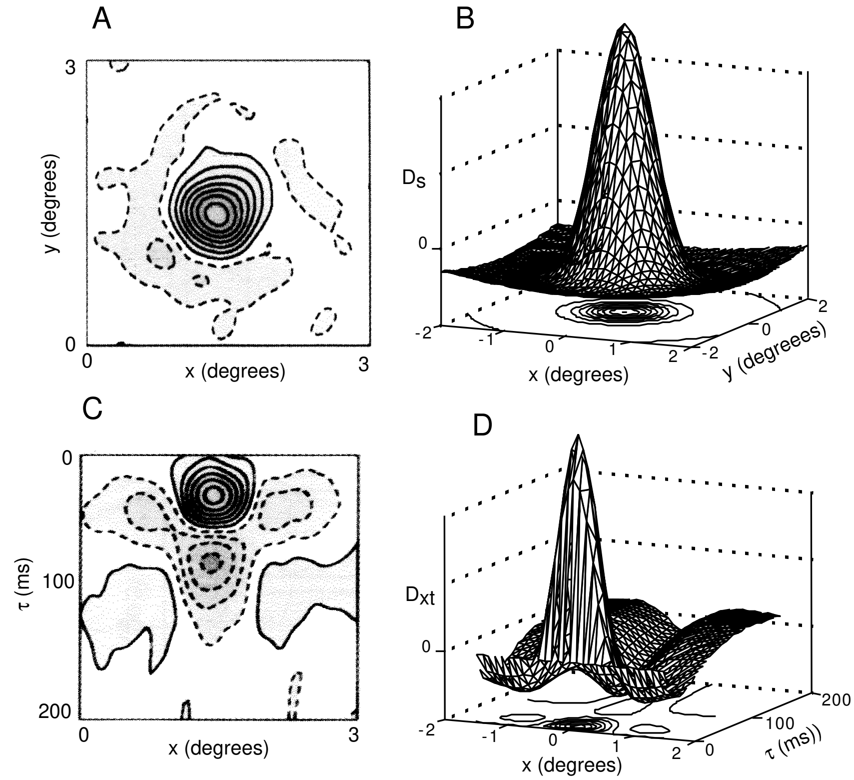

Figure 2.22: Receptive fields of LGN neurons. A) The center-surround spatial structure of the receptive field of a cat LGN X cell. This has a central ON region (solid contours) and a surrounding OFF region (dashed countours). B) A fit of the receptive field shown in A using a difference of Gaussian function (equation 2.39) with $\sigma_{\mathrm{cen}} = 0.3^\circ$, $\sigma_{\mathrm{sur}} = 1.5^\circ$, and $B = 5$. C) The space-time receptive field of a cat LGN X cell. Note that both the center and surround regions reverse sign as a function of τ and that the temporal evolution is slower for the surround than for the center. D) A fit of the space-time receptive field in C using 2.40 with the same parameters for the Gaussian functions as in B, and temporal factors given by equation 2.41 with $1/\alpha_{\mathrm{cen}} = 16$ ms for the center, $1/\alpha_{\mathrm{sur}} = 32$ ms for the surround, and $1/\beta_{\mathrm{cen}} = 1/\beta_{\mathrm{sur}} = 64$ ms. (A and B from DeAngelis *et al*, 1995)

multiply the center and surround, but they can both be described by the same functions using two sets of parameters,

$$D_t^{\mathrm{cen,sur}}(\tau) = \alpha_{\mathrm{cen,sur}}^2 t \exp(-\alpha_{\mathrm{cen,sur}} t) - \beta_{\mathrm{cen,sur}}^2 t \exp(-\beta_{\mathrm{cen,sur}} t) . \qquad (2.41)$$

The parameters α_{cen} and α_{sur} control the latency of the response in the center and surround regions respectively, and β_{cen} and β_{sur} affect the time for the reversal. This function has characteristics similar to the function 2.20 used for V1 neurons, but that the latency effect is less pronounced. Figure 2.22D shows the space-time receptive field of equation 2.40 with parameters chosen to match figure 2.22C.

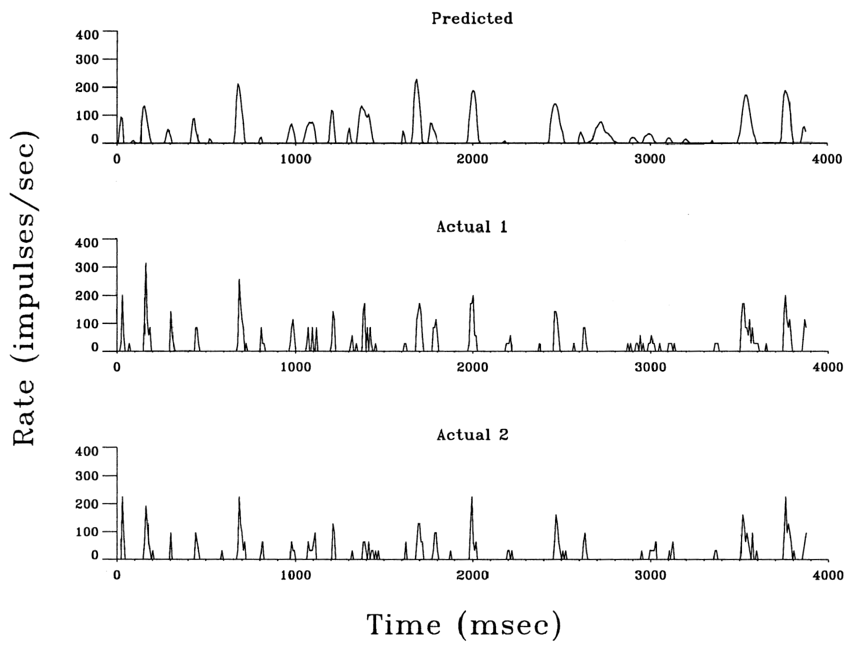

Figure 2.23: Comparison of predicted and measured firing rates for a cat LGN neuron responding to a video movie. The top panel is the rate predicted by integrating the product of the video image intensity and a linear filter obtained for this neuron by spike-triggered averaging a white-noise stimulus. The resulting linear prediction was rectified. The middle and lower panels are measured firing rates extracted from two different sets of trials. (From Dan *et al*, 1996)

Dan, Atick, and Reid (1996) have tested the predictions of a reverse correlation model of an LGN neuron in an interesting way. They first extracted the kernel needed to model a particular LGN cell using a white-noise stimulus. They then used this kernel and a rectifying static nonlinearity to predict the firing rate of the neuron in response to a video movie. The top panel in figure 2.23 shows the resulting prediction while the middle and lower panels show the actual firing rates extracted from two different groups of trials. The correlation coefficient between the predicted and actual firing rates was 0.5, which was very close to the correlation coefficient between firing rates extracted from different trial groups. This means that the accuracy of the prediction was predominantly limited by the variability of the neural response, not by the limitations of the prediction technique.

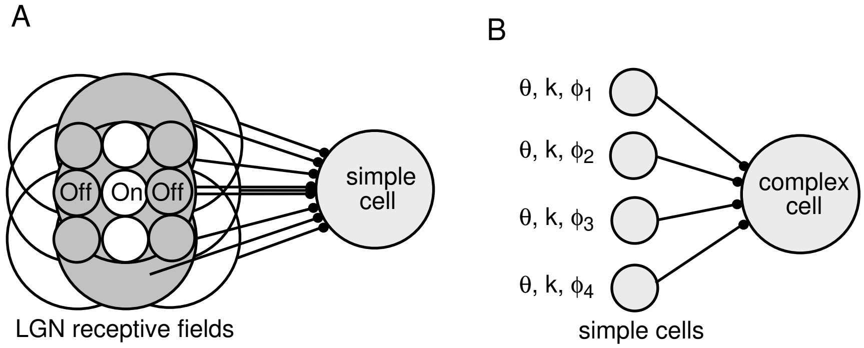

Figure 2.24: A) The Hubel-Wiesel model of orientation selectivity. The spatial arrangement of the receptive fields of nine LGN neurons are shown, with a row of three ON-center fields flanked on either side by rows of three OFF-center fields. White areas denote ON fields and grey areas OFF fields. In the model, the converging LGN inputs are summed linearly by the simple cell. This arrangement produces a receptive field oriented in the vertical direction. B) The Hubel-Wiesel model of a complex cell. Inputs from a number of simple cells with similar orientation and spatial frequency preferences (θ and k), but different spatial phase preferences (ϕ_1, ϕ_2, ϕ_3, and ϕ_4), converge on a complex cell and are summed. This produces a complex cell output that is selective for orientation and spatial frequency, but not for spatial phase. The figure shows four simple cells converging on a complex cell, but additional simple cells can be included to give a more complete coverage of spatial phase .

2.5 Constructing V1 Receptive Fields

The models of visual receptive fields we have been discussing are purely descriptive, but they provide an important framework for studying how the circuits of the retina, LGN, and primary visual cortex generate neural responses. For example, Hubel and Wiesel (1962) showed how the oriented receptive fields of cortical neurons could be generated by summing the input from appropriate selected LGN neurons. Their construction, shown in figure 2.24A, consists of alternating rows of ON-center and OFF-center LGN cells converging on a cortical simple cell. The left side of figure 2.24A shows the spatial arrangement of LGN receptive fields that, when summed, form bands of ON and OFF regions resembling the receptive field of an oriented simple cell. This model accounts for the selectivity of a simple cell purely on the basis of feedforward input from the LGN. We leave the study of this model to exercise 10. Other models, which we discuss in chapter 7, include the effects of recurrent intracortical connections as well.

Hubel and Wiesel (1962) also proposed the circuit model for complex cells depicted in figure 2.24B. Here the phase-invariant response of a complex cell is produced by summing together the responses of several simple cells with similar orientation and spatial frequency tuning, but different preferred spatial phases. In this model, the complex cell inherits its orientation and spatial frequency preference from the simple cells that drive it, but spatial phase selectivity is reduced because the outputs of simple cells with a variety of spatial phases selectivities are summed. Analysis of this model is also left to exercise 10. While the model generates complex cell responses, there are indications that complex cells in primary visual cortex are not exclusively driven by simple cell input. Alternate models are considered in chapter 7.

2.6 Further Reading and References

Further Reading

Carandini M,Heeger DJ & Movshon JA (1996) Linearity and gain control in V1 simple cells. In EG Jones & PS Ulinski (editors) *Cerebral Cortex. Volume X: Cortical Models*. NY:Plenum Press.

DeAngelis, GC, Ohzawa, I & Freeman, RD (1995) Receptive field dynamics in the central visual pathways. *Trends in Neuroscience* **18**:451-458.

De Valois RL & De Valois KK (1990) *Spatial Vision*. NY: Oxford.

Ferster D (1994) Linearity of synaptic interactions in the assembly of receptive fields in cat visual cortex. *Current Opinion in Neurobiology* **4**:563-568.

Gabor D (1946) Theory of communication./ *Journal of Instr. Electrical Engineering* **93**:429-457.

Hubel DH & Wiesel TN (1962) Receptive fields, binocular interaction and functional architecture in the cats visual cortex. *Journal of Physiology* **160**:106-154.

Hubel DH & Wiesel TN (1977) Functional architecture of macaque monkey visual cortex. *Procedings of the Royal Society of London* **B198**:1-59.

Jones, J & Palmer, L (1987a) The two-dimensional spatial structure of simple receptive fields in cat striate cortex. *Journal of Neurophysiology* **58**:1187-1211.

Movshon JA, Thompson ID, Tolhurst DJ (1978a) Spatial summation in the receptive fiels of somple cells in the cat's striate cortex. *Journal of Neurophysiology* **283**:53-77.

Movshon JA, Thompson ID, Tolhurst DJ (1978b) Spatial and temporal contrast sensitivity of neurones in areas 17 and 18 of the cat s visual cortex. *Journal of Neurophysiology* **283**:101-120.

Orban GA (1984) *Neuronal Operations in the Visual Cortex*. Berlin:Springer.

Wandell, B (1995) *Foundations of Vision*. Sunderland, MA:Sinauer.

References

Adelson EH, Bergen JR (1985) Spatiotemporal energy models for the perception of motion. *Journal of the Optical Society of America* **A2**:284-299.

Dan Y, Atick JJ & Reid RC (1996) Efficient coding of natural scenes in the lateral geniculate nucleus: Experimental test of a computational theory. *The Journal of Neuroscience* **16**:3351-3362.

Daugman, J.G. (1985) Uncertainty relation for resolution in space, spatial frequency, and orientation optimization by two-dimensional visual cortical filters. *Journal of the Optical Society of America* **2**:1160-1169.

DeAngelis GC, Ohzawa I & and Freeman RD (1991) Depth is encoded in the visual cortex by a specialized receptive field structure. *Nature* **352**:156-159.

Ermentrout GB & Cowan J (1979) A mathematical theory of visual hallucination patterns. *Biological Cybernetics* **34**:137-150.

Heeger DJ (1992) Normalization of cell responses in cat striate cortex. *Visual Neuroscience* **9**:181-198.

Heeger DJ (1993) Modeling simple-cell direction selectivity with normalized, half-squared, linear operators. *Journal of Neurophysiology* **70**:1885-1898.

Jones J & Palmer L (1987b) An Evaluation of the Two-Dimensional Gabor Filter Model of Simple Receptive Fields in Cat Striate Cortex. *Journal of Neurophysiology* **58**:1233-.

Pollen D & Ronner S (1982) Spatial computations performed by simple and complex cells in the visual cortex of the cat. *Vision Research* **22**:101-118.

Rovamo J & Virsu V (1984) Isotropy of cortical magnification and topography of striate cortex. *Vision Research* **24**:283-286.

Schwarz EL (1977) Spatial mapping in the primate sensory projection: analytic structure and relevance to perception. *Biological Cybernetics* **25**:181-194.

Tootell RB, Silverman MS, Switkes E & De Valois RL (1982) Deoxyglucose analysis of retinotopic organization in primate striate cortex. *Science* **218**:902-904.

Van Essen DC, Newsome WT & Maunsell JHR (1984) The visual field representation in striate cortex of the macaque monkey: Asymmetries, anisotropies, and individual variability. *Vision Research* **24**:429-448.

2.7 Exercises

1. Using the inverses of equations 2.3 and 2.5

$$\epsilon = \epsilon_0 \left(\exp(X/\lambda) - 1\right) \quad \text{and} \quad a = -\frac{180^\circ(\epsilon_0 + \epsilon)Y}{\lambda\epsilon\pi}\,,$$

map from cortical coordinates back to visual coordinates and determine what various patterns of activity in the primary visual cortex would 'look like'. Consider straight lines and bands of constant activity extending across the cortex at various angles. Ermentrout and Cowan (1979) used these results as a basis of a mathematical theory of visual hallucinations.

2. Using equation 2.28 with $\phi = \Phi = 0$, compute the spatial frequencies K_+ and K_- of gratings that produce half of the maximum linear response. Show that the bandwidth $b = \log_2(K_+/K_-)$ is given by equation 2.16 (with $\sigma_x = \sigma$).

3. Compute the integrals in equations 2.22 and 2.23 to obtain the results 2.24 and 2.25 and the selectivity curves in figures 2.12 and 2.13. The integrals can also be done numerically to obtain the results of figures 2.12 and 2.13 directly.

4. Compute the response of a model simple cell with a separable space-time receptive field to a moving grating

$$s(x, y, t) = \cos(Kx - \omega t) \ .$$

For D_s use equation 2.15 with $\sigma_x = \sigma_y = 1^\circ$, $\phi = 0$, and $k = 2/^\circ$. For D_t use equation 2.20 with $\alpha = 1/(15\text{ ms})$. Compute the linear response given by equation 2.12 and assume that the actual response is proportional to a rectified version of this linear response. Plot the response as a function of time for $K = k = 2/^\circ$ and $\omega = 8\pi/\text{s}$. Plot the response amplitude as a function of ω for $K = k = 2/^\circ$ and as a function of K for $\omega = 8\pi/\text{s}$.

5. Construct a model simple cell with the nonseparable space-time receptive field described in the caption of figure 2.15B (with $\theta = 0$). Compute its response to the moving grating of exercise 4. Plot the amplitude of the response as a function of the velocity of the grating, ω/K, using $\omega = 8\pi/\text{s}$ and varying K to obtain a range of both positive and negative velocity values (use negative K values for this).

6. Compute the response of a model complex cell to the moving grating of exercise 4. The complex cell should be modeled by squaring the linear response of the simple cell used in exercise 4, and adding this to the square of the response of a second simple cell with identical properties except that its spatial phase preference is $\phi = -\pi/2$ instead of $\phi = 0$. Plot the response as a function of time for $K = k = 2/^\circ$ and $\omega = 8\pi/\text{s}$. Plot the response amplitude as a function of ω for $K = k = 2/^\circ$ and as a function of K for $\omega = 8\pi/\text{s}$.

7. Determine the extent to which the response of a complex cell model constructed by summing the squared linear responses of a quadrature pair of Gabor filters is independent of the spatial phase for gratings of different spatial frequencies and orientations. Use the exact expression of equation 2.24, not the approximation 2.28.

8. Construct a model complex cell that is disparity tuned but insensitive to the absolute position of a grating. The complex cell is constructed by summing the squares of the responses of two simple cells, but disparity effects are now included. For this exercise, we ignore temporal factors and only consider the spatial dependence of the response. Each simple cell response is composed of two terms that correspond to inputs coming from the left and right eyes. Because of disparity, the spatial phases of the image of a grating in the two eyes, Φ_L and Φ_R, may be different. We write the spatial part of the linear response for a grating with the preferred spatial frequency ($k = K$) and orientation ($\Theta = \theta = 0$) as

$$L_1 = \frac{A}{2}\left(\cos(\Phi_L) + \cos(\Phi_R)\right)$$

assuming that $\phi = 0$ (this equation is a generalization of 2.28). Let the complex cell response be proportional to $L_1^2 + L_2^2$ where L_2 is similar to L_1 but with the cosine functions replaced by sine functions. Show that the response of this neuron is tuned to the disparity, $\Phi_L - \Phi_R$, and is independent of the absolute spatial phase of the grating, $\Phi_L + \Phi_R$. Plot the response tuning curve as a function of disparity. (See Ohzawa *et al*, 1991).

9. Determine the selectivity of the LGN receptive field of equation 2.39 to spatial frequency and of the temporal response function for LGN neurons, equation 2.41, to temporal frequency by computing their integrals when multiplied by cosine functions of space or time respectively. Use $\sigma_c = 0.3^\circ$, $\sigma_s = 1.5^\circ$, $B = 5$, $1/\alpha = 16$ ms, and $1/\beta = 64$ ms. Plot the resulting spatial and temporal frequency tuning curves.

10. Construct the Hubel-Wiesel simple and complex cell models of figure 2.24 (Hubel and Wiesel, 1962). Plot the spatial receptive field of the simple cell constructed in this way. Compare the result of summing appropriately placed LGN center-surround receptive fields (figure 2.24A) with the results of the Gabor filter model of the simple cell that uses the spatial kernel of equation 2.15. Also compare the responses of a complex cell constructed by summing the outputs of simple cells with different spatial phase preferences (figure 2.24B) with the complex cell model obtained by squaring and summing two simple cell responses with spatial phases 90° apart as in equation 2.36.