35
Algotecture of visual cortex

A. B. Watson*

Introduction

Our knowledge of the primary visual pathway is voluminous but fragmentary. We know a great deal about the behavior of cells in the major waystations of this pathway – the retina, lateral geniculate nucleus (LGN), and striate cortex – but know rather little about how they function together to process the visual image. At each level, we know about various classes of cells, and about the spatial, temporal, and chromatic aspects of their receptive fields. But we know rather little about how cells at one level converge upon those at the next, and we know still less about the function of these receptive fields, and why particular receptive fields are employed at each level. An answer to these questions constitutes a theory of the algorithmic architecture (*algotecture*) of the visual pathways. This theory must address data from many sources, and must adhere to numerous mathematical constraints. In what follows, I will enumerate some of the considerations that must guide the attack upon this fundamental problem in vision and brain science. I will then sketch the beginnings of a candidate theory for one part of cortex. The theory raises as many questions as it answers, but that is not necessarily a bad thing, and at least it provides a concrete hypothesis that can be tested, modified, or replaced by a better theory.

I will restrict my focus to one portion of the visual pathway: the oriented simple cells of primate striate cortex. This system, while much less than the whole primary visual pathway, is nonetheless immense and fundamental. It contains a substantial fraction of the visual cortex, and is the destination of the majority of cells in both retina and LGN. It is the putative basis of detailed form vision. If we could understand it, we would understand much of early visual processing. While my focus is on the cortex, a specific theme of this chapter is that cortical structure and function cannot be understood without reference to the structure and function of its tributaries, notable retina and LGN, so that some discussion of their properties is essential.

In pondering the striate cortex, one cannot help but be struck by the orderly, almost crystalline architecture, by the laminar and columnar organization, and by the properties of orientation and frequency tuning that are the hallmark of so many receptive fields. It is clearly an intricate analytical engine, but what does it compute, and how?

Constraints

The constraints that must guide a theory of visual cortex may be organized into the following categories: functional, mathematical, perceptual, physiological, and anatomical. Each category is reviewed briefly below. These reviews are partial and arguable.

Functional constraints

Functional constraints address the question of why the visual cortex does what it does. A most fundamental function of the cortex is to represent the visual image, to provide a 'sensorium' that subsequent visual processes may query (Koenderink & van Doorn, 1987). We may therefore think of the responses of cells of the cortex as providing a code of the visual image. It is important that the representation be complete, in the sense of not neglecting any information present in the image, at least within its chosen 'window of

visibility' (Watson, Ahumada & Farrell, 1986). It has also been argued that an important function of the cortical code is to decorrelate individual elements of the code (Buchsbaum, 1987; Field, 1987; Laughlin, this volume). In the retina, signals from neighboring receptors are highly correlated and hence redundant, while in the cortex, through the construction of more elaborate receptive fields, there is an opportunity to produce cells whose responses are much more nearly independent. This is one aspect of coding efficiency, which we may take as another general functional goal of cortex. Another notable feature of visual cortex is heterogeneity: individual cells differ widely in their receptive field size, shape, and chromatic properties. The functional goal here may be to provide multiple representations of the image suited to various tasks. A specific example is the existence of many sizes of otherwise similar receptive fields, which may permit scale-invariant analysis of the image. Multiple representations may also aid in the segregation of regions within the image by common attributes, such as spatial frequency, orientation, bandwidth, and color and motion (Barlow, 1983). Finally, it is often argued that the cortical code must provide elements that are convenient for subsequent processing. An example is Koenderink's discussion of geometrical measures computed by cortical receptive fields (Koenderink & van Doorn, 1987).

Mathematical constraints

Mathematics cannot dictate physiology, but it can ensure that our theory is sensible and consistent. Furthermore, there is a rich mathematical literature on subjects such as coding, signal processing, and image processing that are of great value in understanding the cortex. Several principles in particular are fundamental.

The first principle is that since cortical information is represented in cell activity, the representation of visual information must be discrete. Each cell holds a *sample* of the visual image, and the receptive field defines the nature of that sample. The sampling process operates over all the dimensions of the input, including spatial, temporal, and chromatic. Sampling theory describes which collections of samples suffice to capture the information in a signal (Shannon, 1949; Jerri, 1977; Dudgeon & Mersereau, 1984). Sampling theory must be the foundation of any theory of the representation of visual information in cell activity (Hughes, 1981; Crick, Marr & Poggio, 1981; Sakitt & Barlow, 1982; Yellott, 1982, 1983; Robson, 1983; Watson, 1983, 1987a,b; Williams, 1985; Geisler & Hamilton, 1986; Maloney & Wandell, 1986; Ahumada & Poirson, 1987). Since spatial imagery is two-dimensional, both the spatial receptive fields and sampling rules must be specified in two dimensions.

In the preceding section I noted the desirability of a complete representation. Sampling theory is one means of ensuring a complete representation. A closely related but somewhat different approach is exemplified in the theory of coding and reconstruction of images (Chellappa & Sawchuck, 1985). One simple principle from this field is that a complete representation must have a reconstruction algorithm, by which one computes the image from the code. Thus we would like to understand at least in principle, how to get back to the image from the cortical code.

Two different codes may be complete, yet differ in efficiency. Efficiency may be quantified in various ways, for example in terms of the number of coefficients (receptive fields), or in terms of the total entropy of the code (Watson, 1987b), or by yet other code statistics (Field, 1987). The efficiency of a code is closely related to the redundancy among coefficients discussed above. In any case, an understanding of the mathematical principles of coding efficiency is likely to enhance our understanding of the visual code.

I have already mentioned one intriguing aspect of visual cortex, namely the existence of multiple sizes and orientations of otherwise similar receptive fields. This feature is also found in image codes that have a *pyramid* structure, so-called because at each level the number of samples declines exponentially (Tanimoto & Pavlidis, 1975; Burt & Adelson, 1983; Woods & O'Neil, 1986; Watson, 1986, 1987a,b). These codes, also described as 'multi-resolution', or 'sub-band', employ kernels (receptive fields) that are self-similar (identical but for a change of scale or orientation). Since the brain appears to use pyramid coding, the principles that underlie these codes may prove helpful in understanding the cortical code.

Perceptual constraints

In the perceptual category, without meaning to disparage psychophysics, I have only one item: *perceptual completeness*. Earlier I said the code must be complete. But it must be complete only up to the perceptual capacity of the observer – it may lose whatever information the human loses. Psychophysics show us what information is retained and what is lost (Watson, Ahumada & Farrell, 1986). The goal is to construct a code that is as complete as the human or primate observer.

As an illustration, Fig. 35.1A is an image represented by 256^2 pixels, each of 24 bits. Elsewhere I have described a code modeled on visual cortex which

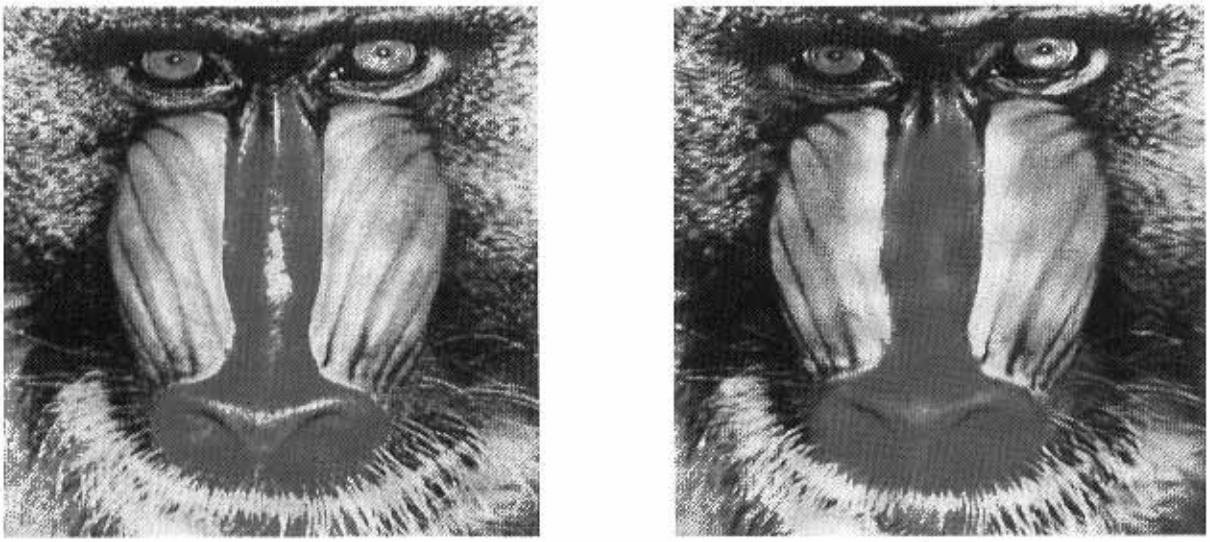


Fig. 35.1. A. An image coded at 24 bits/pixel. B. The same image, reconstructed from a code of 0.85 bits/pixel.

employs achromatic oriented frequency tuned cells and chromatic low resolution non-oriented cells (Watson, 1987a,b). The resulting code for this image uses less than 1 bit/pixel, and hence discards masses of information. But as the reconstruction in Fig. 35.1B illustrates, it is nevertheless perceptually almost complete.

One difficulty with applying perceptual constraints to a theory of visual cortex is that we rarely are able to locate with any precision the anatomical site at which perceptual judgements are made, or more precisely, at which visual information is lost. Nevertheless some general statements can be made. For example, if the physiological pathway under consideration is argued to be the site for high-resolution spatial vision, then it should manifest the general limitations of that capacity, in terms of the contrast sensitivity function and the like.

Physiological constraints

This large category primarily describes properties of the receptive fields of simple cortical cells. Space does not permit a detailed review, and I will highlight only what seem to me the most critical points. Excellent reviews are available elsewhere (Lennie, 1980; Shapley & Lennie, 1985).

Oriented, local, frequency-tuned receptive fields. The fundamental work of Hubel and Wiesel (reviewed in Hubel & Wiesel, 1977) and subsequent observations by others (Movshon *et al.*, 1978a,b; De Valois, Albrecht & Thorell, 1982; De Valois, Yund & Hepler, 1982; Schiller *et al.*, 1976; Foster *et al.*, 1985) has shown that most cortical simple receptive fields consist of a small number of parallel elongated regions of alternating excitation and inhibition. As a consequence, they are relatively compact in space, spatial frequency, and orientation: each cell responds over a limited region of space and to a limited band of spatial frequency and orientation. (The latter two restrictions may be described jointly as a compactness in two-dimensional spatial frequency.) To ensure completeness, there must be many cells, tuned to different spatial frequencies, orientations, and locations. Together the collection of cells must cover the visible regions of space, spatial frequency, and orientation. This distribution of cell properties has in fact been observed (De Valois, Albrecht & Thorell, 1982; De Valois, Yund & Hepler, 1982).

Frequency and orientation bandwidths. Within a simple receptive field, the extent in space (size) and in frequency (bandwidth) are inversely related. The median frequency bandwidth is about 1.5 octaves (De Valois, Albrecht & Thorell, 1982; Foster *et al.*, 1985; Hawken & Parker, 1987). The median orientation bandwidth estimate is about 40 degrees (De Valois, Yund & Hepler, 1982).

Heterogeneity. An important property of the population of cortical receptive fields is heterogeneity. This is perhaps best illustrated by the distributions of frequency and orientation bandwidth. The medians are cited above, but the distributions are broad and relatively shallow. A comprehensive theory must account for these ranges.

Bandwidth vs. peak frequency. In spite of the broad range of frequency bandwidths, we can nevertheless consider the covariation of peak frequency and bandwidth. Two simple cases are possible. If receptive fields were of constant size in degrees, then octave bandwidth would be inversely related to peak frequency. If receptive fields were a constant number of cycles of the underlying modulation (constant number of bars), then octave bandwidth would be constant, and independent of peak frequency. Available evidence indicates that neither result is strictly true. Median bandwidth declines slightly with peak frequency, but is much nearer to the constant octave hypothesis (De Valois, Albrecht & Thorell, 1982).

Phase. There is some evidence that cortical receptive fields exist in quadrature pairs, that is, that their spatial phases are 90° apart (Pollen & Ronner, 1981). Examples might be receptive fields with even and odd symmetry (0° and 90° phases, or so-called bar-detectors and edge-detectors), but current evidence suggests that such canonical phases are not the rule (Field & Tolhurst, 1986).

Aspect ratio. Another quantitative measure of receptive field shape is the aspect ratio, which describes the ratio of receptive field length (along the axis of orientation) to width. In linear cells, this is inversely related to the ratio of frequency and orientation bandwidth. For a cell whose two-dimensional frequency support is circular, the aspect ratio is 1. Estimates vary from cell to cell, with no consensus on a mean value different from 1 (De Valois, Albrecht & Thorell, 1982; Daugman, 1985; Webster & De Valois, 1985).

Linearity. Typically about half of the cells recorded in V1 in primate can be classified as 'simple' (Hubel & Wiesel, 1968; De Valois, Albrecht & Thorell, 1982; Hawken & Parker, 1987). These cells may exhibit output non-linearities, but show approximately linear spatial summation, and thus presumably receive input from linear X-like cells of the geniculate and retina. Thus the convergence of retinal and geniculate cells on the simple cortical cell may be modelled by purely linear operations.

Color. While still a matter of some debate, it appears that the majority of simple oriented V1 cells in primate are achromatic (Lennie, personal communication). In spite of the fact that they receive input from color opponent geniculate cells, they show little evidence of chromatic opponency.

Sensitivity and noise. A number of schemes exist for quantifying the sensitivity of visual cells (Enroth-Cugell *et al.*, 1983; Kaplan & Shapley, 1986), some of which take into account the noise in the cell's response (Derrington & Lennie, 1984; Hawken & Parker, 1984). We define sensitivity as the inverse of the signal contrast which produces a response greater than one standard deviation of the noise on 50% of the trials. This corresponds directly to a signal/noise or d' measure (Green & Swets, 1966).

Kaplan & Shapley (1982) observed that peak sensitivities of parvocellular LGN cells were in general much lower than magnocellular cells. Derrington & Lennie (1984) indicate peak parvocellular sensitivities of around 20. Hawken & Parker (1984) have measured sensitivity of cells in layer 4 of primate cortex. They found a wide range of values (from about 1 to about 200). Within sublaminae receiving direct parvocellular input (IVa and IVcβ) sensitivities were generally lower (between about 1 and 60). As I shall show, specific schemes for the construction of cortical receptive fields from geniculate inputs must accommodate these estimates of sensitivity.

Anatomical constraints

Recent years have seen a profusion of results on the anatomical structure of the striate cortex (reviewed in Hubel & Wiesel, 1977; Gilbert, 1983; Livingstone & Hubel, 1984) and of its tributaries in the retina and LGN (Sterling, 1983; Hendrickson, Wilson & Ogren, 1978; Perry, Oehler & Cowey, 1984). Here again I can only highlight the small fraction of this literature that seems fundamental or particularly germane.

Retinal and geniculate inputs. Primate striate cortex receives all of its visual input from the LGN, which in turn is driven by retinal ganglion cells. Thus cortical cell receptive fields must be constructed from LGN receptive fields, which must in turn be constructed from ganglion cell receptive fields.

The most numerous of the various ganglion cell types in the primate retina are the X-type, which show high spatial resolution and linear responses (De Monasterio, 1978; Kaplan & Shapley, 1986). X receptive fields consist of antagonistic center and surround regions, and can be modelled as the difference of two

radially-symmetric Gaussians. Receptive field size increases with eccentricity, and only one size appears to exist at a given eccentricity (Perry & Cowey, 1981, 1985).

The primate LGN is divided into six lamina. The two dorsal layers contain magnocellular cells, and the four ventral layers contain parvocellular cells (Hubel & Wiesel, 1972). We confine our attention to the parvocellular cells, which are highly linear, have high spatial resolution, and relatively low peak sensitivity (Blakemore & Vital-Durand, 1981; Kaplan & Shapley, 1982; Derrington & Lennie, 1984). Their receptive fields also have a center-surround receptive field, may also be modelled as a difference-of-Gaussian (DOG), and are in general difficult to distinguish from those of X-type retinal cells (So & Shapley, 1981; Derrington & Lennie, 1984). Parvocellular geniculates appear to be driven by retinal X cells with highly similar receptive fields (Kaplan & Shapley, 1986).

Most of the cells in the parvocellular layers of the LGN are chromatically and spatially opponent, most frequently with a red center and green surround, or green center and red surround (Derrington, Krauskopf & Lennie, 1984).

Parvocellular geniculate afferents terminate in layers IVa and IVcβ of the primate cortex (Hubel & Wiesel, 1972; Hendrickson *et al.*, 1978; Blasdel & Lund, 1983). However, in view of intracortical connections, we cannot say with certainty which layers, if any, receive exclusively parvocellular input (Gilbert, 1983).

Hexagonal input lattice. Since the parvocellular LGN appears to provide a transparent conduit between retina and cortex, we can consider the 'parvocellular' retinal cells (Kaplan & Shapley, 1986) as providing input to the cortex. In this light it is interesting to observe that the foveal receptor lattice is nearly hexagonal (Hirsch & Miller, 1987). Furthermore, available evidence suggests that in the foveal region, each ganglion cell center receives input from a single cone (Boycott & Dowling, 1969). This means that the ganglion cell RF centers also form a hexagonal lattice, and if the geniculate cells are likewise matched one-to-one with ganglion cells, then they too have a hexagonal sample lattice. The cells of the cortex must be constructed from this hexagonal array of retinal ganglion cell samples.

Cortical magnification. While the mapping between retina and cortex is topographic, the amount of cortex devoted to each retinal area declines with eccentricity. This relationship may be quantified by the cortical magnification factor, describing the linear mm of cortex devoted to a degree of visual field (Daniel & Whitteridge, 1961). A recent estimate in macaque is about 15–20 mm/deg for the fovea, with peripheral values about 10 times the inverse of eccentricity (Van Essen *et al.*, 1984). The magnification of the fovea relative to periphery is much greater in cortex than in either retina or LGN, so the spatial transformations among these layers cannot be spatially homogeneous (Perry & Cowey, 1985).

Hypercolumns. Area V1 is organized into hypercolumns of approximately 4 mm^2 (Hubel & Wiesel, 1977). Each hypercolumn is believed to perform a complete analysis of a local region of the visual field. Cells with the same ocular dominance tend to lie in parallel slabs orthogonal to the cortical surface. Cells with a common orientation preference also form slabs, or perhaps whirls, which intersect the ocular dominance columns. A hypercolumn consists of a portion of cortex with both ocular dominance columns and a complete set of orientation columns. At the center of each ocular dominance column 'blobs' of non-oriented cells have been found in layers 2 and 3, many of which are color selective (Livingstone & Hubel, 1984). Electrode penetrations tangential to the cortical surface often reveal a progressive change in receptive field orientation with penetration depth, covering 360° of orientation in approximately 1 mm of cortex (Hubel & Wiesel, 1977; Livingstone & Hubel, 1984).

Laminar structure. The striate cortex consists of a number of anatomically distinct layers. There is as yet no clear picture of the functional aspects of cells in each layer, but evidence on this matter is accumulating rapidly (Gilbert, 1983; Livingstone & Hubel, 1984; Hawken & Parker, 1984). Models of striate cortex should acknowledge the laminar structure and functional differences among lamina, and ultimately provide a functional explanation for the lamination.

Oversampling. While we have earlier pointed to efficiency as a plausible goal of the cortical code, and have noted that this may translate into a minimum number of cells, we must acknowledge that the striate cortex effects an enormous multiplication in the number of cells, relative to the number of retinal ganglion cells. This multiplication factor may be as high as 1000 (Barlow, 1983). Since the number of ganglion cells absolutely limits the degrees of freedom of the incoming signal, the cortex clearly does not seek to minimize its total number of cells. However, we have already noted the heterogeneity of visual cortex, and it is

plausible that this multiplicity of cells reflects a multiplicity of codes, rather than an inefficiency in any one. The laminar structure itself suggests a multiplicity of cell types and perhaps codes.

Retinal disorder. Earlier I noted that the retina may impose its hexagonal structure on the cortex, and in this light it is worth noting that the receptor lattice becomes progressively more disordered with eccentricity (Hirsch & Miller, 1987; Ahumada & Poirson, 1987). This disorder may give rise to some of the heterogeneity at the cortical level, but must at any rate be taken into account in a description of the mapping from retina to cortex.

Receptive field models

There have been attempts to describe detailed receptive field shape with mathematical models. Marr & Hildreth have proposed $\nabla^2 G$, the Laplacian of a Gaussian (Marr & Hildreth, 1980). This is inadequate because it lacks orientation tuning. Young (1985) and Koenderink & van Doorn (1987) have proposed $D_N G$, the successive derivatives of a Gaussian. This is problematic because it predicts a constant linear bandwidth, and because it has no plausible extension into two dimensions. The Gabor function, the product of a Gaussian and a sinusoid, fits simple cell receptive fields well in many respects (Marcelja, 1980; Marcelja & Bishop, 1982; Webster & De Valois, 1985; Daugman, 1985; Field & Tolhurst, 1986). But detailed comparisons with empirical tuning functions show that the Gabor has too much sensitivity at low frequencies (Hawken & Parker, 1987). Put another way, the Gabor is highly asymmetrical in log–log coordinates, whereas actual tuning profiles are more nearly symmetrical. As an alternative Hawken & Parker have proposed the difference-of-separated-DOGs (d-DOG-s) model, describing the one-dimensional receptive field profile as the sum of 3 DOGs, a large positive central one flanked by two smaller negative ones. The fits of this model are generally much better than a Gabor, and are quite good overall. The objections to this model are, first, that it requires nine parameters to get this good fit. Second, it is not specified in the second spatial dimension, which would require still more parameters.

The chexagon model

I turn now to a description of a new model, which attempts to deal with many of the constraints I have described. While many of these constraints guided the design of this model, two were particularly influential. The first was the simple fact that cortical cells must be constructed from retinal inputs with DOG receptive fields. The second was that foveal receptors form an approximately hexagonal lattice. This means that X ganglion cells, and parvocellular cells of the LGN, must likewise have receptive fields forming a hexagonal lattice.

While the direct input to cortex is from the LGN, this input is in turn derived from retinal ganglion cells. Furthermore, there is as yet little evidence regarding how geniculate cells alter the retinal signals. I will therefore speak of the retinal ganglion cells as 'inputs' to the cortex. Accordingly, we begin with a retinal parvocellular receptive field, which we model as a DOG. The center Gaussian has a radius w, and the surround Gaussian a radius qw. The center Gaussian has volume $\varkappa$, and the surround Gaussian $s\varkappa$. Note that when s is 0, the frequency response is low-pass, when it is 1, the response is band-pass with no DC response.

The receptive fields are arranged in a hexagonal lattice with spacing λ. The DOG center radius can be expressed as a proportion of λ,

$$w = r\lambda \tag{1}$$

A value of $r=1$, illustrated in Fig. 35.2, provides adequate sampling, and is in general agreement with what is known about the coverage factor of primate X parvocellular geniculate cells (the ratio of cell density to receptive field size) (Perry & Cowey, 1985).

In the remainder of this paper I will consider cortical receptive fields constructed from linear com-

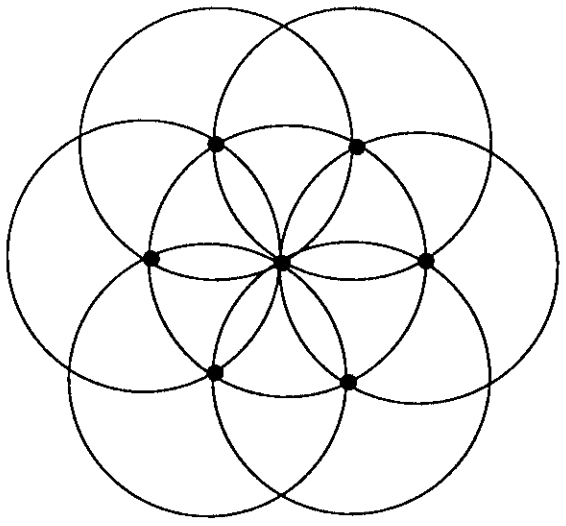


Fig. 35.2. Portion of a hexagonal lattice of ganglion cells. The ganglion cells are modelled by a difference-of-Gaussians (DOG). Circles represent center Gaussians at half height. In this example, the center Gaussian radius is equal to center spacing.

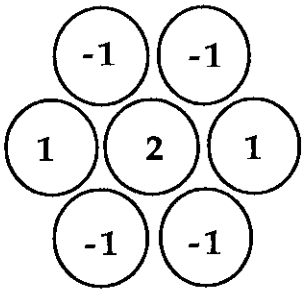


Fig. 35.3. Simple chexagon kernel. The numbers are weights assigned to each ganglion cell in the lattice of Fig. 35.2.

binations of these hexagonally spaced DOG inputs. Each receptive field is defined by the weights assigned to each point in the lattice. I will call this set of weights, and its topographic arrangement, a *kernel*. I will describe two schemes in which the kernels employ only a point and its six neighbors in the lattice, which I will call *chexagons* (cortical hexagon). The name is not to be taken too seriously, but provides a useful shorthand. In the discussion I will consider generalizations to other kernels that may employ larger numbers of weights.

Simple chexagon

Kernel. Our first scheme consists of kernels using a point and its six neighbors. We select weights such that the resulting receptive fields will have no response to uniform illumination (0 DC), will be oriented, and will be band-pass in spatial frequency. A particularly simple arrangement is shown in Fig. 35.3. In other words, we combine the signals from these seven DOG receptive fields, weighting each by the value shown. Two more receptive fields can be constructed by rotating the kernel by angles of 60° and 120°, providing a set of three oriented receptive fields, as shown in Fig. 35.4.

Figure 35.5 shows the 1D spectrum of one receptive field, measured at optimal orientation. Also shown are the spectrum of the underlying DOG, and the spectrum of a Gabor function of the same peak frequency and bandwidth. The 1D bandwidth of the chexagon is about 1.6 octaves, near to the median estimate for primate cortical cells. The chexagon spectrum is considerably narrower than the DOG spectrum, and in effect selects a slice, or sub-band of the DOG spectrum. The chexagon bandshape is much more nearly symmetrical (on log–log coordinates) than the Gabor, and is therefore likely to provide a much better fit to measured 1D spectra of cortical cells (Hawken & Parker, 1987).

Pooling. I have constructed a receptive field tuned to a particular frequency band, and note that it is at the upper end of the pass-band of the DOG spectrum, and thus of the pass-band of the system as a whole. It is, in other words, a high-frequency cell. I will call these the *level* 0 cells. Now we ask: where do we get the lower-frequency cells? We cannot simply increase the separation between geniculate inputs λ, as that will produce severe undersampling and aliasing. Likewise we cannot vary w, since retinal ganglion cell centers do not vary in size at one eccentricity.

The solution is to create *pooling units*, through linear combination of the original retinal inputs. There are various pooling kernels that might be used. For simplicity, we consider a Gaussian function of distance from the central sample point. The pooling units then have individual receptive fields that are the convolution of a DOG and the Gaussian kernel. These will resemble the original DOGs but will be somewhat larger, depending on the size of the Gaussian. These pooling units now provide the input to the next set of kernels, which are identical to those at the previous level except that they are enlarged by a factor of two. In other words, the distance between weights in the kernel is now 2λ. These new kernels generate the receptive fields of cells at level 1. The same procedure may be repeated, successive applications yielding larger and larger pooling units, and lower and lower frequency cells.

Receptive fields and spectra. Figure 35.6 illustrates the receptive fields, and their frequency spectra, at four levels. The figure also shows the spectra of the components used to construct the receptive field: the DOG, the hexagonal kernel, and the pooling kernel. The receptive field spectrum is the product of these three spectra. The receptive fields have a shape that roughly resembles cortical cells. The spectra are smooth, band-pass, and oriented, and generally consistent with what is known about 2D spectra of single cortical cells (De Valois, Albrecht & Thorell, 1982; De Valois, Yund & Hepler, 1982).

Sensitivity. Earlier we noted that the sensitivity of a cell may be expressed in signal/noise terms. Given a signal/noise description of retinal ganglion cell sensitivity, the sensitivity of the computed cortical receptive field can be expressed in the same terms. Specifically, assume that each ganglion cell has identical independent Gaussian noise, and that no additional noise is introduced at geniculate or cortical levels (this assumption can be relaxed in a more elaborate formulation). Then the noise at the cortical

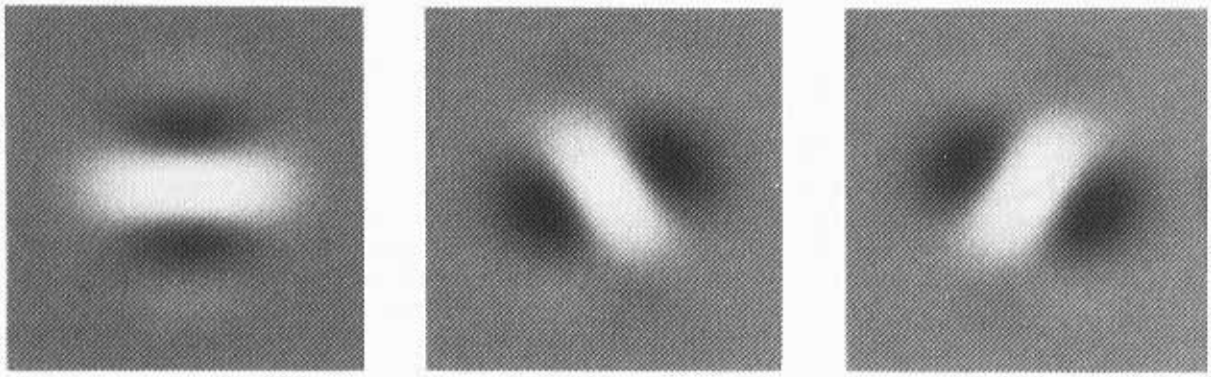


Fig. 35.4. Receptive fields of the simple chexagon model. The three orientations are produced by rotating the kernel of weights by 0°, 60°, and 120°. The dimensions of each image are 8λ. Parameters are $r = 1$, $q = 6$, $s = 0.5$.

cell will have a variance equal to the sum of the squared kernel weights times the retinal variance. To express sensitivities in signal/noise terms, the kernel must be normalized, so that the sum of the squared kernel weights is 1 (the kernel has a unit norm). For the higher order cells, this unit norm applies to the convolution of the pooling function and the chexagon kernel.

Figure 35.7 shows the 1D spectra at four levels, as well as the spectrum of the retinal DOG. The curves are correctly placed on the vertical axis. Note that the cortical sensitivity can be substantially above that of the ganglion cells, because of spatial pooling. At each level the spectra are similar in shape, but displaced downwards on the frequency axis by approximate factors of two. The spectra have approximately constant log bandwidth, consistent with the data cited earlier.

Subsampling. The centers of the kernels of the level 0 cells are coincident with the centers of the retinal inputs. Since there are three orientations, there are therefore three times as many of these level 0 cells as there are retinal inputs. The size of the pooling Gaussian may be chosen so as to permit subsampling at each level without substantial aliasing. By subsampling we mean the construction of receptive fields at a subset of the locations in the input sample lattice. As an example, we consider subsampling by a factor of two in each dimension at each level. The distance between receptive field centers at level n is then $\lambda 2^n$. This results in a total number of cortical cells equal to four times the number of retinal inputs. The various levels of cells together form a pyramid structure as discussed above. Figure 35.8 illustrates several levels of this pyramid.

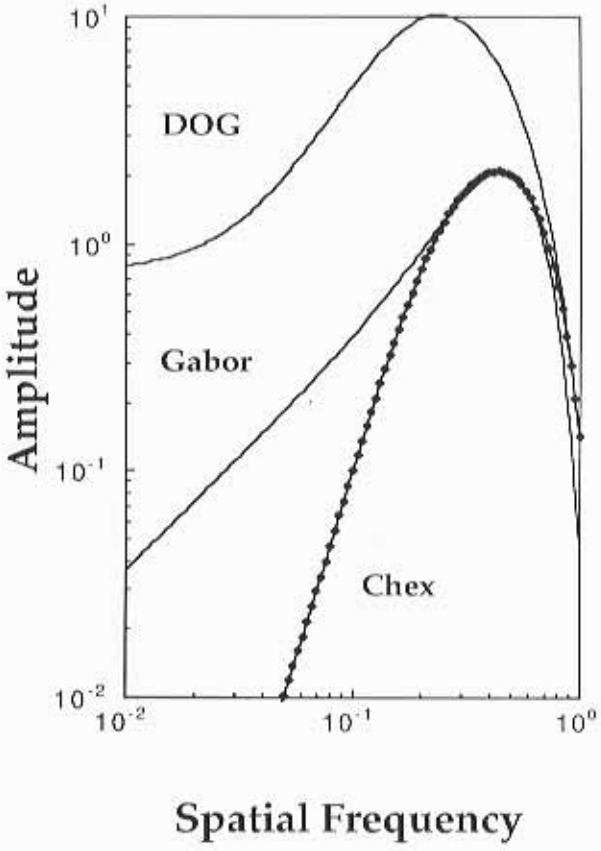


Fig. 35.5. 1D spectrum of the chexagon receptive field at optimal orientation. Also shown are the spectrum of the retinal DOG receptive field and of a Gabor function of similar frequency and log bandwidth. Parameters are $\alpha = 15$, $\lambda = 0.5$, $r = 1.15$, $q = 3$, $s = 0.95$.

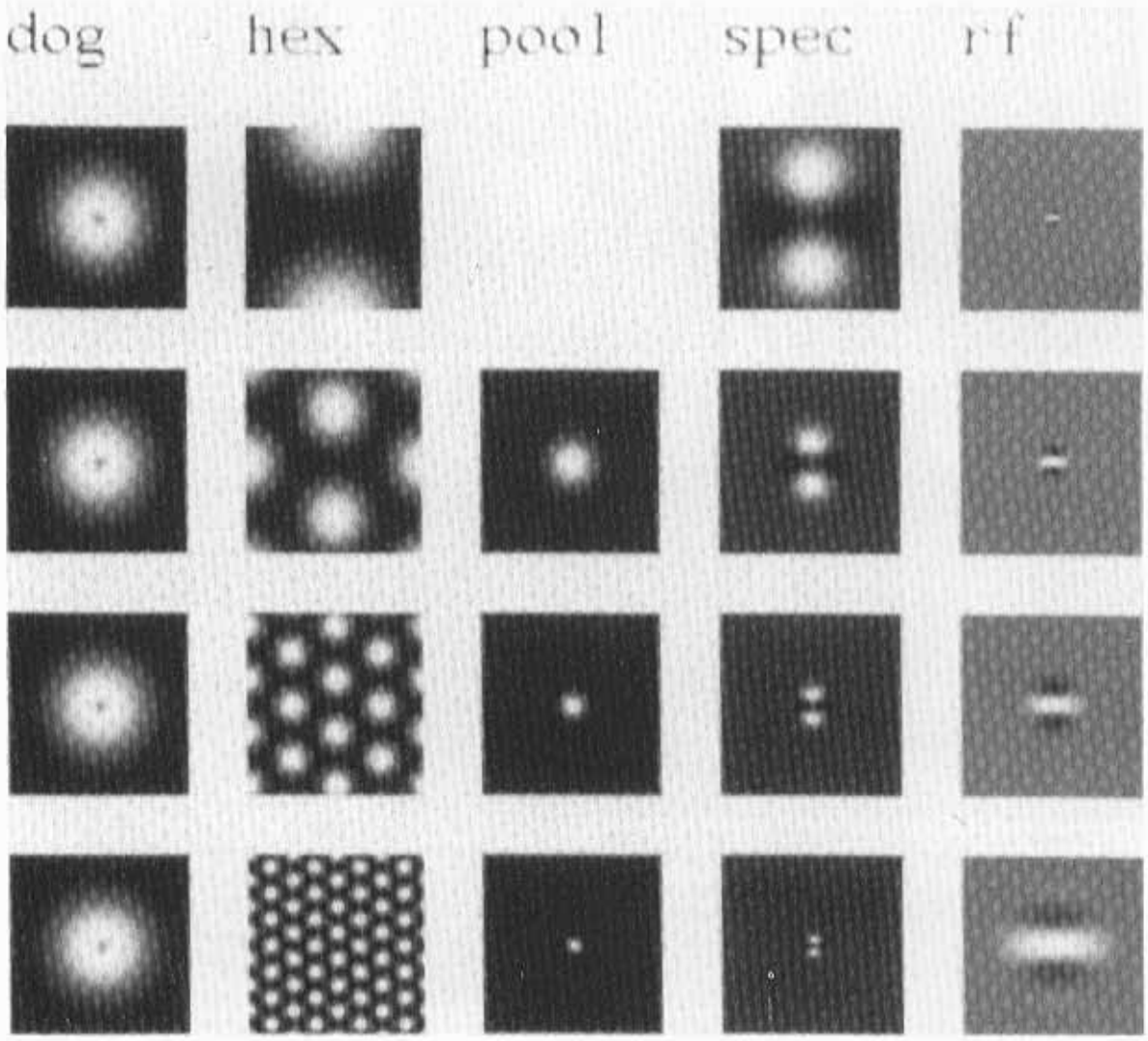


Fig. 35.6. Receptive fields and frequency transforms for four levels of the simple chexagon pyramid. Only the even receptive field at one orientation is shown. From left to right, the columns show the spectrum of the retinal inputs (DOG), the spectrum of kernel of weights, the spectrum of the pooling kernel, the spectrum of the cortical receptive field, and the receptive field. The fourth column is equal to the product of the first three columns, and the last column is the inverse Fourier transform of the fourth column. Rows from top to bottom show levels from 0 to 3. There is no pooling kernel at level 0. Parameters are $r = 1$, $q = 6$, $s = 0.5$. Dimensions of each image are 64λ.

Hyperchexagons. The process of generating larger and larger cells must stop somewhere. The number of levels corresponds to what have elsewhere been called channels, which have been estimated psychophysically to number between three and eight (Wilson & Bergen, 1979; Watson, 1982). At the top of the pyramid, the receptive fields will have a particular size. It is useful to be able to talk about all the receptive fields in the area occupied by this largest receptive field, so I will call them collectively a *hyperchexagon*. Each hyperchexagon is itself a pyramid which analyzes a sub-region of the visual field.

Efficiency. Although we do not demonstrate it here, it is worth noting that the simple chexagon pyramid provides a complete representation of the image pro-

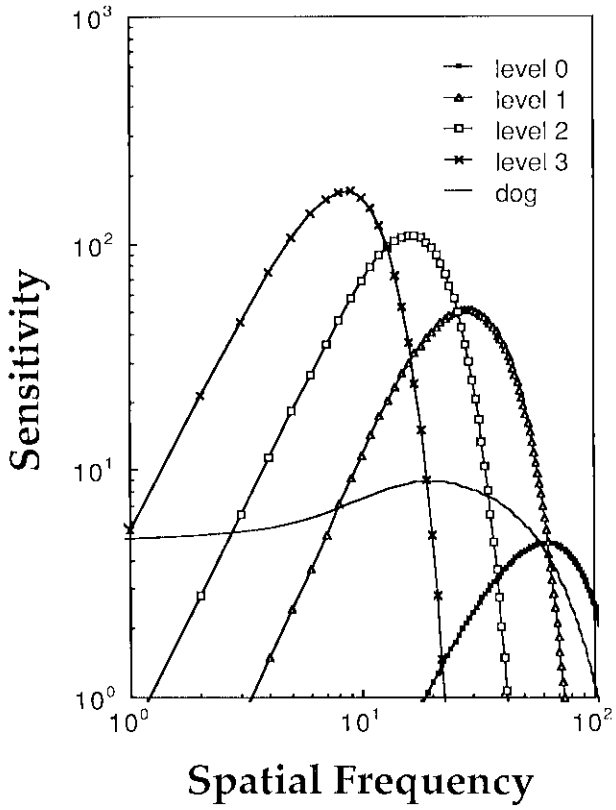


Fig. 35.7. 1D frequency spectra for four levels of the simple chexagon receptive field. The spectrum of the retinal input DOG is also shown. Sensitivity is expressed in signal/noise terms. Parameters are $\kappa = 10$, $r = 1$, $q = 6$, $s = 0.5$. Frequency is in units of $1/(256\lambda)$.

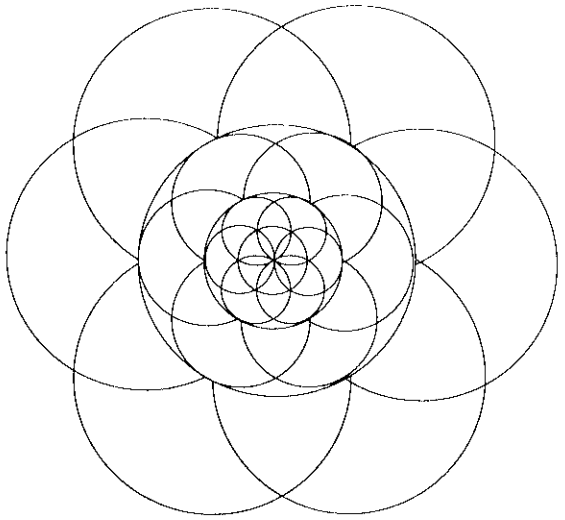


Fig. 35.8. Three levels of the simple chexagon pyramid. The smallest circles represent receptive field centers of retinal ganglion cells. The larger circles represent pooling units. Only one chexagon is shown at each level.

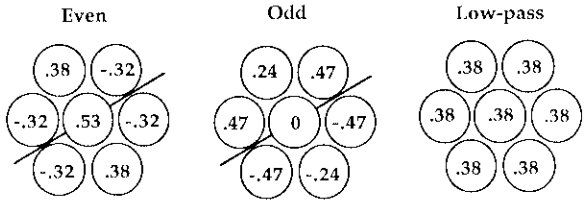


Fig. 35.9. Orthogonal chexagon kernels. The oblique lines indicate one axis of symmetry for the high-pass kernels.

vided by the input ganglion cells. In fact, as might be suspected from the multiplication by four of the total number of cells, the representation is redundant. Elsewhere we have noted that the cortex as a whole greatly expands the number of cells, relative to the retinal input. This over-representation may in part serve to immunize the system against noise and loss of cells, and so we should not necessarily expect a cortical representation that is highly efficient. Nevertheless, it is worth considering the form such an efficient code might take, if only to understand the redundancy imposed by less efficient codes. This leads us to the second form of chexagon pyramid.

Orthogonal chexagons

Kernels. In this case we have sought a set of kernels that are band-pass and oriented as above, but that are also orthogonal, and that form odd–even (quadrature) pairs. One consequence of orthogonality is that the number of cells constructed is equal to the number of input cells, hence the code is maximally efficient in this sense. The orthogonal code is complete, and thus permits exact reconstruction of the input samples. In addition, reconstruction of the image employs the same kernels used in coding. The derivation of the kernels is described elsewhere (Watson & Ahumada, 1987). The resulting kernel values are shown in Fig. 35.9.

There are seven kernels, of which only three are shown in Fig. 35.9. They are even, odd, and low-pass. The four not shown are produced by rotations of even and odd kernels by angles of 60° and 120°. The receptive fields are shown in Fig. 35.10.

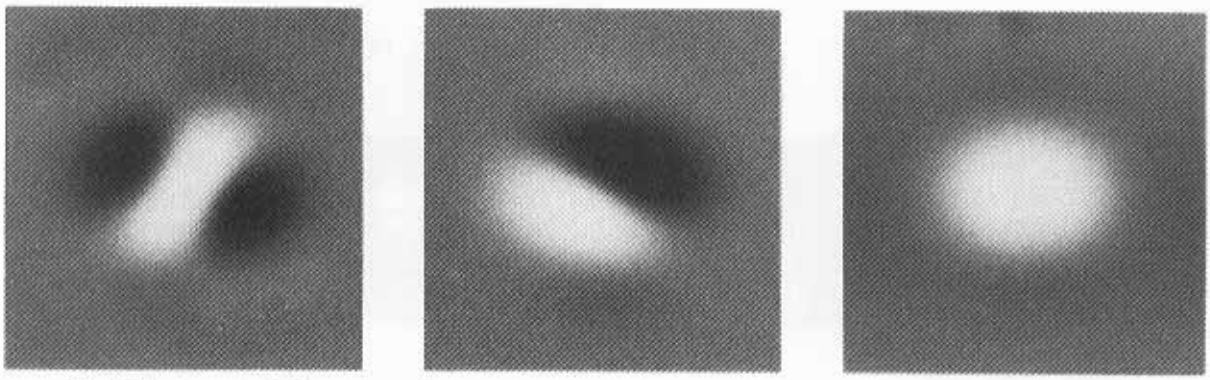
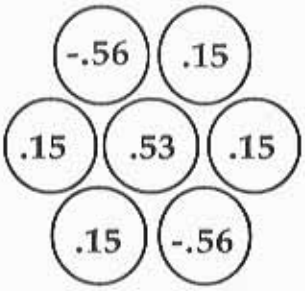


Fig. 35.10. Receptive fields generated by orthogonal chexagon kernels. From left to right they are: even, odd, and low-pass. Dimensions of each image are 8λ.



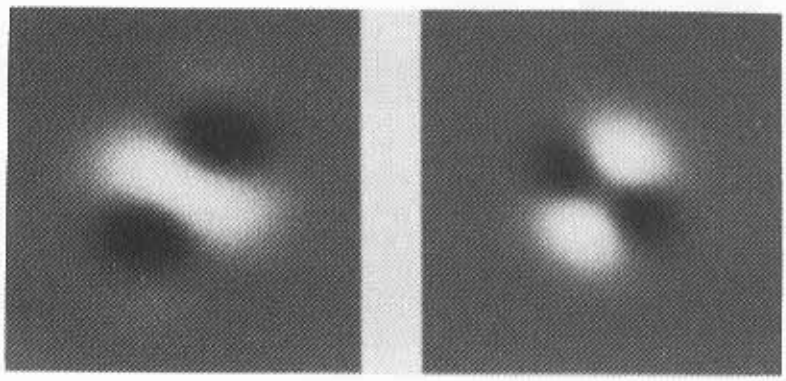


Fig. 35.11. Alternate form for the even kernel, with corresponding receptive field and frequency transform.

Alternate even kernel. We define the *orientation* of a receptive field as the orientation of the peak of the frequency spectrum, that is, the orientation of a sinusoidal input at which the kernel gives the largest response. An interesting feature of the resulting kernels is that while the axis of symmetry was fixed at 30°, the orientation of the even kernel is actually orthogonal to this axis at 120°. This places its orientation axis on the hexagonal lattice. The constraints under which the kernels were derived (essentially orthonormality, completeness, and quadrature pairs) generate another alternate solution for the even kernel (Fig. 35.11).

Calling the first solution type 0 and the second type 1, the latter has an orientation of 30°, equal to that of the odd receptive field. Thus if it is desired to have quadrature pairs with equal orientation, the type 1 even kernel must be used. But as can be seen in Fig. 35.11, the type 1 kernel suffers relative to the type 0 in having substantially greater sensitivity at the non-preferred orientation.

Pooling. In this scheme, the pooling is done by the low-pass kernel. The next level of the pyramid is constructed from these pooling units. Because the low-pass kernel is orthogonal to all six high-pass kernels, the high-pass kernels at level 1 will be orthogonal to the high-pass kernels at level 0 (since they are linear combinations of the low-pass kernel). By extension of this argument, it is clear that all kernels in the pyramid are orthogonal.

Receptive fields and spectra. The kernels at each level are self-similar. The receptive field will be only approximately self-similar, because while the kernels scale in size at each level, the retinal DOG function with which they are convolved remains fixed in size. Figure 35.12 illustrates an even receptive field and its spectrum at four scales. The spectra are approximately band-pass in frequency and orientation, with bandwidths similar to those encountered physiologically. For this even filter, there is substantial sensitivity at 90° to the preferred orientation. Such secondary peaks in

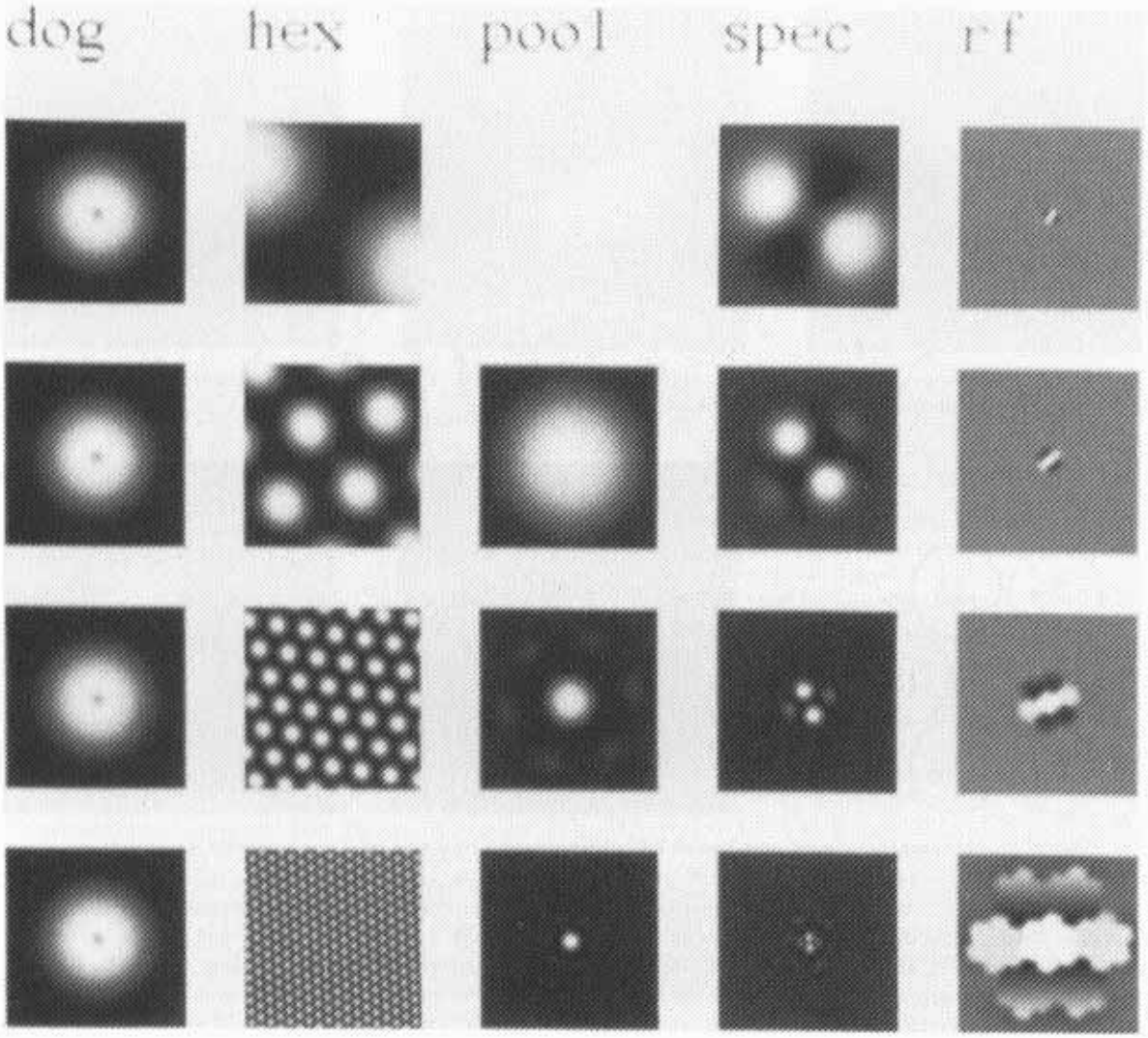


Fig. 35.12. Receptive fields and frequency transforms for four levels of the orthogonal chexagon pyramid. Only the even type 0 receptive field at one orientation is shown. From left to right, the columns show the spectrum of the retinal inputs (DOG), the spectrum of the kernel of weights, the spectrum of the pooling kernel, the spectrum of the cortical receptive field, and the receptive field. The fourth column is equal to the product of the first three columns, and the last column is the inverse Fourier transform of the fourth column. Rows from top to bottom show levels from 0 to 3. There is no pooling kernel at level 0. Parameters are $r = 1$, $q = 6$, $s = 0.5$. Dimensions of each image are 64λ.

orientation tuning are quite often found in cortical cells, but may not be a general feature (De Valois, Yund & Hepler, 1982). A closer look at a one-dimensional cross-section through each filter at the optimal orientation (Fig. 35.13) shows considerable ringing in the higher order filters.

Since the filters are orthogonal, this creates no problems for coding or reconstruction, but does render this exact form of the filters less like cortical cells. However, only the ringing in the highest level cell would be detectable in physiological measurements, due to their limited precision. This ringing can be reduced to some extent through judicious choice of filter parameters, but is an inherent property of these filters attributable to the small area and abrupt borders of the pooling region. It is not evident in the previous

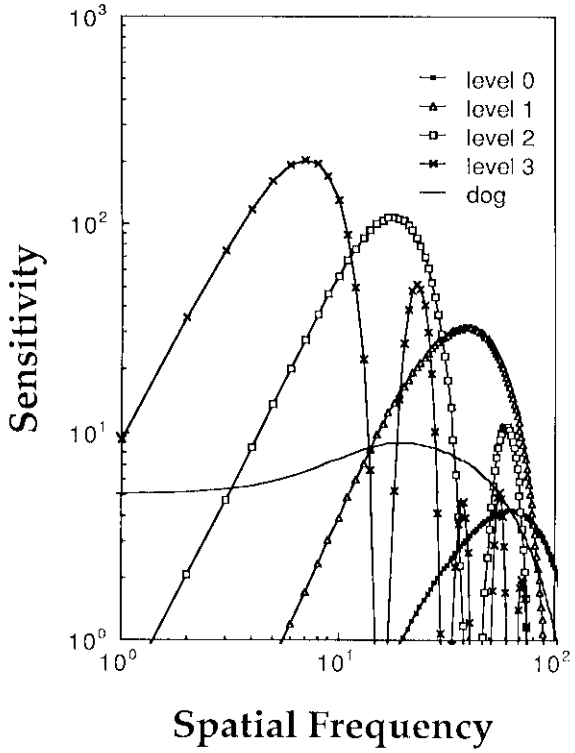


Fig. 35.13. 1D frequency spectra for four levels of the orthogonal chexagon receptive field. The spectrum of the retinal input DOG is also shown. Sensitivity is expressed in signal/noise terms. Parameters are $\kappa = 10$, $r = 1$, $q = 6$, $s = 0.5$. Frequency is in units of $1/(256\lambda)$.

scheme (Fig. 35.5) because there pooling employs a Gaussian extending over a larger area.

Subsampling. Since the information in each hexagonal neighborhood of seven retinal inputs is completely captured in the seven cortical receptive fields constructed therefrom, there is no need for overlap of adjacent kernels. We therefore tesselate the retinal input lattice with hexagonal kernels, as illustrated in Fig. 35.14. Because each kernel consumes exactly seven inputs, the number (or density) of receptive fields of a given type (e.g. odd at 60°) will be one seventh the number (or density) of retinal inputs. The kernel centers form a new hexagonal lattice with a sample distance of $\sqrt{7}\,\lambda$. The lattice is rotated by an angle of $\tan^{-1}(\sqrt{3}/5) \approx 19.1°$ relative to the retinal sample lattice. The kernels of level 1 are constructed from this new lattice, using the low-pass cells of level 0. Likewise at each successive level new kernels are constructed from the low-pass units at the previous

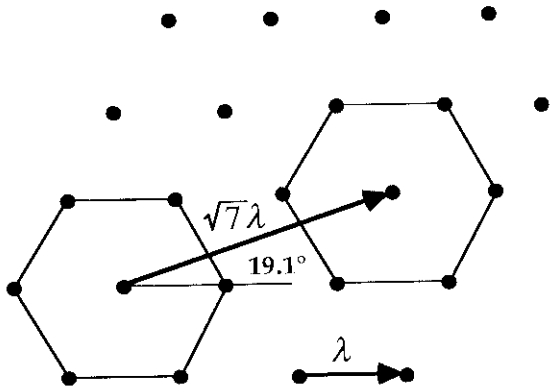


Fig. 35.14. Construction of receptive fields at level 0. Each point is the center of a retinal ganglion cell receptive field. Each hexagon shows a set of inputs that are combined to form the level 0 receptive fields.

stage, on a hexagonal lattice whose sample distance increases by $\sqrt{7}$ at each level, and which rotates by $\pm 19.1°$ at each level. After n levels, there are n sizes (spatial frequencies) of receptive field, and the total number of receptive fields is equal to the number of retinal inputs. The complete pyramid is illustrated in Fig. 35.15.

Discussion

In this section I will make some observations that apply to all of the schemes described above.

Hypercolumns. In the first scheme, the radius of the kernel increases by a factor of two at each level, so that at level n it is $\lambda 2^n$. This defines the size of the hyperchexagon. For example, if $n=4$, (5 levels), and $\lambda = 0.01°$ (Hirsch & Miller, 1987), then a single hyperchexagon has a radius of about 16 receptors or 0.16°. Foveal cortical magnification has been estimated in macaque at between 5 and 13 mm/°, so that a single hypercolumn of radius 1 mm corresponds to about 0.11°. While the numbers used in this calculation are not known with precision, and are further confused by the change in scale with eccentricity, they nevertheless suggest an identity between the hyperchexagon and the hypercolumn. This in turn suggests a speculation concerning the topography of cells within a hypercolumn.

Hypercolumn Architecture. In the pyramid schemes the number of cells per level is in proportion to the pass-band of that level, many for high-frequency levels, few for low-frequency levels. Adding to this the observation that the cortex is relatively homogeneous

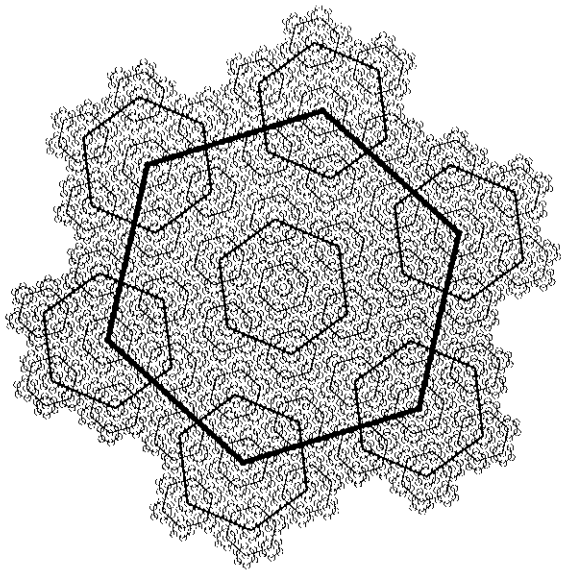


Fig. 35.15. Orthogonal chexagon pyramid. The level 0 kernels are represented by the smallest hexagons. The vertices of these smallest hexagons correspond to centers of the retinal ganglion cell receptive fields (and hence to receptors). Kernels at successive levels correspond to larger and larger hexagons. The vertices of each hexagon at each level coincide with the centers of the hexagons at the previous level. At each level, the kernels are applied to the low-pass responses from the previous level. This hexagonal fractal was constructed by first creating the largest hexagon, then placing at each of its vertices a hexagon rotated by $\tan^{-1}(\sqrt{3}/5) \approx 19.1°$ and scaled by $1/\sqrt{7}$. The same procedure is then applied to each of the smaller hexagons, down to a terminating level of 5. The sample lattice (defined by the vertices and centers of the smallest hexagons) is then a finite-extent periodic sequence on a hexagonal sample lattice. The sample lattice has 7^6 points, the same as a rectangular lattice of 343^2. The perimeter is a 'Koch curve' with a fractal dimension of $\log 3/\log \sqrt{7} \approx 1.19$ (Mandelbrot, 1983, p. 46). The program used to create this image is given in Watson & Ahumada (1987).

in cell density, we confront the problem of how to arrange the cells within each hypercolumn. A simple solution is to arrange them as a replica of the spectrum itself, as shown in Fig. 35.16. This would lead naturally to the orientation columns found anatomically, as well as to the progressive changes in orientation found along certain tangential electrode tracks. This is speculative, of course, but amenable to test.

Pooling units. The pooling process described in the above schemes is essential to generate narrowband

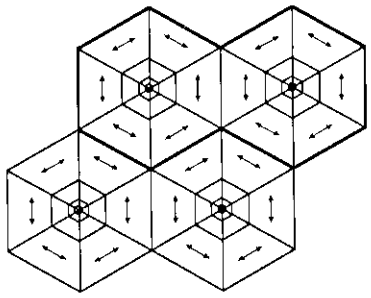


Fig. 35.16. Schematic of a possible arrangement of frequency, orientation, and ocular dominance among cortical hypercolumns. Each large hexagon encloses one ocular dominance column. Each quadrilateral region encloses cells with the same orientation and spatial frequency. The orientation is indicated by the arrows. Spatial frequency is lowest at the center of each hexagon, and increases centrifugally. The heavy line encloses two ocular dominance columns forming one hypercolumn.

spectra for low-frequency cells. There are at least three natural possibilities for the site at which this pooling occurs. First, Derrington & Lennie (1984) have found a wide distribution of center sizes for foveolar parvocellular geniculate cells. While they suggest this may be due to measurement errors, it may reflect a pooling operation within the geniculate. A second and perhaps more likely possibility is that the pooling units may be the various non-oriented cortical cells such as those found in the cytochrome oxidase blobs at the center of each hypercolumn in layers 2 and 3 (Livingstone & Hubel, 1984), or the non-oriented cells found in layers 4cβ (Hawken & Parker, 1984). A third possibility is that the pooling may be at the level of the oriented cell itself, although, for a low-frequency cell, this calls for the convergence of very large numbers of geniculate cells upon a single cortical cell.

Multiplexing of luminance and color information. To this point I have spoken of retinal and geniculate inputs as being of two types: on- and off-center. But Derrington & Lennie have shown that the parvocellular geniculate cells are chromatically, as well as spatially opponent. They are of two types. By far the more numerous are R–G cells, in which center and surround receive opposed contributions from R and G cones (red on-center, red off-center, green on-center, green off-center). Much rarer are B–(R&G), in which B cones, in either center or surround, are opposed by some combination of R and G cones in the other spatial region. Furthermore, from the ratio of receptors to ganglion cells it appears that foveal retinal ganglion cell centers

are driven by a single cone. Thus the lattice of inputs, at least in the fovea, is presumably color opponent. Since oriented cortical cells are generally found to have negligible color opponency, we may ask how they are constructed from these color opponent inputs.

The simplest answer is to combine spatially coincident inputs of opposite center color before forming the kernel. For example, if a retinal input in a previous scheme was assigned a weight of 1, we now assign weights of 1 to both a red on-center cell and a coextensive green on-center cell. The resulting cortical receptive field will have no color opponency. It has been argued that retinal and geniculate cells multiplex both luminance and color information (Ingling & Martinez, 1983; Derrington, Krauskopf & Lennie, 1984). By combining chromatically opponent cells we 'de-multiplex' the luminance information.

At level 0 of the pyramid, this scheme is problematic, since it requires both R and G centers at each point in the initial hexagonal sample lattice. But as noted above, for foveal ganglion cells there is evidently *only* an R *or* a G center at each point, but not both. This is only a problem for the highest-frequency cells (level 0), and it may be that without the demultiplexing operation, chromatic abberation and the quasi-random assignment of cone types to input nodes renders these highest-frequency cells effectively achromatic.

Push–pull inputs. Assigning a positive weight to an input may mean any of three things: excitatory input from an on-center cell, inhibitory input from an off-center cell, or both. The last scheme has the advantage of rejecting certain nonlinearities common to both on- and off-center cells (Derrington, this volume). It also has the advantage of rejecting noise common to both cells that arises distal to the image. In the present case, use of push–pull inputs would mean that each of the seven nodes in the hexagon of inputs would receive four spatially coextensive inputs: on- and off-centers of both R and G types. Whether push–pull inputs are used in the construction of cortical cells is not yet known, though they appear to be used in the retina (Sterling, 1983).

Spatial inhomogeneity and disorder. The constructions above are based on a hexagonal lattice with spacing λ, identified with the cone spacing. As noted above, in the primate retina this spacing increases with distance from the fovea (Hirsch & Miller, 1987; Perry & Cowey, 1985). There is also a general increase in the disorder of the cone lattice, proceeding from foveal sections of nearly perfect hexagonal packing to peripheral sections with high disorder. There are several ways to generalize the chexagon models to the case of an inhomogeneous, disordered retina. One is to establish cortical connections with an ordered, homogeneous retina, and to preserve these connections during a process of retinal warping and perturbation. A second is to apply the chexagon construction to the six nearest neighbors, regardless of their precise locations. These speculations are perhaps better deferred until a clearer picture of retinal and cortical development is available.

The chexagon constructions described above specify a fixed convergence from receptors to ganglion cells, to geniculate, and to cortex. But it is known that the ratio of numbers of these four cell types devoted to a given retinal area vary considerably with eccentricity. Specifically, the ratio of cones to ganglion cells is about 1 within 5 mm (about 20°) of the fovea, increasing to 2 or 3 in the far periphery (Perry & Cowey, 1985). In contrast, the ratio of parvocellular geniculate cells to the P_β ganglion cells (anatomically distinct cells believed to be the inputs to the parvocellular pathway in the geniculate) is approximately 1 throughout the visual field (Perry & Cowey, 1985). Finally, the ratio of cortical cells to retinal cells is much higher in the fovea than in the periphery. The cortex devotes about 35% of its cells to the central 10 degrees, while the retina devotes only about 7%. These spatial variations in the pattern of convergence are a challenge for future models.

Conclusions

I have reviewed a number of the facts and principles that must be observed by a theory of operation of striate visual cortex. This list has necessarily been fragmentary. I have presented a general scheme in which oriented, band-pass cortical receptive fields are constructed from hexagonal kernels of weights applied to retinal or geniculate inputs. Two specific schemes were considered: one with greater physiological plausibility, the other with greater mathematical elegance. We may hope that a future scheme will combine both virtues. Both schemes are successful, however, in producing the general features of cortical processing, in particular band-pass receptive fields with moderate bandwidths in frequency and orientation. In the case of the orthogonal filters, quadrature pairs are also produced. Both schemes are exciting in that they suggest a wealth of questions and experiments, as well as refinements of the theory. A natural direction for future development is to consider kernels of more than seven inputs. This would allow greater degrees of freedom in design of the receptive field,

and hence a better match to physiological data. Another research direction is to consider how the construction of chexagon receptive fields might be manifest in the microcircuitry of the cortex (Gilbert, 1983).

In exploring this domain we frequently experience a tension between mathematical simplicity and biological complexity. We are reminded of Niels Bohr's dictum that 'truth' is the variable complementary to 'clarity'. It may help to consider two possible states of nature. In one, the visual brain obeys some simple architectural principle, and the variability and imperfection we observe are merely developmental perturbations. In this case, there is good reason to search for the guiding principle. In the second case, cortical 'structure' is a byproduct of some mindless evolutionary and developmental process. But while such a process may result in a less orderly structure, it is not random and can also be probed for its 'principle'. Thus in both cases it is worth searching for a canonical algotecture.

The cortical code poses an exciting and fundamental challenge for neuroscience, for cognitive science, and for vision science. It is a game that must be played by the rules, from experimental data to abstract principles of visual coding. A solution to the puzzle must work on all levels, physiological, anatomical, mathematical, perceptual and functional. The scheme I have proposed, while it does not fit every aspect of the data, at least provides an example of an hypothesis that works at every level, and that may encourage newer and better proposals. We will not solve the puzzle until we start to put the pieces together.

References

Ahumada, A. J., Jr. & Poirson, A. (1987) Cone sampling array model. *J. Opt. Soc. Am. A*, **4**, 1493–502.

Barlow, H. B. (1983) Understanding natural vision. In *Physical and Biological Processing of Images*, ed. O. J. Braddick & A. C. Sleigh, pp. 2–14. New York: Springer-Verlag.

Blakemore, C. B. & Vital-Durand, F. (1981) Distribution of X- and Y-cells in the monkey's lateral geniculate nucleus. *J. Physiol. (Lond.)*, **320**, 17–18P.

Blasdel, G. G. & Lund, J. S. (1983) Terminations of afferent axons in macaque striate cortex. *J. Neurosci.*, **3**, 1389–413.

Boycott, B. B. & Dowling, J. E. (1969) Organization of primate retina. *Phil. Trans. Roy. Soc.*, B**255**, 109–84.

Buchsbaum, G. (1987) Image coding in the visual system. *Optic News*, **13**, 16–19.

Burt, P. J. & Adelson, E. H. (1983) The laplacian pyramid as a compact image code. *IEEE Transactions on Communications*, COM-31, 532–40.

Chellappa, R. & Sawchuck, A. (ed.) (1985) *Digital Image Processing and Analysis Volume 1: Digital Image Processing*. Silver Spring, MD: IEEE Computer Society Press.

Crick, F. H. C., Marr, D. C. & Poggio, T. (1981) An information processing approach to understanding the visual cortex. In *The Organization of the Cerebral Cortex*, ed. F. O. Schmitt, F. G. Worden, G. Adelman & S. G. Dennis. Cambridge, Mass.: MIT Press.

Daniel, P. M. & Whitteridge, D. (1961) The representation of the visual field on the cerebral cortex in monkeys. *J. Physiol. (Lond.)*, **159**, 203–21.

Daugman, J. G. (1985) Uncertainty relation for resolution in space, spatial frequency and orientation optimized by two-dimensional visual cortex filters. *J. Opt. Soc. Am. A*, **2**, 1160–9.

De Monasterio, F. M. (1978) Properties of concentrically organized X and Y ganglion cells of macaque retina. *J. Neurophysiol.*, **41**, 1394–417.

Derrington, A. M. & Lennie, P. (1984) Spatial and temporal contrast sensitivities of neurones in lateral geniculate nucleus of macaque. *J. Physiol. (Lond.)*, **357**, 219–40.

Derrington, A. M., Krauskopf, J. & Lennie, P. (1984) Chromatic mechanisms in lateral geniculate nucleus of macaque. *J. Physiol. (Lond.)*, **357**, 241–65.

De Valois, R. L., Albrecht, D. G. & Thorell, L. G. (1982) Spatial Frequency Selectivity of Cells in Macaque Visual Cortex. *Vision Res.*, **22**, 545–59.

De Valois, R. L., Yund, E. W. & Hepler, H. (1982) The orientation and direction selectivity of cells in macaque visual cortex. *Vision Res.*, **22**, 531–44.

Dudgeon, D. A. & Mersereau, R. M. (1984) *Multidimensional Digital Signal Processing*. Englewood Cliffs, NJ: Prentice-Hall.

Enroth-Cugell, C., Robson, J. G., Schweitzer-Tong, D. & Watson, A. B. (1983) Spatio-temporal interactions in cat retinal ganglion cells showing linear spatial summation. *J. Physiol. (Lond.)*, **341**, 279–307.

Field, D. J. (1987) Relations between the statistics of natural images and the response properties of cortical cells. *J. Opt. Soc. Am. A*, **4**(12), 2379–94.

Field, D. J. & Tolhurst, D. J. (1986) The structure and symmetry of simple-cell receptive-field profiles in the cat's visual cortex. *Proc. Roy. Soc. Lond.*, **228**, 379–400.

Foster, K. H., Gaska, J. P., Nagler, M. & Pollen, D. A. (1985) Spatial and temporal frequency selectivity of neurones in visual cortical areas V1 and V2 of the macaque monkey. *J. Physiol. (Lond.)*, **365**, 331–63.

Geisler, W. S. & Hamilton, D. B. (1986) Sampling-theory analysis of spatial vision. *J. Opt. Soc. Am. A*, **3**, 62–70.

Gilbert, C. D. (1983) Microcircuitry of the visual cortex. *Ann. Rev. Neurosci.*, **6**, 217–47.

Green, D. M. & Swets, J. A. (1966) *Signal Detection Theory and Psychophysics*. New York: Wiley.

Hawken, M. J. & Parker, A. J. (1984) Contrast Sensitivity and

orientation selectivity in lamina IV of the striate cortex of old world monkeys. *Exp. Brain Res.*, **54**, 367–72.

Hawken, M. J. & Parker, A. J. (1987) Spatial properties of neurons in the monkey striate cortex. *Proc. Roy. Soc. Lond.*, B**231**, 251–88.

Hendrickson, A. E., Wilson, J. R. & Ogren, M. P. (1978) The neuroanatomical organization of pathways between the dorsal lateral geniculate nucleus and visual cortex in old and new world primates. *J. Comp. Neurol.*, **182**, 123–36.

Hirsch, J. & Miller, W. H. (1987) Does cone positional disorder limit resolution? *J. Opt. Soc. Am. A*, **4**, 1481–92.

Hubel, D. H. & Wiesel, T. N. (1968) Receptive fields and functional architecture of monkey striate cortex. *J. Physiol. (Lond.)*, **195**, 215–43.

Hubel, D. H. & Wiesel, T. N. (1972) Laminar and columnar distribution of geniculo-cortical fibres in macaque monkey. *J. Comp. Neurol.*, **146**, 421–50.

Hubel, D. H. & Wiesel, T. N. (1977) Functional architecture of the macaque monkey visual cortex. Ferrier Lecture. *Proc. Roy. Soc. Lond. (Biology)*, **198**, 1–59.

Hughes, A. (1981) Cat retina and the sampling theorem: the relation of transient and sustained brisk-unit cut-off frequency to a- and b-mode cell density. *Exp. Brain Res.*, **42**, 196–202.

Ingling, C. R. & Martinez, E. (1983) The spatiochromatic signal of the r–g channel. In *Colour Vision*, ed. J. D. Mollon & L. T. Sharpe, pp. 433–44. London: Academic Press.

Jerri, A. J. (1977) The Shannon sampling theorem—its various extensions and applications: a tutorial review. *Proceedings of the IEEE*, **65**, 1565–96.

Kaplan, E. & Shapley, R. M. (1982) X- and Y-cells in the lateral geniculate nucleus of macaque monkeys. *J. Physiol. (Lond.)*, **330**, 125–43.

Kaplan, E. & Shapley, R. M. (1986) The primate retina contains two types of ganglion cells, with high and low contrast sensitivity. *Proc. Natl. Acad. Sci. USA*, **83**, 2755–7.

Koenderink, J. J. & van Doorn, A. J. (1987) Representation of local geometry in the visual system. *Biol. Cybern.*, **55**, 367–75.

Kulikowski, J. J., Marcelja, S. & Bishop, P. O. (1982) Theory of spatial position and spatial frequency relations in the receptive fields of simple cells in the visual cortex. *Biol. Cybern.*, **43**, 187–98.

Lennie, P. (1980) Parallel visual pathways: a review. *Vision Res.*, **20**, 561–94.

Livingstone, M. S. & Hubel, D. H. (1984) Anatomy and physiology of a color system in the primate visual cortex. *J. Neurosci.*, **4**, 309–56.

Maloney, L. T. & Wandell, B. A. (1986) Color constancy: a method for recovering surface spectral reflectance. *J. Opt. Soc. Am. A*, **3**, 29–33.

Mandelbrot, B. B. (1983) *The Fractal Geometry of Nature*. New York: Freeman.

Marcelja, S. (1980) Mathematical description of the responses of simple cortical cells. *J. Opt. Soc. Am.*, **70**, 1297–300.

Marr, D. & Hildreth, E. (1980) Theory of edge detection. *Proc. Roy. Soc. Lond.*, B**207**, 187–217.

Movshon, J. A., Thompson, I. D. & Tolhurst, D. J. (1978a) Spatial summation in the receptive fields of simple cells in the cat's striate cortex. *J. Physiol. (Lond.)*, **283**, 53–77.

Movshon, J. A., Thompson, I. D. & Tolhurst, D. J. (1978b) Spatial and temporal contrast sensitivity of neurones in areas 17 and 18 of the cat's visual cortex. *J. Physiol. (Lond.)*, **283**, 101–20.

Perry, V. H. & Cowey, A. (1981) The morphological correlates of X- and Y-like retinal ganglion cells in the retina of monkeys. *Exp. Brain Res.*, **43**, 226–8.

Perry, V. H. & Cowey, A. (1985) The ganglion cell and cone distributions in the monkey's retina: implications for central magnification factors. *Vision Res.*, **25**, 1795–810.

Perry, V. H., Oehler, R. & Cowey, A. (1984) Retinal ganglion cells that project to the dorsal lateral geniculate nucleus in the macaque monkey. *Neurosci.*, **12**, 1101–23.

Pollen, D. A. & Ronner, S. F. (1981) Phase relationship between adjacent simple cells in the visual cortex. *Science*, **212**, 1409–11.

Robson, J. G. (1983) Frequency domain visual processing. In *Physical and Biological Processing of Images*, ed. O. J. Braddick & A. C. Sleigh, pp. 73–87. Berlin: Springer-Verlag.

Sakitt, B. & Barlow, H. B. (1982) A model for the economical encoding of the visual image in cerebral cortex. *Biol. Cybern.*, **43**, 97–108.

Schiller, P. H., Finlay, B. L. & Volman, S. F. (1976) Quantitative studies of single cell properties in monkey striate cortex. III. Spatial frequency. *J. Neurophysiol.*, **39**, 1334–51.

Shannon, C. E. (1949) Communication in the presence of noise. *Proceedings IRE*, **37**, 10–21.

Shapley, R. M. & Lennie, P. (1985) Spatial frequency analysis in the visual system. In *Annual Review of Neuroscience*, Vol. 8, ed. W. M. Cowan, pp. 547–83. Palo Alto, CA: Annual Reviews, Inc.

So, Y. T. & Shapley, R. (1981) Spatial tuning of cells in and around lateral geniculate nucleus of the cat: X and Y relay cells and perigeniculate interneurons. *J. Neurophysiol.*, **45**, 107–20.

Sterling, P. (1983) Microcircuitry of the cat retina. *Ann. Rev. Neurosci.*, **6**, 149–85.

Tanimoto, S. & Pavlidis, T. (1975) A Hierarchical data structure for picture processing. *Computer Graphics and Image Processing*, **4**, 104–19.

Van Essen, D. C., Newsome, W. T. & Maunsell, J. H. R. (1984) The visual field representation in striate cortex of the macaque monkey: asymmetries, anisotropies, and individual variability. *Vision Res.*, **24**, 429–48.

Watson, A. B. (1982) Summation of grating patches indicates many types of detector at one retinal location. *Vision Res.*, **22**, 17–25.

Watson, A. B. (1983) Detection and identification of simple spatial forms. In *Physical and Biological Processing of Images*, ed. O. J. Braddick & A. C. Sleigh, pp. 100–14. Berlin: Springer-Verlag.

Watson, A. B. (1986) *Ideal Shrinking and Expansion of Discrete*

Sequences. NASA Technical Memorandum 88202. Washington DC: NASA.

Watson, A. B. (1987a) The cortex transform: rapid computation of simulated neural images. *Computer Vision, Graphics, and Image Processing*, **39**, 311–27.

Watson, A. B. (1987b) Efficiency of an image code based on human vision. *J. Opt. Soc. Am. A*, **4**, 2401–17.

Watson, A. B. & Ahumada, A. J., Jr. (1987) *An Orthogonal Oriented Quadrature Hexagonal Image Pyramid*. NASA Technical Memorandum 100054. Washington DC: NASA.

Watson, A. B., Ahumada, A. J. Jr. & Farrell, J. (1986) Window of visibility: psychophysical theory of fidelity in time-sampled visual motion displays. *J. Opt. Soc. Am. A*, **3**, 300–7.

Webster, M. A. & De Valois, R. L. (1985) Relationship between spatial-frequency and orientation of striate-cortex cells. *J. Opt. Soc. Am. A*, **2**, 1124–32.

Williams, D. R. (1985) Aliasing in human foveal vision. *Vision Res.*, **25**, 195–205.

Wilson, H. R. & Bergen, J. R. (1979) A four mechanisms model for threshold spatial vision. *Vision Res.*, **19**, 19–33.

Woods, J. W. & O'Neil, S. D. (1986) Subband coding of images. *IEEE Transactions on Acoustics, Speech, and Signal Processing ASSP 34*, 1278–88.

Yellott, J. I. (1982) Spectral analysis of spatial sampling by photoreceptors: typological disorder prevents aliasing. *Vision Res.*, **22**, 1205–10.

Yellott, J. I. (1983) Spectral consequences of photoreceptor sampling in the Rhesus retina. *Science*, **221**, 382–5.

Young, R. A. (1985) *The Gaussian Derivative Theory of Spatial Vision: Analysis of Cortical Cell Receptive Field Line-Weighting Profiles*. General Motors Research Report GMR-4920. Warren MI: General Motors.