Two-dimensional spatial structure of receptive fields in monkey striate cortex

A. J. Parker and M. J. Hawken

University Laboratory of Physiology, Parks Road, Oxford OX1 3PT, England, UK

Received April 14, 1987; accepted October 14, 1987

Measurements of the spatial contrast sensitivity function and orientation selectivity of visual neurons in the foveal striate cortex (VI) of primates were interpreted within the context of a model of the two-dimensional spatial structure of their receptive fields. Estimates of the spatial dimensions of the receptive fields along the axis of preferred orientation were derived from the application of the model and were compared with estimates of the smallest spatial subunit in the dimension orthogonal to the preferred orientation. Some measure of agreement was found with corresponding estimates of parameters for psychophysical channels in human foveal vision.

INTRODUCTION

Biological visual systems have evolved over a long period of time and presumably represent a state of highly efficient organization for the computational task of visual information processing. This consideration has led to suggestions that there is special significance attached to the exact form of the spatial organization of the mechanisms responsible for early visual processing. The properties of these spatial filters have been revealed by psychophysical experiments designed to measure the detection of stimuli as a function of contrast, the identification of stimuli at threshold, the localization of targets with the precision of hyperacuity, and the consequences of noise perturbation of the stimulus on performance with these tasks. In this paper we make use of an empirically based model for the behavior of VI neurons in response to stimuli of optimal orientation together with measurements of their orientation selectivity to give estimates of spatial summation along the axis of preferred orientation. These estimates of receptive-field “height” are compared with estimates of the dimensions of spatial filters inferred from psychophysical experiments.

BACKGROUND

The qualitative nature of the two-dimensional structure of visual receptive fields in cat area VI was made clear by the experiments of Hubel and Wiesel, who proposed the existence of two types of cell, simple and complex, based on the spatial organization of their receptive fields. The receptive fields of simple cells could be divided into ON and OFF zones, which were found to be elongated in shape, with the long axis parallel to the preferred orientation for a bar stimulus. Knowledge of the shape and spatial arrangement of these zones allowed qualitative predictions of the responses of simple cells to a variety of stimuli. By contrast, complex cells could not be mapped into spatially discrete ON and OFF zones, even though they resembled simple cells in several other respects: orientation selectivity, existence of binocularity and direction selectivity, and, at any given eccentricity, a preference for roughly similar sizes of bar stimulus. Subsequently, it was shown that complex cells also resembled simple cells in the peaks and bandwidths of their spatial-frequency tuning and their temporal properties.

The generation and validation of a full quantitative two-dimensional model for receptive fields in cortical area VI is a more-complicated problem. Even major quantitative studies have so far succeeded in making only a partial contribution. Differences in the spatial summation behavior of simple and complex cells for stimuli of optimal orientation have been delineated quantitatively in cat and in monkey. The space-domain and spatial-frequency-domain descriptions of the receptive field are in reasonable agreement for simple cells but differ markedly for complex cells. Many of the similarities and differences of simple and complex cells can be resolved by supposing that complex cells are constructed from a number of subunits, each of which shows linear spatial summation, but which are combined in a nonlinear way. Each linear subunit would be similar in size and spatial arrangement to the simple cells in nearby regions of cortex, though the subunits of complex cells need not necessarily be derived directly from simple cells. For both simple and complex cells in the cat, summation in the dimension parallel to the preferred orientation can be described by a single Gaussian, provided that a straightforward correction is made for the existence of a threshold nonlinearity. More accurate tests of the summation-to-threshold model for the length summation behavior of cat simple cells confirm this conclusion.

A function describing receptive-field organization in the dimension orthogonal to the preferred orientation needs to exhibit the qualitative property of multiple ON and OFF zones in the receptive field. Our recent analysis of the ability of a number of models to describe spatial contrast sensitivity functions led to the formation of a model in which each major ON and OFF zone in a simple cell’s receptive field is described by a difference-of-Gaussians (DOG) function. The peaks of these DOG functions are spatially offset from each other to correspond with the spatial separation of the ON and OFF zones. For a complex cell, we would conjecture that this spatial arrangement would apply to each linearly summing subunit, but the number and distribution of spatial subunits as well as the exact nonlinear form of their combination require further investigation.
In this paper we make use of a one-dimensional model defined by a difference-of-DOG's (d-DOG-S) function to describe the receptive fields of V1 neurons in the dimension orthogonal to their preferred orientation:

\[
\begin{align*}
  &k_c \exp[-(x/x_c)^2] - k_s \exp[-(x/x_s)^2] \\
  &- (1-g)(k_c \exp[-(x+S/x_c)^2] - k_s \exp[-(x+S/x_s)^2]) \\
  &- g(k_c \exp[-(x-S/x_c)^2] - k_s \exp[-(x-S/x_s)^2]).
\end{align*}
\]

(1)

In this expression, \(x\) represents position in Cartesian coordinates; \(x_c, x_s\) are the space constants of the component Gaussians; and \(k_c, k_s\) control the scaling of the Gaussians. Each line of the equation represents a one-dimensional DOG function. The first line is a DOG whose peak is at the origin, whereas the second and third line represent DOG's whose peaks are at flanking positions \(+S\) and \(-S\) relative to the origin. The relative strength of the flanking DOG's is controlled by the parameter \(g\), which allows for differences in the spatial phase organization of different receptive fields. Some constraints were enforced during the fitting of this model to the data. These are presented in detail elsewhere; briefly, they consisted of the following: \(x_c \leq x_s; x_c \leq x_{1c}; x_c < x_{1s}; x_c < x_{2s}; S < 2(x_c + x_s); 0 \leq g \leq 1\). The primary reason for choosing the d-DOG-S function to describe the one-dimensional behavior of the receptive field is its accuracy, but other models that are sufficiently accurate would be equally valid for analyzing the results presented here. The model can be expanded into a two-dimensional form that consists of the product of a d-DOG-S function with a single Gaussian describing the "height" of the receptive field parallel with the axis of preferred orientation. The expression describing the height dimension of the receptive field is

\[
\exp[-(y/y_h)^2],
\]

(2)

where \(y\) denotes position in Cartesian coordinates and \(y_h\) denotes the space constant of the Gaussian describing the height dimension. Figure 1 shows an example of a perspective plot of the function defined by the two-dimensional model.

The model makes the twin assumptions of mathematical separability in Cartesian coordinates and linearity of spatial summation in a full two-dimensional form, both of which are in need of more extensive formal validation for neurons in area V1. In favor of the idea that these assumptions are not unreasonable are the above-mentioned findings on the nature of spatial summation; recent findings on the nature of excitatory and inhibitory postsynaptic potentials recorded intracellularly in cat cortical cells; and the results of attempts to probe the question of polar separability in cortical neurons. However, against the model are the facts that it does not deal adequately with end inhibition in cortical receptive fields or with inhibitory interactions in the orientation domain (or, indeed, any domain). If the restrictions of this framework are accepted, measurements of the orientation bandwidth of these cells permits an estimate of the space constant \(y_h\) of the Gaussian describing the height of the receptive field, thus yielding a full two-dimensional description of the spatial organization of the receptive field.

The purpose of this paper is to present the distribution of values for the space constant of the Gaussian describing the height dimension for a set of neurons in V1 with foveal receptive fields, to examine some of the interrelationships of these values with other spatial properties of the receptive field, and to compare these values with those from previous neurophysiological investigations and with the properties of psychophysical channels in human foveal vision.

**METHODS**

A detailed description of methods has been given elsewhere. Extracellular recordings were made from 119 single units (eccentricities between 0.2 and 3.5 deg) in the cortical area V1 of anesthetized, paralyzed Old World monkeys.
tive fields were initially tested with hand-held stimuli to assess the orientation selectivity, ocular dominance, and spatial summation properties. The spatial summation properties were investigated quantitatively with phase-reversing gratings. Simple cells were taken as those responding predominantly at the same frequency as the phase reversal where the response changed systematically with the spatial phase of the stimulus, whereas complex cells responded predominantly at twice the frequency of the phase reversal, and the response was typically invariant with the spatial phase of the stimulus. Orientation selectivity was measured quantitatively by using drifting sinusoidal gratings with spatial and temporal frequencies that gave an optimal response from the cell. Contrast sensitivity for gratings of optimal orientation, speed, and direction of drift was measured as a function of spatial frequency; a staircase procedure was used to analyze the variability of the resting discharge in the absence of a stimulus and then to search for the minimum contrast required to bring the cell's discharge reliably above the resting firing rate.

RESULTS

An example of the type of results from which parameters were extracted is shown in Fig. 2. The graph in Fig. 2(a) shows the contrast sensitivity of a simple cell as a function of spatial frequency. The smooth curve through the data is the best-fitting version of the Fourier amplitude spectrum of a d-DOG-S function. (The contrast sensitivity data are from one of the cells that were presented in detail for the analysis of the shape of the one-dimensional model; that paper contains details of the procedure for fitting curves and of the parameters extracted by application of the model.) The values of the parameters associated with the best-fitting curve are consistent with estimates of the spatial properties of foveal monkey cortical cells obtained with bar stimuli and with the proposal that the zones of simple cell receptive fields are made up from inputs closely matched to the characteristics of receptive fields in the lateral geniculate nucleus.

After a description of the contrast sensitivity function was obtained by using the one-dimensional d-DOG-S function, the orientation bandwidth of the cell was used to obtain the space constant of the Gaussian describing the profile along the preferred orientation. The solid symbols in Fig. 2(b) show responses to a grating of 6 cycles/deg at orientations around the peak in 10-deg steps. The smooth curve shows the predicted orientation tuning function from the two-dimensional model. Orientation selectivity was measured by using a response measure rather than a sensitivity measure, which may introduce some bias into our analysis, but results in the cat suggest that orientation bandwidth, the relevant parameter for these purposes, changes relatively little with the contrast of the test grating. The smooth curve indicating the model's prediction for orientation selectivity has three free parameters: the height and location of the peak.

![Graph showing contrast sensitivity as a function of spatial frequency](image1)

**Fig. 2.** (a) Spatial contrast sensitivity as a function of spatial frequency for a direction-selective simple cell recorded in layer 6 of area V1. Each data point shows the mean of 12 determinations of contrast sensitivity derived from a staircase procedure for determining the threshold. The smooth curve through the data is the best-fitting version of the d-DOG-S model. Details of the procedure for fitting the curve, the exact specification of the model, and the resulting values of parameters were presented previously. (b) Orientation tuning function for the same neuron as in (a). The abscissa shows the orientation of the high-contrast test grating; the ordinate shows the number of spikes per presentation. Each presentation lasted 0.77 sec and consisted of two complete cycles of drifting grating. The smooth curve through the data illustrates the shape of the orientation tuning curve predicted by the two-dimensional model. The height and the location of the peak of the smooth curve were allowed to vary freely, and these two parameters reflect, respectively, the peak response and the preferred orientation of this cell at the test spatial frequency; the spread of the orientation tuning function is determined by the parameters from the fit of the d-DOG-S function to spatial contrast sensitivity measurements, as shown in (a), and by the space constant $(\gamma_0)$ of the Gaussian describing the height dimension, which was free to vary. The model function describes the data quite accurately. The value of the height parameter $(\gamma_0)$ inferred from this procedure is 11.3 arcmin.
and the spread of the curve at half-height. (The height of the peak is a free parameter because the orientation selectivity was estimated by using a response measure rather than a sensitivity measure.) The parameters needed for the d-DOG-S function were taken from the fit of that function to the spatial contrast sensitivity measurements [Fig. 2(a)]. Within the framework of the model, the orientation bandwidth fixes a value for the space constant of the Gaussian describing the height ($y_m$) of the receptive field.

Measurements of orientation selectivity were analyzed from 56 simple cells and 49 complex cells. In addition, seven cells with measurable contrast sensitivity functions were excluded because they were so strongly end stopped with extended gratings that an orientation tuning curve could not be generated, and a further seven cells were excluded because they were nonoriented or very weakly biased for orientation and their orientation bandwidth could not be defined. The orientation bandwidth of each cell (the spread of orientation selectivity in degrees at half the height of the peak) was measured directly from plots of the response tuning
data. The distribution of orientation bandwidths for simple and complex cells is shown in Fig. 3. An iterative search procedure was used to find the value of height \( y_h \) that was consistent with the orientation bandwidth, with the spatial frequency at which the orientation selectivity was measured, and with the values of the spatial parameters resulting from the best-fitting version of the d-DOG-S model to the contrast sensitivity function of the cell in question. Given the generally smooth nature of the functions describing the model, the resulting estimate of the height \( y_h \) is unique. The distribution of estimates for height is quite broad for both simple cells [Fig. 4(a)] and complex cells [Fig. 4(b)] and ranges from <5 to >40 arcmin. These values are comparable with those found by direct measurement of length summation in other neurophysiological experiments.\(^{12}\)

It is informative to compare the values for height \( y_h \) with those for other parameters of the receptive field. This allows some assessment of the extent to which the scatter in

![Histograms of the values of receptive-field center sizes \( x_{c0} \) derived from the fit of a d-DOG-S function to spatial contrast sensitivity measurements: (a) simple cells and (b) complex cells. It is notable that there is a group of simple cells with large center sizes (>5 arcmin), whereas there is only one complex cell whose center size exceeds this value.](image)

![Histograms of the ratio of values of the height dimension \( y_h \) relative to the receptive-field-center size \( x_{c0} \): (a) simple cells and (b) complex cells. It can be seen that the modal value for both groups of cells is about 3.0, indicating that the length or height of the receptive field is on average 3 times the size of the smallest excitatory region. However, there is a considerable range of values.](image)
values represents changes in the size of the receptive field by an overall scaling factor as opposed to changes in the shape of the receptive field independent of scale. When a model based on DOG functions is used, one reasonable choice for comparison is the smallest Gaussian subunit (or subunit) in the receptive field. Just as for retinal ganglion cells, the space constant of this smallest subunit \( x_0 \) is determined by the shape and the location of the contrast sensitivity function near its high-frequency cutoff (measured with gratings of preferred orientation in the case of cortical cells). Indeed, at spatial frequencies close to the cutoff, the effect of other Gaussian mechanisms in a model based on DOG functions is negligible, and for this reason all three variants of DOG-based models explored in our previous paper estimate similar values for this parameter. By analogy with retinal and lateral geniculate nucleus neurons, this parameter will be called the receptive-field center size.

Figure 5 shows the distribution of the receptive-field center size (\( x_0 \)). As before, simple cells are shown in the upper histogram [Fig. 5(a)], and complex cells are shown in the lower histogram [Fig. 5(b)]. The values range from around 1 to 10 arcmin for simple cells over the eccentricity range studied here (0.2–3.5 deg). This distribution agrees well with certain estimates of the range inferred for physiological channels from measurements of contrast detection and spatial acuity, where the range of center sizes over the same eccentricity range is slightly less than 1 arcmin up to 10.7 arcmin. However, there are also clear differences between the psychophysical results and the predictions of psychophysical models: cells with the same peak spatial frequency and eccentricity may have markedly different bandwidths, some cells having bandwidths broader than those assumed in most psychophysical models. For complex cells, these measurements could be interpreted as indicating the size of the smallest discrete zone or zones within the receptive field; this could, for example, be an ON or OFF zone within a linear subunit, or several similarly sized subunits. It is interesting to note that the center sizes of complex cells assessed in this way do not cover such a wide range as simple cells, owing to the apparent absence of complex cells with large receptive-field center sizes.

Figure 6 shows the ratio of height \( y_0 \) relative to receptive-field center size (\( x_0 \)). This pair of histograms makes it clear that height is always larger than center size; sometimes it is many times larger. Psychophysical measurements of this parameter measured by orientation specific masking find the ratio of height to center size to be 3.2, a value close to the means of the distributions in Fig. 6 for both simple and complex cells. The value of 3.2 is also appropriate for modeling psychophysical hyperacuity thresholds, such as the changes in Vernier acuity with line length. Even though Wilson used a different form of DOG model for his psychophysical channels instead of the d-DOG-S function used here, the fact that both models have a single center mechanism that dominates at high spatial frequencies and a single Gaussian mechanism to describe the height dimension means that these particular parameters can be fairly compared with each other.

**DISCUSSION**

It must be acknowledged that, while this study has produced a set of plausible estimates for some features of the two-dimensional spatial profile of receptive fields of V1 neurons, there is a need for independent verification of the approach and the parameter values calculated by using it. At best this would consist of proof that the model, or a modification of it, would be capable of simulating the response of cortical cells to arbitrary spatial patterns, but this ideal goal can only be approached in a piecewise fashion. We are currently comparing the length summation characteristics of neurons in area V1 measured directly with those derived from the indirect approach adopted here in order to explore the basis of the current calculations in more detail.

**Shapes of Spatial Filters**

It would be natural to expect that results of this kind would point to a definition of the canonical form of spatial filters responsible for early processing in primate foveal vision. Certainly, there is good agreement in some cases between the estimates from psychophysical measurements and those obtained here. Elsewhere, we have presented detailed arguments in favor of the d-DOG-S model on the grounds that it is physiologically more realistic. First, the center space constants \( x_{c1}, x_{c2} \) of all the subregions of receptive fields in area V1 estimated by using this model fall within the range of space constants for receptive-field centers in the lateral geniculate nucleus. Second, the spatial separation of the subregions acts to enhance the sensitivity of the cells at intermediate spatial frequencies. Although the d-DOG-S model is both accurate and physiologically realistic, it is not obviously in accord with any theoretical analysis of the computational constraints on visual information processing. Thus one may expect future developments of theory to bring forward proposals for mathematical descriptions of spatial filters that would reveal clearly the relationship of the empirical model utilized here and any novel theoretical model based on an improved understanding of visual information processing.

However, the present analysis shows some of the diversity of spatial properties (notably the ratio of height to center size, i.e., \( y_0/x_0 \)) within the population of receptive fields in V1. This diversity is genuine; altering the values of parameters within a relatively restricted range often made the fit to the data significantly worse. At first sight, this diversity would appear to stymie any attempts to find a simple principle underlying the spatial organization of receptive fields. However, the population of neurons in area V1 serves a wide range of information-processing tasks (binocular matching, motion detection, spatial analysis, color), and it is reasonable that this diversity of needs should be reflected in a diversity of properties. Thus the path to identifying the principles underlying the shapes of spatial filters in the visual system is to consider in more detail the needs of the separate information-processing tasks individually and to make comparisons between cortical cells and theory only after these issues have been clarified.

The restricted size of foveal receptive fields in primate area V1 offers some insight into the apparent discrepancies between psychophysical contrast sensitivity for extended grating targets, which may be as high as 1000 at some spatial frequencies, and the contrast sensitivity of individual neurons at the same spatial frequencies, which is never in excess of 100. Each neuron collects information about an extended grating target through a window of visibility defined by the shape of its own receptive field. If this window of visibility
is simulated for a psychophysical observer by physically windowing the grating target on the display screen, then observers' thresholds agree well with the most sensitive cells in the population of V1 neurons. We are currently extending this approach by measuring the psychophysical sensitivity of human observers to test stimuli that exactly match the spatial profile of cortical receptive fields.

Cortical Organization of Receptive Fields
The elongation of V1 receptive-field zones implied by these interpretations of orientation selectivity is of considerable interest in the context of current attempts to understand cortical circuitry. The original proposal of Hubel and Wiesel was that this elongation arises as a consequence of the alignment of a number of circularly symmetric lateral geniculate nucleus receptive fields in a row, all contributing to the simple cell's receptive field. Although this proposal has been questioned, recent intracellular investigations of excitatory and inhibitory postsynaptic potentials indicate that evidence for alternative proposals is hard to come by. If the alignment hypothesis is correct, then the ratios of height to center size ($y_0/x_0$) should give some clue as to the minimum number of lateral geniculate nucleus inputs that would need to be aligned. However, it is not clear that all neurons in area V1 acquire their orientation selectivity in this way, since some are almost certainly driven by other cortical cells via intrinsic interconnections among different cortical laminae. Moreover, it is clear from our own data that the ratios of height to center size have a differential distribution within the cortical laminae.

Figure 5 suggests the possible existence of a group of linear simple cells with receptive-field center sizes ($x_0$) in the range 6–10 arcmin, much larger than all except one complex cell. Not all aspects of the functional significance of this difference are clear. The presence of simple cells with larger receptive-field center sizes is required to match conclusions about the range of spatial channels at the fovea that are drawn from psychophysical experiments on contrast detection. Some of these large-field simple cells are direction selective and are located anatomically in cortical layers 4b and 4c, so these direction-selective simple cells may be sending information to the extrastriate visual area MT that is thought to support motion detection. Thus large-field simple cells might be part of the pathway responsible for the identification of the direction of a moving target close to contrast threshold.

ACKNOWLEDGMENTS
This work was supported by a Wellcome Trust Major Equipment Grant, by the U.S. Air Force Office of Scientific Research under grant AFOSR-85-0396, and by Medical Research Council grant 7800491 to Colin Blakemore. During part of this work, A. J. Parker was supported by a Light research fellowship at St. Catherine's College, Oxford, UK.

REFERENCES