Neural Responses to Polar, Hyperbolic, and Cartesian Gratings in Area V4 of the Macaque Monkey

JACK L. GALLANT, CHARLES E. CONNOR, SUBRATA RAKSHIT, JAMES W. LEWIS, AND DAVID C. VAN ESSEN

Division of Biology, California Institute of Technology, Pasadena, California 91125; and Department of Anatomy and Neurobiology, Washington University School of Medicine, St. Louis, Missouri 63110

SUMMARY AND CONCLUSIONS

1. We studied the responses of 103 neurons in visual area V4 of anesthetized macaque monkeys to two novel classes of visual stimuli, polar and hyperbolic sinusoidal gratings. We suspected on both theoretical and experimental grounds that these stimuli would be useful for characterizing cells involved in intermediate stages of form analysis. Responses were compared with those obtained with conventional Cartesian sinusoidal gratings. Five independent, quantitative analyses of neural responses were carried out on the entire population of cells.

2. For each cell, responses to the most effective Cartesian, polar, and hyperbolic grating were compared directly. In 18 of 103 cells, the peak response evoked by one stimulus class was significantly different from the peak response evoked by the remaining two classes. Of the remaining 85 cells, 71 had response peaks for the three stimulus classes that were all within a factor of 2 of one another.

3. An information-theoretic analysis of the trial-by-trial responses to each stimulus showed that all but two cells transmitted significant information about the stimulus set as a whole. Comparison of the information transmitted about each stimulus class showed that 23 of 103 cells transmitted a significantly different amount of information about one class than about the remaining two classes. Of the remaining 80 cells, 55 had information transmission rates for the three stimulus classes that were all within a factor of 2 of one another.

4. To identify cells that had orderly tuning profiles in the various stimulus spaces, responses to each stimulus class were fit with a simple Gaussian model. Tuning curves were successfully fit to the data from at least one stimulus class in 98 of 103 cells, and such fits were obtained for at least two classes in 87 cells. Individual neurons showed a wide range of tuning profiles, with response peaks scattered throughout the various stimulus spaces; there were no major differences in the distributions of the widths or positions of tuning curves obtained for the different stimulus classes.

5. Neurons were classified according to their response profiles across the stimulus set with two objective methods, hierarchical cluster analysis and multidimensional scaling. These two analyses produced qualitatively similar results. The most distinct group of cells was highly selective for hyperbolic gratings. The majority of cells fell into one of two groups that were selective for polar gratings: one selective for radial gratings and one selective for concentric or spiral gratings. There was no group whose primary selectivity was for Cartesian gratings.

6. To determine whether cells belonging to identified classes were anatomically clustered, we compared the distribution of classified cells across electrode penetrations with the distribution that would be expected if the cells were distributed randomly. Cells with similar response profiles were often anatomically clustered.

7. A position test was used to determine whether response profiles were sensitive to precise stimulus placement. A subset of Cartesian and non-Cartesian gratings was presented at several positions in and near the receptive field. The test was run on 13 cells from the present study and 28 cells from an earlier study. All cells showed a significant degree of invariance in their selectivity across changes in stimulus position of up to 0.5 classical receptive field diameters.

8. A length and width test was used to determine whether cells preferring non-Cartesian gratings were selective for Cartesian grating length or width. Responses to Cartesian gratings shorter or narrower than the classical receptive field were compared with those obtained with full-field Cartesian and non-Cartesian gratings in 29 cells. Of the four cells that had shown significant preferences for non-Cartesian gratings in the main test, none showed tuning for Cartesian grating length or width that would account for their non-Cartesian responses. However, tuning for Cartesian grating length or width was demonstrated in five other cells in the sample.

9. The population of V4 neurons displayed a clear bias in their responses in favor of polar and hyperbolic stimuli, and some cells were highly selective for these stimuli. The Cartesian stimuli alone could not explain the responses of most cells to non-Cartesian stimuli. The fact that nearly all cells conveyed significant information about all three stimulus classes, and that most had identifiable tuning curves in multiple classes, suggests that V4 cells are neither simple feature detectors nor simple filters within a single restricted stimulus space. Tuning for multiple stimulus classes may reflect a particular visual processing function or a general principle such as efficient image encoding.

INTRODUCTION

Form vision involves several hierarchical stages of processing, from area V1 to the inferior temporal complex (Van Essen and Gallant 1994). One important intermediate area in this hierarchy is V4. Although originally thought to be primarily involved in color vision (Zeki 1975, 1983a), V4 is now recognized to be critical for form vision as well (Desimone and Schein 1987; Heywood et al. 1992; Schiller and Lee 1991). Studies of form processing in V4 have demonstrated selectivity for simple dimensions such as orientation, spatial frequency, length, and width, as well as several interesting nonlinear properties (Desimone and Schein 1987). Yet all of these properties have been observed to some degree in earlier stages, including area V1 (De Valois and De Valois 1990; Hubel and Wiesel 1968; Knierim and Van Essen 1992). Do neurons in V4 display specific selectivities that are not found in earlier areas, and that might be related to form processing?

Our approach to this problem has been to explore stimuli that are more complicated than the simple bars and Cartesian...
gratings used in most past studies of area V4 (Desimone and Schein 1987; Zeki 1983b). Rather than conducting a breadth-first analysis of a large range of stimuli (cf. Kobatake and Tanaka 1994), we examined a smaller set of stimuli in much more detail. Our stimulus set is composed of three separate stimulus classes: polar and hyperbolic sinusoidal gratings (collectively called non-Cartesian gratings), and conventional Cartesian gratings (see Fig. 1).

These three classes were chosen for several reasons. First, they collectively capture some of the richness of natural shape while allowing quantitative analysis of selectivity along well-defined dimensions. Second, they are related to motion patterns used in studies of the middle temporal area and the medial superior temporal area (area MST); Cartesian and non-Cartesian gratings are analogous to optical flow patterns such as translation, expansion, rotation,
and shear, produced by the motion of an animal through its environment. Such patterns are particularly effective stimuli for cells in the dorsal subdivision of area MST (MSTd), an area apparently specialized for processing optical flow (Duffy and Wurtz 1991a,b; Graziano et al. 1994; Orban et al. 1992; Saito et al. 1986; Tanaka and Saito 1989; Tanaka et al. 1989). Because the visual system often uses similar mechanisms to process information in time and space, it is logical to search for comparable responses to stationary non-Cartesian stimuli in area V4. A third reason for using non-Cartesian stimuli comes from theoretical studies proposing that polar or hyperbolic filters constitute an intermediate stage of form processing (Dodwell 1983; Eagleson 1992; Hoffman 1966; Koenderink and van Doorn 1990; Perona 1991). A better understanding of how cortical neurons respond to non-Cartesian stimuli should help to clarify the neurobiological relevance of such models.

In an earlier study (Gallant et al. 1993) we demonstrated that some V4 neurons are highly selective for specific non-Cartesian stimuli. At that time we sampled the non-Cartesian stimulus spaces relatively sparsely (20 non-Cartesian and 30 Cartesian stimuli), and our primary analysis was a comparison of the peak responses obtained with Cartesian versus non-Cartesian stimuli. In the present study we sampled the non-Cartesian stimulus space more densely (60 non-Cartesian and 30 Cartesian stimuli) in order to identify cells with significant non-Cartesian responses that might have been missed by our earlier screen. We also expanded our data analyses. One set of analyses is aimed at comparing responses across the stimulus classes. These include a more extensive analysis of peak response rates, an information-theoretic analysis, and estimation of tuning curves in Cartesian, hyperbolic, and polar stimulus spaces. Another set of analyses aims to classify cells according to their response profiles, and to test for an anatomic correlate of the physiological classification. Finally, we carried out subsidiary tests on a subset of neurons to determine the effects of stimulus position, length, and width on responses obtained with Cartesian and non-Cartesian gratings.

METHODS

Single-unit recordings were made from 10 macaque monkeys (Macaca nemestrina and Macaca mulatta) in acute procedures lasting from 2 to 7 days. Monkeys were prepared for recording either at the beginning of the acute recording session or in a separate surgery occurring 1–9 days before recording. Methods for surgery and acute recording were similar to those used previously in this laboratory (Felleman and Van Essen 1987; Olavarria et al. 1992). All procedures were carried out under institutionally approved protocols and conformed to the National Institutes of Health Guidelines for the Care and Use of Animals.

Aspetic surgery was conducted under general anesthesia (3–5% isoflurane in air containing 2.5% CO₂). Anesthesia level was assessed by electrocardiogram and electroencephalogram. During surgery a stainless steel base chamber was affixed to the skull over the prelunate gyrus of one hemisphere, ~2 cm lateral to the midline. After recovery surgeries, Demerol (0.5–1.0 mg/kg) or Buprenex (0.01–0.03 mg/kg) and Tylenol were administered to minimize discomfort.

At the outset of the recording session a small craniotomy (~1 cm anteroposterior and 0.5 cm mediolateral in extent) was made under isoflurane anesthesia. The prelunate gyrus was located between the lunate and superior temporal sulci, which were usually visible through the dura. Several small (1–2 mm) slits were made partly or entirely through the dura. Once surgical procedures were complete, a bolus of sufentanil citrate (5–8 mg/kg) was administered intravenously, the isoflurane was discontinued, and animals were switched to a continuous infusion of sufentanil citrate (5–8 mg·kg⁻¹·h⁻¹ iv). Anesthesia was adjusted for each animal throughout the experiment by monitoring electrocardiogram rate (90–150 beats per min) and electroencephalogram state (predominance of slow-wave activity), and by periodically testing for absence of electrocardiogram or electroencephalogram responses to toe pinch. After it was ensured that a proper anesthesia level was obtained, paralysis was induced with gallamine triethiodide (10 mg·kg⁻¹·h⁻¹ iv). Animals were respired through a tracheal cannula with a mixture containing 2.5% CO₂ in air or with air alone. After paralysis, atropine (2%) and neosynephrine drops were used to produce mydriasis and cycloplegia. The eyes were then fitted with neutral contact lenses and focused onto a tangent screen (114 cm distant) with appropriate corrective lenses and artificial pupils (4 mm diam). Foveal position was determined for each eye with a reversing-beam ophthalmoscope. To avoid binocular interactions, stimuli were presented only through whichever eye was most effective for each cell, and the other eye was occluded with an eye shutter. Eye condition and alignment were checked periodically throughout the experiment. On occasion a small bolus of Nembutal (5 mg/kg) was used to reduce burstiness of neural activity. In some cases recordings were interrupted after several days for up to 8 h while anatomic tracer injections were made in the opposite hemisphere, outside of area V4, as part of a separate experiment.

Recording techniques and data acquisition

Extracellular single-unit recordings were made with epoxy-coated tungsten microelectrodes (125 or 250 μm diam, 5 or 12 MΩ impedance, A-M Systems, Seattle, WA) advanced with a microdrive mounted on a sealed, oil-filled chamber. The microdrive was equipped with an X-Y stage that allowed accurate positioning of the electrode within the chamber. V4 penetrations were made most frequently on the prelunate gyrus, and also down both posterior and anterior banks of the prelunate gyrus (up to a maximum depth of 5 mm). Receptive field centers were located 3–12° from the fovea in the inferior visual field, and were between 1.6° and 3° diam. This range for eccentricity and receptive field size is diagnostic for area V4 on this part of the prelunate gyrus (Gattass et al. 1988). Recordings made posterior to the lunate sulcus fell in area V2, where the receptive fields were significantly smaller and more foveal. In some cases electrolytic lesions (5–10 mA for 7 s) were made to confirm electrode placement. Cells whose locations could not be determined unambiguously were excluded from this report.

Experiments were controlled by a Macintosh computer running custom software written under LabView (National Instruments). The Macintosh generated pseudorandomized stimulus sequences and passed these specifications to a Silicon Graphics Indigo workstation for stimulus generation and display. The electrode signal was amplified and single cells were isolated with the use of a window discriminator. Spikes were fed to a data acquisition board on the Macintosh and correlated with synchronization signals that were flashed on the graphics monitor and detected with a photodiode. Results could be viewed on-line to determine appropriate parameters for further tests.

Experimental procedures and stimulus construction

Data collection for each cell involved a fixed sequence of steps. First, the location and boundaries of the classical receptive field (CRF) of the cell were estimated manually with the use of a mouse-
driven receptive field plotting program that permitted complete control over the size, position, color, and motion of rectangular bars and various grating patterns. Next, a preliminary test was run to determine the optimal color for the cell, and this color was then used for all gratings shown in the main test. When possible we also ran subsidiary tests on the same cell to examine the effects of grating position, length, and width.

Visual stimuli were presented on a 19-in. gamma-corrected red-green-blue monitor driven by the graphics workstation. The monitor was placed 114 cm in front of the eyes, providing coverage of \(17 \times 14^\circ\) of the visual field at any one time. Stimuli were displayed on a neutral gray background (\(\sim 10 \text{ cd/m}^2\)) and were shown in one of eight different colors: white, dark blue, green, light blue, red, magenta, and yellow. The colors were not equated for luminance, but it is unlikely that this substantially biased our results; the test reported here used only a single, optimal color for each cell, and each of the eight colors was most effective for a comparable portion of cells.

The experimental stimuli were sinusoidally modulated luminance gratings belonging to three distinct classes: Cartesian, polar, and hyperbolic. Each class refers to the coordinate system, which determines the pattern of luminance modulation of the grating. Thus the Cartesian stimuli, illustrated in Fig. 1A, are taken from a stimulus space whose cardinal axes represent horizontal and vertical gratings; other locations in the space represent Cartesian gratings at oblique orientations. The radial component specifies the spatial frequency of the gratings; low frequencies are represented in the center of the space, and frequency increases as one moves outward. This space forms a basis, because its cardinal axes are linearly independent and span the space. The equations used to generate Cartesian gratings were

\[
L_c(x, y) = L_0 [1 + m (\cos (2\pi f_x x + \theta) + \cos (2\pi f_y y + \phi))]
\]

where \(L_0\) gives the mean luminance pedestal of the stimuli, \(m\) is a constant giving modulation depth, \(\theta\) gives phase, \(f\) is the spatial frequency of the grating, and \(\phi\) is the grating orientation.

The polar stimuli, illustrated in Fig. 1B, were taken from a stimulus space whose cardinal axes represent concentric and radial gratings; intermediate locations in the space produce spiral gratings that have either clockwise or counterclockwise polarity. The radial component specifies the frequency of the gratings. The equations governing polar gratings were

\[
L_p(x, y) = L_0 [1 + m (\cos (2\pi f_x x + \theta) + \cos (2\pi f_y y + \phi))]
\]

where \(L_0\) gives the mean luminance pedestal of the stimuli, \(m\) is a constant giving modulation depth, \(\theta\) gives phase, \(f\) is the spatial frequency of the grating, and \(\phi\) is the grating orientation.

The hyperbolic stimuli, illustrated in Fig. 1C, were taken from a stimulus space whose cardinal axes represent hyperbolic gratings differing in orientation by \(45^\circ\); intermediate orientations in the space produce hyperbolic gratings at intermediate orientations. The radial component specifies the frequency of the gratings. The equations governing hyperbolic gratings were

\[
L_h(x, y) = L_0 [1 + m (\cos (2\pi f_x x + \theta) + \cos (2\pi f_y y + \phi))]
\]

Grating frequency was always calculated in cycles per receptive field rather than cycles per degree of visual angle. Therefore each cell was studied with an identical stimulus set relative to the CRF size rather than a stimulus set that was absolutely identical for all cells.

The stimulus set used in the main test consisted of 90 gratings, including 30 Cartesian gratings, 40 polar gratings, and 20 hyperbolic gratings (see Fig. 1). Each grating was presented at two or three different spatial phases, separated by 120 or 180, and all gratings were the same size as the CRF of the cell under study. Stimulus presentation was blocked, so that each stimulus was shown once before another repetition was begun, but the order of presentation within a block was randomized. A test usually involved seven repetitions of the stimulus set and took \(~80\text{ min}\) to run. Cells that were lost before completion of the test were included in the analysis if sufficient data were available, as determined by the number of repetitions obtained and the variance of the observed responses. About 20% of the cells we encountered were isolated according to their spontaneous discharges, but could not be driven adequately during preliminary tests to warrant detailed investigation. These cells were excluded from further analysis.

Two subsidiary tests were run to examine the neural mechanism underlying non-Cartesian responses in more detail. The first examined whether non-Cartesian responses depended critically on the placement of gratings in and around the CRF. This test included six Cartesian gratings (increments of \(30^\circ\) orientation), two orthogonal polar gratings (1 concentric and 1 radial), and two orthogonal hyperbolic gratings differing in orientation by \(45^\circ\) (Fig. 2, top). All gratings were shown at one spatial frequency, optimized for the cell under study. The gratings were the same size as the CRF and were presented either centered within the CRF or at one of four positions centered on the edge of the CRF.

The second test examined whether cells that gave larger responses to non-Cartesian than to Cartesian gratings were actually selective for Cartesian gratings shorter or narrower than the CRF. It included three Cartesian grating orientations (increments of \(60^\circ\) orientation) that were truncated parallel and/or orthogonal to the direction of sinusoidal modulation (Fig. 2, bottom), so that they were between 0.125 and 1.0 times the size of the CRF. The test also included a subset of 16 non-Cartesian gratings consisting of 4 hyperbolic gratings whose orientation differed by 22.5, 4 radial gratings, 4 concentric gratings, and 4 left-handed spiral gratings. Cartesian and hyperbolic gratings were shown at one spatial frequency, optimized for the cell under study. Polar gratings spanned a concentric and radial frequency range bracketing the optimal Cartesian frequency of the cell.

Data analysis

RANDOMIZATION STATISTICS. Throughout this study we used randomization tests rather than classical statistical tests to estimate significance and bias. In a randomization test, the distribution of the statistic under the null hypothesis is estimated directly from the data set, by computing the test statistic on a large number of random permutations of the original data. Significance is determined by comparing the test statistic actually obtained in the experiment with the distribution of the randomized statistic. In this report significance indicates a probability due to chance of <5% (\(P < 0.05\)).

Randomization tests are related to bootstrapping procedures (Manly 1991), but randomization is used to estimate significance rather than a distribution parameter of the data. These tests have rarely been used in neurophysiology (Edgington and Bland 1993), although they are commonly used in other areas of biology (see Manly 1990). They may be better suited to neurophysiological data than conventional statistical tests because they do not rely so heavily on assumptions of normality and independence and they permit construction of new tests in situations where no standard method is available.
EVOKE RESPONSE RATES. Evoked responses were calculated for each trial by subtracting the firing rate obtained during the 500 ms preceding the stimulus (the before period) from the rate obtained during the 500 + 16 ms period that the stimulus was on the screen (the during period). The significance of evoked responses for each cell was assessed by comparing the difference in before-period and during-period response rates with the distribution of differences obtained from 1,000 random assignments of the responses to "before" and "during" categories. All 103 area V4 cells reported here showed significant stimulus-evoked responses by this measure.

PHASE. Selectivity for grating phase was examined by comparing the responses obtained with different phases of the most effective grating in the stimulus set. The test statistic was equivalent to a one factor F statistic (analysis of variance) computed across phase, and significance was determined by comparing this value with the distribution of values obtained from 1,000 random permutations of the data from the most effective grating.

PEAK RESPONSE ANALYSIS. The analysis of peak responses was based entirely on the data obtained with the most effective Cartesian, polar, and hyperbolic gratings for each cell. The primary
goal was to identify highly selective cells in which the response to one stimulus class was significantly larger than the response to the remaining two classes. The test statistic was the difference between the peak response obtained with the most effective stimulus and that obtained with the next most effective stimulus from the remaining two classes. (Responses to the least effective stimulus class were not included.) Significance was determined by comparing this value with the distribution of values obtained from 1,000 random permutations of the data from these three stimuli. Because there were always fewer stimuli belonging to the most effective class than to the other two classes combined, this test is biased against finding significant differences (i.e., it is a conservative test). We also carried out an analogous test to identify cells in which the peak response to one stimulus class was significantly smaller than the peak response to the remaining two classes.

Selectivity for the polarity of spiral polar gratings was evaluated by comparing the peak response of each cell to the optimal spiral grating with the response of the same cell to the equivalent spiral grating of opposite polarity. The test statistic was the difference in response rates obtained with these two spiral gratings, and significance was determined by comparing this value with the distribution of values obtained from 1,000 random permutations of the data from the two spiral gratings.

**INFORMATION ESTIMATION.** The information-theoretic analysis followed the procedures developed in Optican et al. (1991) and Tovee et al. (1993) for computing the information present in a mean response rate code. Given a set of stimuli, $S$, and their associated neural responses, $R$, the average information, $I$, transmitted by any single response, $r$, about the ensemble of stimuli in the set is given by 

$$I_r = \sum_{s \in S} \sum_{r \in R} P(s,r) \log_2 \frac{P(s,r)}{P(s)P(r)}$$

where $P(s,r)$ is a matrix giving the probability of observing a particular response, $r$, for each stimulus, $s$. We used the following procedure to estimate the probability matrix $P(s,r)$ (Optican et al. 1991; Tovee et al. 1993): each stimulus, $s$, is associated with a set of response rates, $R$, evoked over a series of repeated presentations. To remove the DC bias, the mean response rate obtained across all stimuli is subtracted from $R$. Next, the set of responses is converted to a probability density function; the set of responses is histogrammed within the range $+m$, where $m = \max(|R|)$, and smoothed by convolution with a Gaussian filter whose SD is equal to the SD of $R$, and the result is normalized to unity. The probability density function is then discretized into $D$ bins across the data range $\pm m$. The number of bins must be suitably large; information rates typically asymptote when $D > 10-12$, and we chose $D = 16$. The area within each bin is the estimate of $P(s,r)$ for a given $s$ and $r$. $P(r)$ and $P(s)$ are obtained by collapsing across the irrelevant dimension.

An important source of bias in information estimates is the small number of trial replications used in most physiology experiments (Optican et al. 1991; Tovee et al. 1993). We used a bootstrap procedure to estimate this bias as the mean of the distribution of values obtained from 500 random permutations of the relevant data. Significance was determined by comparing this value with the distribution of values obtained from 500 random permutations of the data from these two stimulus classes. (On each permutation, new corrected information estimates were computed, and the randomized difference was that between the highest and second highest value.) Because this statistic incorporates the average information transmitted about each member of a stimulus class, it is an unbiased test. We performed an analogous test to identify cells that transmitted significantly less information about one stimulus class than the remaining two classes.

**FITTING GAUSSIAN TUNING CURVES.** For each cell, we attempted to fit a parametric tuning curve model to the data corresponding to each stimulus class. The data obtained with each stimulus class were normalized to the range 0–1 and mapped into the original stimulus space (see Fig. 1). To eliminate wraparound artifacts in the data, the responses were replicated and rotated across the horizontal axis of the space. They were then fit with the following model

$$G(x,y,\phi) = \exp \left\{ -2.77 \left( \frac{R_s(x,y,\phi)^2}{\sigma^2_s} + \frac{R_s(x,y,\phi)^2}{\sigma^2_s} \right) \right\}$$

where $\phi$ gives the angle of rotation of the Gaussians. Note that this model specifies two mirror-symmetric Gaussians rather than one, to take account of the mirror symmetry of the original stimulus spaces (see Fig. 1). Parameters of the best-fit model were determined with the use of a nonlinear optimization procedure (routine nonlin, Splus, Statistica), although there were many cases where this procedure failed to fit any model. In all cases where the optimization procedure produced a valid model, the goodness of fit was significant as assessed in a randomized $\chi^2$ test (100 permutations). To ensure that we were not fitting noise, we discarded some valid fits in which the body of the tuning curve lay outside the region of stimulus space where data were collected (i.e., cases in which all the data points lay on the tails of the Gaussian model). Thus in this report the existence of an acceptable fit for a cell indicates that the data are a good match for the Gaussian model, but the absence of a fit does not imply that the data are incompatible with such a model.

**HIERARCHICAL CLUSTER ANALYSIS AND MULTIDIMENSIONAL SCALING.** Hierarchical cluster analysis and metric multidimensional scaling are multivariate statistical techniques that summarize the relationships within a data set according to some measure of similarity computed between all of the points in the set (see Jain and Dubes 1988). This summary can be used to derive a classification scheme in which cells that have a similar pattern of responses
across the stimulus set are grouped together. Because these are complementary methods that use different algorithms, findings that are consistent across the two analyses confer a high degree of confidence.

Both hierarchical cluster analysis and metric multidimensional scaling operate on a matrix reflecting the similarities between cells’ responses across the entire stimulus set. We obtained consistent and robust results by treating the responses of each cell as a 90-dimensional vector and computing the angle between the vectors associated with every pair of cells (cf. Di Lorenzo 1990). The angular similarity measure, unlike the more commonly used Euclidean metric, normalizes the data with respect to differences in overall responsiveness across cells (but see Erickson et al. 1993).

Hierarchical cluster analysis (routine hclus, Splus, S-Plus) converts a similarity matrix into a binary tree (‘‘dendrogram’’) representing a nested partitioning of the data. The algorithm iteratively groups cells or clusters of cells together. At each iteration the two most similar objects (cells or clusters) are merged, ensuring that the final tree is binary. The results of the analysis depend on the rule used to decide which objects should be merged. We obtained the best results with the complete link method, which merges two objects if their least similar members are more similar than the least similar members of any other pair of objects.

Metric multidimensional scaling (routine mds, Splus, S-Plus) projects a similarity matrix into a low-dimensional space (we used 2 dimensions) in which the distances between data points reflects their similarity: more similar points are closer together, whereas less similar ones are farther apart. The algorithm begins with an arbitrary placement of the data points and iteratively shifts points in order to reduce the mismatch (‘‘stress’’) between the interpoint distances and the original similarities. Of course, when the data space is reduced from 90 to 2 dimensions the distances between all data points rarely match the similarity matrix exactly. The algorithm ensures that this distortion is as small as possible.

Both algorithms produce a solution that reflects, as closely as possible, the similarities between all possible pairs of cells. To interpret the analyses we cut the dendrogram at a point near the top of the tree, dividing the cells into three distinct cell groups. The response profiles of different groups were compared by averaging the profiles of all cells in a group and projecting this average profile back into the stimulus space. Angles, rather than response rates, were used to compute the average so as to normalize for differences in overall response rates. Such a projection reflects the stimulus selectivity of the group as a whole, discounting differences in the absolute responsiveness of members of the cluster.

**Statistical Assessment of Anatomical Distribution of Classified Cells.** To determine objectively whether cells with similar physiological response properties were clustered anatomically, we designed an appropriate randomization test. Ideally, such a test would incorporate information about the three-dimensional locations at which different types of cells were observed in each animal. However, our data set was too sparse for such a test; it included 23 separate electrode penetrations in 10 different animals, between 1 and 10 neurons were studied on each penetration, and we could not confirm the laminar position of all cells. Therefore, we based the test only on the distribution of cell groups across electrode penetrations. The statistic was the difference between the expected and observed frequency of encountering cells belonging to each group on a particular electrode penetration, integrated across all penetrations. For this analysis the major subbranches of the binary tree produced by the hierarchical cluster analysis were treated as distinct classes, and the statistic was computed separately for each.

Given that a particular class \( i \) contains \( n \) cells, the prior probability, \( P(n_i;S) \), of encountering neurons of a particular class is simply the proportion of neurons of that class in the total population

\[
P(n_i;S) = \frac{n_i}{S}
\]

where \( S \) is the sample size (103 in this experiment). A measure of clustering, \( C_i \), for class \( i \), was calculated by integrating the difference between the class density observed on an electrode penetration and the prior, across all penetrations, \( j \)

\[
C_i = \sum_j \left( \left| \frac{P(n_i;S) - n_i/m_j}{m_j/S} \right| \right)
\]

where \( n_j \) is the number of cells of class \( n \) encountered on electrode penetration \( j \), and \( m_j \) is the the total number of cells encountered on electrode penetration \( j \). Because penetrations containing more cells should provide more reliable evidence than those containing fewer cells, the summation is weighted to take into account the number of cells observed on each penetration. To determine the statistical significance of this measure for each cell class, we compared the obtained value of \( C_i \) with the distribution of values obtained from 10,000 random permutations of the assignment of cells to penetrations.

The sensitivity of this measure, as for any statistic, depends on sample size and is greater for larger cell classes than for smaller ones (because of the increased variance accompanying small samples). Thus a small value for this statistic constitutes positive evidence that the cells were not distributed randomly across electrode penetrations. The reverse is not true, because a large value might result instead from the reduced sensitivity accompanying small sample sizes.

**Position Test.** Position invariance was assessed by comparing relative responses to the best Cartesian, polar, and hyperbolic gratings across stimulus position. First, the position that evoked the greatest response for any of the Cartesian and non-Cartesian stimuli in the set was determined. The maximum Cartesian, polar, and hyperbolic responses at that position defined the optimal vector. The test statistic was the average angular difference between the optimal vector and the vectors obtained at the other four positions. Significance was determined by comparing this value with a random distribution of angles between 0° and 90° (Monte Carlo analysis).

**Length and Width Test.** Selectivity for the length and width of Cartesian gratings was assessed by determining which Cartesian stimulus (of any length and width) produced the highest response rate. If the most effective Cartesian stimulus was a full-field grating, the test statistic was the difference between the response obtained with this stimulus and the most effective truncated Cartesian grating having the same orientation, with the use of a randomization test. If the most effective Cartesian stimulus was a truncated grating, it was compared with the full-field Cartesian grating having the same orientation. Significance was determined by comparing this value with the distribution of values from 1,000 random permutations of the data from these two stimuli.

To determine whether non-Cartesian selectivity was actually the result of a cell’s being tuned for Cartesian gratings shorter or narrower than the CRF, the response obtained with the optimal Cartesian grating (of any length and width) was compared with that obtained with the optimal non-Cartesian grating. The test statistic was the response difference, and significance was determined by comparing this value with the distribution of values from 1,000 random permutations of the data from these two stimuli. This is a conservative test, because the optimal Cartesian grating orientation and frequency was used in the test but only a small subset of polar and hyperbolic gratings were used.

**Results**

**Responses to Cartesian, polar, and hyperbolic gratings**

The 103 V4 neurons in our sample showed wide diversity in their responses to various stimuli, in the average informa-
tion transmitted about each stimulus class, and in the quality of tuning for various stimulus dimensions. Figure 3, A–C, summarizes data from three representative cells that reflect this diversity. The cell shown in Fig. 3A was highly selective for hyperbolic stimuli. Each icon represents one grating from the stimulus set, and the color of the icons corresponds to the mean response obtained with each stimulus. This cell was selective for a narrow range of hyperbolic gratings of the appropriate orientation and spatial frequency. Most stimuli produced little or no response above background (blue to green), but a few hyperbolic stimuli produced strong responses (yellow to red).

The cell shown in Fig. 3B was more broadly tuned. The strongest responses were evoked by concentric gratings. High-frequency spiral gratings also produced a large response, whereas radial gratings elicited almost no response. Both clockwise and counterclockwise spirals were effective, although there was a slight preference for clockwise spirals. Responses to Cartesian and hyperbolic stimuli, although weaker overall, were not random: the Cartesian responses were strongest for both low and high spatial frequencies and weakest for midrange spatial frequencies. Hyperbolic responses were stronger for intermediate to high spatial frequencies.

Figure 3C shows a cell that responded well to gratings from all three stimulus classes. The cell was narrowly tuned for horizontal low-frequency Cartesian gratings. It was more broadly tuned within the polar space, although it responded best to midfrequency concentric gratings. Finally, it gave moderate responses to low frequency hyperbolic gratings of any orientation.

These examples show marked differences both in the maximum responses evoked by the various stimulus classes and in the breadth of tuning within different classes. To capture the complexity of these responses, we performed several different analyses on the data set. 1) Peak responses. We first compared the peak responses obtained within each of the three stimulus classes. This comparison is a straightforward measure of the relative efficacy of different stimulus classes in driving V4 neurons. Its major disadvantage is that it does not include information about responses to suboptimal stimuli within each class. 2) Information-theoretic analysis. We estimated the amount of information each cell transmitted about the average stimulus within each class and about the stimulus set as a whole. Estimates of transmitted information reflect the distribution of responses across the stimulus set, but they reveal nothing about whether there is an orderly tuning of responses along one or another dimension. 3) Tuning curves. The responses of many neurons in visual cortex to simple dimensions such as orientation or spatial frequency are conveniently described in terms of Gaussian tuning curves along the relevant dimension. We therefore attempted to fit two-dimensional Gaussian tuning curves to the responses obtained within each stimulus class.

As a preliminary step in our analysis, we examined the sensitivity of all cells to stimulus phase, by comparing responses obtained to different spatial phases of the most effective grating in the set. Of 103 cells, 11 showed a significant difference in their responses to different phases, a proportion consistent with a previous report of phase selectivity in area V4 in which Cartesian gratings were used (Desimone and Schein 1987). Because a detailed analysis of phase sensitivity was not a central aim of this study, for other analyses we combine responses to all stimuli that are identical except for phase. This increased the signal-to-noise ratio of our data.

**Peak responses.** The analysis of peak responses used only the data obtained with the most effective Cartesian, polar, and hyperbolic gratings for each cell. Figure 4A summarizes this analysis in a form that represents the three peak response values for each cell as a vector in a three-dimensional response space. Responses have been normalized so that the vectors for all cells are of equal length; this is analogous to projecting each vector onto the surface of one sector of a sphere. The figure is oriented so that the vector representing equivalent peak responses to the three stimulus classes is projecting out of the page directly toward the viewer. The position of each symbol in the figure reflects the relative peak responses of a cell to the three stimulus classes; cells lying within the triangular zone in the interior of the plot had peak responses to Cartesian, polar, and hyperbolic stimuli that were all within a factor of 2 of one another. The size of the symbols reflects the maximum evoked response rate obtained for each cell across the entire stimulus set (the symbols have been scaled to the range of responses obtained across all 103 cells).

The diamond symbols with arrows in Fig. 4A identify the three cells illustrated in Fig. 3, A–C. The cell in Fig. 3A lies in the bottom right corner of the plot, on the hyperbolic axis. The most effective Cartesian stimulus evoked 3.6 spikes/s from this cell, whereas the most effective polar stimulus evoked 2.1 spikes/s and the most effective hyperbolic stimulus evoked 25 spikes/s. The cell illustrated in Fig. 3B lies in the left half of the plot, on the polar axis. Its most effective Cartesian, polar, and hyperbolic stimuli evoked 37, 61, and 27 spikes/s, respectively. The cell in Fig. 3C also lies along the polar axis, but falls within the zone in which peak responses for all stimulus classes were within a factor of 2 of one another. Its most effective Cartesian, polar, and hyperbolic stimuli evoked 74, 73, and 55 spikes/s, respectively.

The distribution of cells in Fig. 4A is fairly continuous and is concentrated near the center of the plot, where cells give comparable peak responses to the three stimulus classes. The population as a whole is biased toward the polar sector and away from the Cartesian sector. This is particularly noticeable for the subset of neurons that gave more than twice the peak response to one stimulus class than to another, and that therefore lie outside the zone of similar response levels in the center of the plot. There are 10 such cells in the polar sector, 8 in the hyperbolic sector, and only 2 in the Cartesian sector. (Because of response variability and stimulus sampling biases, not all these cells had statistically significant preferences for 1 stimulus class over another.) The only strong hint of clustering comes from a group of four cells in the bottom right corner of the plot (1 is hidden) that were highly selective for hyperbolic stimuli.

To illustrate the effectiveness of the different stimulus classes across the sample, we calculated the average response of the 103 cells to each of the stimuli in the set. In Fig. 3D these averages are shown using the same color code
FIG. 3. Response summary of 3 V4 cells, and the population average. Each icon represents a particular stimulus, and its color represents the mean response obtained with that stimulus over several repetitions, relative to the spontaneous firing rate of the cell. The color scale is shown at the top of each panel; dark blue refers to low evoked responses, whereas red refers to high evoked responses. A: neuron that was most selective for hyperbolic stimuli. A narrow range of hyperbolic gratings evoked a large response, but Cartesian and polar gratings produced no response above the spontaneous rate. B: neuron that was most selective for polar stimuli. Concentric gratings evoked the largest response, although the cell also responded well to a broad range of high-frequency spiral gratings. Cartesian and hyperbolic gratings were relatively ineffective. C: multimodal cell that responded approximately equally to the most effective Cartesian, polar, and hyperbolic stimuli. The cell was tuned for a narrow range of stimuli within each class. D: average response to each stimulus in the set, taken across all 103 cells in the sample. The population response is greatest to medium-frequency concentric and hyperbolic gratings.

Robert P.QUEST: POLAR, HYPERBOLIC, AND CARTESIAN GRATINGS IN V4

FIG. 4. A: relative responses for each cell obtained with the most effective Cartesian, polar, and hyperbolic stimuli. Responses are treated as a 3-dimensional vector, normalized to unit length, and projected onto the surface of a unit sphere. The figure is oriented so that the origin of the coordinate system is at the center, and the unit vector (where all responses are equal) is projecting out of the page directly toward the viewer. The positions of each symbol reflect the relative peak firing rates for each cell obtained with the most effective Cartesian, polar, and hyperbolic stimulus. Cells within the smaller triangle on the interior of the plot had peak responses to Cartesian, polar, and hyperbolic stimuli that were within a factor of 2 of each other. The size of each symbol reflects the peak evoked response rate obtained for a cell across the entire stimulus set, scaled relative to the range of peak responses obtained across the entire sample. Black symbols: cells whose peak response to the most effective stimulus class was significantly greater than the peak obtained with the remaining 2 classes. Gray symbols: cells that were significant in the corresponding test for lesser responses to 1 class. The 3 diamonds represent the example cells illustrated in Fig. 3, as noted by the arrows. B: relative average transmitted information for each cell estimated separately for Cartesian, polar, and hyperbolic stimulus classes. The size of each point corresponds to the maximum average transmitted information for the 3 stimulus classes, scaled relative to the maximum rates obtained across the entire sample. Black symbols: cells that transmitted significantly more information about 1 stimulus class than about the remaining 2 classes. Gray symbols: cells that were significant in the corresponding test for lesser average transmitted information to 1 class. The 3 diamonds again represent the example cells illustrated in Fig. 3, as noted by the arrows.

used for the individual examples. The population bias in favor of polar stimuli is clearly visible in the figure. Overall, the average response to polar stimuli was ~20% greater than to Cartesian stimuli. (The population average response was 8.7 spikes/s to Cartesian gratings, 11.1 spikes/s to polar gratings, and 10.0 spikes/s to hyperbolic gratings.) There was considerable variability in the maximum response rates of different neurons, between 0.9 and 111 spikes/s (mean 20.1 spikes/s). However there was no correlation between overall peak response rates (given by the size of the symbols in Fig. 4) and stimulus selectivity as denoted by the distance of cells from the origin of Fig. 4A (r = 0.06). Large peak response rates were about as common in more selective cells as in broadly tuned or nonselective cells.

To identify cells that were significantly more responsive to one stimulus class than to the other two, we compared the peak responses obtained with the most effective stimulus with those obtained with the next most effective stimulus (from either of the other two stimulus classes). No cells preferred Cartesian over polar or hyperbolic gratings, but seven cells preferred polar gratings (including the cell shown in Fig. 3B) and six preferred hyperbolic gratings (including the cell shown in Fig. 3A). These 13 cells are highlighted in black in Fig. 4A. Because this test is conservative, these estimates represent a lower bound on the proportion of cells with significant peak response differences.

Gray symbols in Fig. 4A indicate the five cells that were significantly less responsive to one stimulus class than to the other two. One cell gave a significantly smaller peak response to the most effective Cartesian grating than to the most effective polar or hyperbolic gratings, two gave a significantly smaller response to polar gratings, and one gave a significantly smaller response to hyperbolic gratings.

We evaluated selectivity for the polarity of spiral polar gratings by comparing the peak response of each cell to the optimal spiral grating with the response of the same cell to a similar spiral grating with opposite polarity. Only 3 of 103 cells gave a significantly different response to the two spiral directions.

AVERAGE INFORMATION TRANSMISSION. The information estimates were based on the distribution of response rates obtained for all stimuli in each of the three stimulus classes. Figure 4B summarizes the estimated information transmitted...
about the average member of each stimulus class for all 103 cells in the sample, in a format similar to that used for the peak response analysis in Fig. 4A. The average transmitted information estimates obtained with each stimulus class for each cell were treated as a three-dimensional vector, normalized so that the vectors for all cells had equal length, and projected onto the surface of one sector of a sphere. Thus the positions of the symbols in Fig. 4B reflect the relative information transmitted about the average member of each of the three stimulus classes. Cells lying within the triangular zone in the interior of the plot transmitted similar amounts of information (within a factor of 2) about each of the three stimulus classes. The size of the symbols reflects the estimated information transmitted about the average stimulus within the set taken as a whole (scaled to the range of transmitted information estimated across all 103 cells).

The diamond symbols with arrows in Fig. 4B identify the three cells illustrated in Fig. 3. The cell in Fig. 3A lies in the bottom right corner of the plot, along the hyperbolic axis. Information analysis showed that this cell transmitted 0.05 bits about the average Cartesian stimulus, 0.03 bits for the average polar stimulus, and 0.2 bits for the average hyperbolic stimulus. Thus the analyses of peak evoked responses and transmitted information for this cell are consistent. The cell transmitted 0.04 bits about the average stimulus in the whole set, reflecting the fact that it gave little or no response to the great majority of stimuli in the set. The cell in Fig. 3B lies in the bottom left corner of the plot, near the polar axis. It transmitted 0.11, 0.37, and 0.04 bits for the average Cartesian, polar, and hyperbolic stimulus, respectively, and 0.25 bits for the average stimulus in the set as a whole. For this cell the information estimate is slightly more biased in favor of polar stimuli than was the analysis of peak responses. The cell in Fig. 3C lies nearer the top of the plot. It transmitted 0.15, 0.06, and 0.07 bits for the average Cartesian, polar, and hyperbolic stimulus, respectively, and 0.16 bits for the average stimulus in the set as a whole. For this cell the information estimate is strongly biased in favor of Cartesian stimuli, but the peak responses were roughly equivalent for Cartesian and polar stimuli. The difference arises because the two analyses evaluate different aspects of the data. The comparison of peak responses focuses on differences in absolute responsiveness between the various stimulus classes and depends only on the most effective stimulus in each class, whereas the information-theoretic analysis normalizes data with respect to absolute response differences and depends on the entire set of responses.

The distribution of cells shown in Fig. 4B is more uniform than that in Fig. 4A. There were 35 cells in the polar sector, 33 in the Cartesian sector, and 35 in the hyperbolic sector. Proportions were similar for cells that gave more than twice as much information about one class than the others. Although the absolute levels of information transmitted by different cells were highly variable, all but 2 of the 103 cells transmitted a significant amount of information about the stimulus set as a whole. The range of information transmission rates for these 101 cells was highly variable, between 0.016 and 0.44 bits per stimulus (mean 0.08 bits). These rates are similar to those reported in other studies for mean rate codes (Optican and Richmond 1987; Optican et al. 1991; Richmond et al. 1987; Tovee et al. 1993). There was no correlation between the information transmitted about the average member of the entire stimulus set and the distance of cells from the center of Fig. 4B ($r = -0.05$). Large information transmission rates occurred equally often in selective and nonselective cells.

Most cells transmitted significant levels of information about each of the three stimulus classes considered individually: 99 cells transmitted significant amounts of information about the members of the Cartesian stimulus set, 101 cells about the polar set, and 98 cells about the hyperbolic set. There was a significant correlation between the peak firing rate obtained for a given cell and the information the cell transmitted about the average member of the stimulus set as a whole ($r = 0.46$, significant by a randomization test of correlation, 1,000 permutations). Thus the peak response of a cell predicts ~20% of the variance in the amount of information the cell transmits about the stimulus set as a whole.

To identify cells that transmitted significantly more information about one stimulus class than the others, we compared the average transmitted information rate for the best stimulus class with that for the second best class. Five cells were significantly more informative about Cartesian stimuli than about the other two classes, three were more informative about polar stimuli, and one was more informative about hyperbolic stimuli. These nine cells are highlighted in black in Fig. 4B.

To identify cells that transmitted significantly less information about the average member of one stimulus class than the remaining two classes, we compared the average transmitted information rate for the worst stimulus class with that for the second worst class. Five cells were significantly less informative about Cartesian stimuli than about the other two classes; six were less informative about polar stimuli, and four were less informative about hyperbolic stimuli. These 13 cells are highlighted in gray in Fig. 4B.

**GAUSSIAN TUNING CURVES WITHIN STIMULUS CLASSES.** For each cell, we attempted to fit separate Gaussian tuning curves to the data obtained with each of the three stimulus classes. Figure 5 illustrates the Gaussian fits obtained for the broadly tuned cell shown in Fig. 3C. The fit for Cartesian data is illustrated in the top row, the fit for polar data in the middle row, and the fit for hyperbolic data in the bottom row. In the left column, the responses for each stimulus class have been transformed into the coordinates of the original stimulus spaces (see Fig. 1). To preserve the inherent circular symmetry of the spaces, the responses were first replicated and rotated across the horizontal axis of the spaces, and then smoothed and coded as gray values, with darker values representing higher responses. The points overlying the plots show the locations of the experimental stimuli (and their reflections) within the original stimulus spaces. The middle column shows the best fit Gaussian tuning curves for each stimulus class, and the right column shows the residual responses not accounted for by the models. The residuals are fairly small for all three stimulus classes, indicating that the Gaussian model is a good description of this cell’s tuning curves in all three stimulus spaces.

Successful Gaussian fits were made for at least one stimulus class in 99 of 103 neurons, although only 37 permitted
fits in all three stimulus classes. The Cartesian data were fit in 58 cells, whereas the polar data were fit in 98 cells and the hyperbolic data were fit in 65 cells. The high proportion of fits achieved for polar stimuli may partly reflect the fact that the polar stimulus space was sampled densely, although denser sampling does not account for the greater incidence of hyperbolic compared with Cartesian fits.

In many cases no significant fit could be achieved. This occurred most often when the data were noisy, although in some cells the responses were quite obviously not distributed according to a unimodal Gaussian tuning curve. Although both narrow and broad tuning curves were fit in different cases, the tuning curves for some extremely narrowly tuned hyperbolic cells could not be fit, probably because the hyperbolic stimulus space was not sampled densely enough.

The likelihood of fitting tuning curves to a cell’s most effective stimulus class was not significantly greater than the likelihood of fitting tuning curves to less effective classes. Of the 24 cells whose largest response was obtained with hyperbolic gratings, hyperbolic tuning curves were fit to 15 (63%); of the 50 cells whose largest response was obtained with polar gratings, polar tuning curves were fit to 47 (94%); and of the 29 cells whose largest response was obtained with Cartesian gratings, Cartesian tuning curves were fit to 16 (55%).

The tuning curves obtained for the entire sample of cells collectively covered all three stimulus spaces. There was substantial variability in the sizes, shapes (elongation), and positions of the Gaussian tuning curves for all stimulus classes, and we did not find any differences between the parameters for tuning curves obtained with different stimulus classes across the sample. There was also no obvious relationship between receptive field size and parameters of tuning curves, which, as noted earlier, were calculated relative to receptive field diameter rather than in absolute spatial coordinates.

**Physiological classification and anatomic distribution**

In the previous analyses we compared the responses obtained for each cell across the three stimulus classes. To summarize the relationships between the responses of neurons across the entire stimulus set, we used two objective methods, hierarchical cluster analysis and metric multidimensional scaling. The summaries were then used to group cells according to their response profiles across the stimulus set.

**Hierarchical cluster analysis and metric multidimensional scaling.** The results of hierarchical cluster analysis are shown in Fig. 6, top. Each of the terminal branches of the binary tree (dendrogram) represents one cell, and the branch points of the tree reflect the similarity between the response profiles of cells or groups of cells. Branch points near the bottom of the dendrogram connect cells or clusters that are highly similar, whereas those near the top connect cells or clusters that are highly dissimilar. In this figure cells are numbered sequentially from left to right, according to their position on the dendrogram. At its highest level the tree has two distinct branches: the left branch contains only seven cells (group G1, +), and the right branch contains the rest of the sample. The right branch is further subdivided into a large branch in the middle containing the majority of cells (group G2, -), and a smaller branch on the right containing 20 cells (group G3, 0). The results of two-dimensional metric multidimensional scaling, performed with the same similarity data, are shown in Fig. 6, bottom. Cells are numbered according to their position on the dendrogram shown in Fig. 6, top to make comparison easier. The cells appear to be divided into two distinct groups. The small
group on the left contains nine cells (those belonging to group G₁₋₇, plus cells 99 and 100), and the large group on the right includes the remaining cells.

To make comparison of these groups easier, Fig. 7A shows a schematic of the dendrogram from Fig. 6, top with each of the major cell groups outlined: group G₁₋₇ is outlined in yellow, G₈₋₋₃ in red, and G₈₄₋₁₀₃ in blue. Figure 7B shows the multidimensional scaling solution from Fig. 6, bottom; cells are colored in yellow, red, or blue according to their membership in the cell groups outlined in Fig. 7A. Note that groups of cells that tend to be clustered in one analysis also tend to be clustered in the other analysis. The seven cells in group G₁₋₇ in Fig. 7A all lie within the cluster of nine cells at left in Fig. 7B. The cells in the remaining two groups in Fig. 7A lie in virtually nonoverlapping regions of Fig. 7B.

The response profiles of the cells within each of these three groups are summarized in Fig. 7, C–E. The angular response vectors of all the cells in each group were averaged together and projected back into the stimulus space, and are represented as color maps (cf. Fig. 3). Thus these panels reflect the stimulus selectivity of the average member of each group, discounting differences in absolute response rates. Figure 7C summarizes the responses of group G₁₋₇ (yellow). This group was highly selective for mid- to low-frequency hyperbolic gratings; Cartesian and polar gratings were relatively ineffective stimuli for these cells. Although it included only seven cells, this was the most selective group of cells in the entire sample.

Figure 7D summarizes the responses of group G₈₋₋₃ (red). A wide range of polar gratings was effective stimuli for this group, with the exception of high-frequency radial gratings. The group also displayed secondary selectivity for low-frequency Cartesian and hyperbolic gratings. Subgroups of cells (lower branches on the dendrogram) were tuned for different parts of the polar stimulus space, but there was no subgroup whose primary selectivity was for Cartesian or hyperbolic stimuli. Although this group constitutes the majority of the cells in our sample, the response profile is much narrower than the total population response shown in Fig. 3D.
Figure 7E summarizes the responses of group G84-103 (blue). The group was selective for high-frequency radial gratings, although many patterns were moderately effective. Note that this group contains two cells that lie at the far left of the multidimensional scaling space, near the group of cells selective for hyperbolic gratings. Inspection of the response profiles of these two cells revealed a strong hyperbolic response, suggesting that these cells were misclassified by the hierarchical cluster analysis but were classified more accurately by the multidimensional scaling solution.

To confirm the reliability of the results, both the cluster analysis and multidimensional scaling were repeated several times with half of the stimuli or half of the cells randomly removed. Removal of half of the stimuli produced little change in the classifications, suggesting that the stimulus spaces were highly oversampled. Removal of half of the cells caused some changes in classification, but the results were qualitatively similar to those obtained with all cells.

Considering that the two classification methods used here rely on fundamentally different algorithms, their correspondence is remarkably good. Both procedures emphasize the distinction between the cells selective for hyperbolic gratings and the rest of the cells in the sample. The substructure of the polar-prefering classification is also represented consistently; cells lying on particular branches of the hierarchical cluster analysis tree are segregated in the metric multidimensional scaling solution space as well.

ANATOMIC CLUSTERING. To investigate whether the cell classes identified physiologically were segregated anatomically as well, we analyzed the statistical distribution of cells across electrode penetrations. This analysis compares the observed distribution of identified cells across electrode penetrations with the frequencies that would be expected under the null hypothesis of a random distribution (see METHODS). The results are shown in Fig. 8. The numbers above branch points on the dendrogram represent the estimated probability of obtaining the observed amount of anatomic clustering by chance, for the cells on the corresponding branch. This value was 2% for the large polar group (G1-1) illustrated in Fig. 7D, indicating that anatomic clustering of this group was nonrandom. The value was 9% for the hyperbolic group (G1-1) shown in Fig. 7C, and 14% for the radial group (G84-103) shown in Fig. 7E; these values suggest a nonrandom anatomic distribution for these groups as well. As noted in the METHODS, these clustering probability estimates are sensitive to sample size, and this is reflected in the probabilities shown in Fig. 8; large groups of cells generally have a smaller probability of being randomly distributed than do smaller groups of cells.

SUBSIDIARY TESTS

POSITION TEST. The position test was designed to determine whether cells’ response preferences for different stimulus types were affected by the position of the stimuli with respect to the CRF. In essence, this is a question about the position invariance of V4 response profiles. The test compared responses to Cartesian, concentric, radial, and orthogonal hyperbolic gratings at the preferred spatial frequency of the cell (see Fig. 2A). For some cells the pattern of responses across stimulus types was strikingly stable regardless of stimulus position. For example, Fig. 9A shows the responses in the position test of the cell shown in Fig. 2B. Although the peak response rate varies with stimulus position, at each position concentric gratings produce the largest response.

The position test was run on 13 of the 103 cells in the current data set, and on 28 cells that were examined with this test as part of an earlier study (Gallant et al. 1993). The test was identical in both cases, but quantitative analysis of the earlier data has not been reported previously. Position invariance was assessed by comparing the average angle between the three-dimensional vector representing optimal Cartesian, polar, and hyperbolic stimuli shown at the most effective position, and the vectors representing the remaining four positions (see METHODS). Figure 10 displays the data for the 41 cells analyzed in a format similar to that shown previously in Fig. 3. Each cell is represented by four rays that connect the vector associated with the most effective stimulus position to those associated with the remaining four stimulus positions. These rays tend to be compact, indicating that cells generally have a high degree of position invariance. In fact the statistical test showed that all 41 cells had a statistically significant degree of position invariance.

LENGTH AND WIDTH TEST. The length and width test was designed to determine whether non-Cartesian response preferences might be produced by the length or width selectivity reported previously for V4 cells (Desimone and Sreen 1987). Responses to truncated Cartesian gratings that were shorter or narrower than the CRF were compared with those obtained with Cartesian gratings that filled the CRF and to a subset of the non-Cartesian gratings used in the main test (see Fig. 2B).

Figure 9B illustrates the results obtained in the length and width test for one cell that had a significant preference for radial polar gratings in the main test. This cell gave insignificant responses to all Cartesian gratings (both truncated and full field), including long thin bars that one might suspect would drive a radial-prefering cell. Thus conventional selectivity for length and width does not account for the non-Cartesian selectivity of this cell.

Of the 29 cells examined in this test, the main test identified 4 as significantly non-Cartesian (either polar or hyperbolic). In the length and width test, these four cells gave statistically indistinguishable responses to the most effective full-field Cartesian and truncated gratings. Three cells gave a significantly larger response to a non-Cartesian grating than to any full-field or truncated Cartesian grating. One cell gave a significantly larger response to full-field Cartesian than to non-Cartesian gratings, but this cell’s preferred non-Cartesian grating was not included in the stimulus subset used in the length and width test. This analysis indicates that length and width selectivity generally does not account for the significant non-Cartesian response selectivity observed in the main test.

On the other hand, we found some cells that were tuned for the length of Cartesian gratings and others that were tuned for grating width. Figure 9C summarizes the responses of one cell that gave a significantly larger response to extremely short Cartesian gratings than to Cartesian gratings that filled the CRF. In the main test this cell had a nonsignificant preference for polar over Cartesian and hyperbolic...
Cluster Average: Hyperbolic

Cluster Average: Polar (Concentric)

Cluster Average: Polar (Radial)
The branch points of the dendrogram (Fig. 6A) have been labeled with the probabilities that the cells on each branch were not distributed randomly across electrode penetrations. The probability estimate is more reliable for larger than smaller groups. The large polar cluster was highly unlikely to be distributed randomly across electrode penetrations. The hyperbolic cluster was also unlikely to be randomly distributed, whereas the radial cluster has a marginal probability of being nonrandomly distributed.

Interestingly, the cell was not particularly sensitive to the orientation of its preferred short Cartesian gratings. The preference for short, wide gratings (i.e., a periodically interrupted bar) over long, narrow gratings (i.e., thin solid bars) of any orientation cannot be accounted for by conventional tuning for stimulus length and width of the type reported by Desimone and Schein (1987).

Of 29 cells tested, 5 gave a significantly larger response to the most effective truncated Cartesian grating than to the corresponding full-field Cartesian grating of the same orientation, confirming the earlier report of length and width selectivity in a subset of V4 cells. In contrast, none of the cells gave a significantly larger response to a full-field Cartesian grating than to the most effective truncated grating.

**DISCUSSION**

In this study we systematically examined the responses of V4 neurons to three classes of geometric patterns, and we carried out several complementary analyses on the responses of individual cells and on the population as a whole. The results provide several insights about the representation of form in V4. Our discussion will be organized around four general issues. 1) **Specificity**: to what degree are V4 neurons uniquely responsive to or selective for these various stimulus classes? Can cells be grouped into distinct response categories according to their response profiles? 2) **Non-Cartesian responses**: to what extent can non-Cartesian responses be accounted for in terms of responses along simpler dimensions? How might non Cartesian responses be contructed...
FIG. 9  

A: position test results for the cell shown in Fig. 3B. Although overall response level is modulated by the precise position of the pattern with the CRF, at each position the concentric grating generates a larger response than any other stimulus. B: length and width test results for a cell that, in the main test, gave a significantly larger response to radial gratings than to hyperbolic or Cartesian gratings. This response preference was preserved in the length-width test; the response obtained with radial gratings was significantly larger than that obtained to any full-field or truncated Cartesian grating. C: length and width test results for a cell that gave significantly larger responses to truncated Cartesian gratings substantially shorter than the CRF. The orientation of the truncated gratings was not important. According to the main test this was a nondifferential cell that gave the largest response to polar gratings.
from cells with simpler receptive field properties? 3) V4 versus MST: how does non-Cartesian pattern selectivity in area V4 compare with non-Cartesian motion selectivity reported previously in area MST? 4) Functional implications: what is the functional significance of the non-Cartesian selectivity encountered in this study? How does it relate to form processing in other visual areas?

Specificity

V4 cells displayed a wide range of selectivity for the three stimulus classes. About 70% of cells gave comparable responses (within a factor of 2) to the most effective member of each class. However, cells did not simply respond to every stimulus; most cells had clear tuning profiles within each stimulus space, and virtually all cells conveyed significant information about each class. The results suggest that the population of V4 cells processes information about both Cartesian and non-Cartesian stimulus classes.

Although the majority of cells responded to all three stimulus classes, across the population non-Cartesian stimuli evoked a 20% larger response than did Cartesian stimuli. A minority of cells (17%) gave significantly larger responses to the most effective non-Cartesian grating than to any Cartesian grating. This percentage is almost the same as we reported previously (16%) (Gallant et al. 1993). In contrast, there were no cells that gave significantly larger responses to full-field Cartesian gratings. In our earlier report 8% of cells fit into this category, but we suspect that most of these cells would have been classified differently had they been tested with the larger set of non-Cartesian stimuli used in the present study. Only a few cells were extremely selective, responding to at most a few highly similar stimuli. Most of these were selective for hyperbolic gratings. Thus there is a general bias in V4 in favor of non-Cartesian over full-field Cartesian stimuli, and a subset of cells is highly selective for non-Cartesian gratings.

The broad bias in favor of non-Cartesian patterns is consistent with the results of Kobatake and Tanaka (1994), who used a completely different method to estimate the optimal stimuli for cells in several extrastriate areas. They began with a broad array of three dimensional objects and tried to determine which two-dimensional features of particular objects were most effective in driving each cell. The two most common effective two-dimensional patterns reported in their summary figure for area V4 were an “X” shape, analogous to the hyperbolic stimulus class, and a bullseye, star, or circular shape, analogous to the polar class. Kobatake and Tanaka also report peak responses for other complex stimuli that do not have analogues in the present study, suggesting, as one might expect, that Cartesian and non-Cartesian stimuli represent only a subset of the response preferences existing in area V4. Further experiments are likely to reveal additional dimensions useful for describing V4 response properties.

Cell Classes. It has often proven difficult to determine whether cells in visual cortex that differ by some neurophysiologic measure represent distinct categories or whether they instead represent points on a continuum. By some measures, such as the relative responses to the most effective stimulus within each class and the information transmitted about each class (cf. Fig. 3), we found cells to be distributed along a continuum. The cluster and multidimensional scaling analyses generally confirmed this for the majority of the sample, which appears to form a broad continuum whose endpoints represent selectivity for concentric and radial polar gratings, respectively. However, there appears to be a subset of cells that are highly selective for hyperbolic gratings, and these might form a distinct physiological class. Interestingly, there is no cluster centered on Cartesian gratings, although such selectivity is common in earlier visual areas (De Valois and De Valois 1990).

Even if cells with different non-Cartesian response properties do not form physiologically distinct classes, our analysis suggests that they are anatopically clustered in the sense that neighboring cells are likely to have similar tuning characteristics. Such clustering is consistent with that observed at both earlier (Hubel and Wiesel 1977) and later (Fujita and Tanaka 1992) stages of the visual processing hierarchy, and with other evidence that area V4 contains compartments or modules (DeYoe et al. 1994; Zeki 1983b). Anatomic clustering of cells with similar non-Cartesian response properties suggests that these stimuli reflect important aspects of the representation of form in this area.

Non-Cartesian responses

Although many V4 neurons showed orderly tuning profiles for more than one stimulus class, the responses obtained with any given class generally did not predict or account for the responses obtained with the other two classes. Although a significant subset of cells was tuned for the length and...
width of Cartesian gratings, this tuning also failed to account for non-Cartesian response profiles. The lack of predictability of non-Cartesian from Cartesian responses in area V4 suggests the existence of mechanisms that pool across orientation and frequency to shape a cell’s responses to a particular stimulus class.

In principle, selectivity for concentric gratings could arise if a V4 cell received convergent excitatory input from vertically oriented Cartesian cells along the left and right flank of the CRF and from horizontally oriented cells along the top and bottom of the CRF. Selectivity for radial gratings could analogously arise from a complementary arrangement of inputs from horizontally oriented cells along the left and right flanks of the CRF and from vertically oriented cells along the top and bottom. However, this mechanism would generate receptive fields that are sensitive to the exact position of stimuli with respect to the CRF; a shift in stimulus position of one half CRF would completely change the tuning characteristics of the cell.

The position invariance we observed in all V4 cells tested suggests that non-Cartesian selectivity involves more complex, probably nonlinear interactions. For example, V4 receptive fields might be generated from a set of position-sensitive receptive field subunits like those described above, with similar preferences but different center positions. Such subunits might be represented explicitly by cells in lower visual areas (e.g., area V2), analogous to the hypothesized construction of complex cells from simple cells in area V1 (Hubel and Wiesel 1962). Alternatively, the subunits might be implemented directly within area V4 by nonlinear interactions on dendrites, as has been suggested for position-independent rotation-selective cells in area MSTd (Tanaka et al. 1989; Zhang et al. 1993).

V4 versus MST

Our use of non-Cartesian stimuli was motivated in part by analogy with area MSTd selectivity for components of non-Cartesian optical flow patterns. There are many similarities between the Cartesian and non-Cartesian responses we observed in area V4 and those reported previously for area MSTd. Most obviously, both areas contain cells that give robust responses to Cartesian and non-Cartesian stimuli: translation in the image plane is analogous to motion of the Cartesian gratings, expansion is analogous to motion of concentric gratings, and rotation is analogous to motion of radial gratings. Cells in both areas can also give tuned responses to stimuli from more than one class. In the case of area V4, the majority of cells have similar responses (within a factor of 2) to patterns from all three stimulus classes. This issue has not been examined as exhaustively in area MSTd, but many MSTd cells give similar responses to the most effective Cartesian and polar optical flow patterns (Duffy and Wurtz 1991a,b; Graziano et al. 1994).

The population of V4 cells is biased toward concentric and spiral selectivity and away from radial selectivity (see Fig. 2D), although a significant subset of cells is selective for radial stimuli. A similar bias has been reported in area MSTd, where more cells are selective for expansion than for rotation (Graziano et al. 1994; Tanaka and Saito 1989). Moving non-Cartesian grating patterns were actually used as stimuli in one study of area MSTd, and cells showed similar patterns of responses for gratings and for random dot displays with similar motion properties (Tanaka and Saito 1989).

V4 cells tend to maintain their stimulus selectivity even when stimulus position is shifted by as much as half the diameter of the CRF. The existence of substantial position invariance suggests that non-Cartesian form selectivity is based on the configuration of the entire stimulus near the CRF, rather than the pattern of stimulation at each local subregion. MSTd cells also show a high degree of position invariance, although they have much larger receptive fields (Duffy and Wurtz 1991b; Saito et al. 1986). They can also be driven by appropriate patterns of motion that are significantly smaller than the CRF. Zhang et al. (1993) have shown that these properties can arise in a neural network model of MSTd that receives input from many fixed, local receptive fields selective for a single direction of (Cartesian) motion. As noted above, a similar mechanism may underlie position invariance in area V4 as well.

One major difference between cells in the two areas is found in their responses to spirals of opposite polarity (clockwise vs. counterclockwise). V4 cells that respond to spiral patterns generally show only a slight bias in their preference for one spiral polarity or the other. In contrast, many MSTd cells are highly selective for only one direction of spiral motion (Duffy and Wurtz 1991a,b; Graziano et al. 1994; Tanaka et al. 1989).

The similarities in responses of cells in areas V4 and MST to non-Cartesian patterns are particularly notable considering differences in their presumed visual functions (Ungerleider and Mishkin 1982; Van Essen and Gallant 1994). Area MSTd lies in the dorsal processing stream, which is thought to be involved in the analysis of spatial relationships. The response preferences of area MSTd neurons are therefore explicable in terms of the mathematics of optical flow arising from an animal’s motion through the environment (Duffy and Wurtz 1991a,b; Graziano et al. 1994; Orban et al. 1992; Saito et al. 1986; Tanaka and Saito 1989). Area V4 lies in the ventral processing stream, which is thought to be involved in the analysis of the structure of objects. The precise function of non-Cartesian responses in form vision is not yet clear, but the similarities with area MSTd suggest that some common computational principles underlie many aspects of form and motion processing.

Functional implications

FEATURE DETECTORS VERSUS FILTERS. Feature detectors and filters represent two distinct perspectives for interpreting the functional significance of visual neurons. In essence, a feature detector signals the presence or absence of a particular feature (or conjunction of features) within the cell’s receptive field (Barlow 1972; Tanaka 1993), whereas a filter produces a graded estimate of the degree to which the image composition within the receptive field matches the sensitivity profile of the cell (De Valois and De Valois 1990). The existence of a few cells in our study that were highly selective for a narrow range of non-Cartesian stimuli makes it tempting to wonder whether they might indeed represent feature detectors. However, the majority of cells we encoun-
tered gave vigorous responses to a wide range of stimuli, including gratings within more than one stimulus class. Of course, these cells might have given even larger responses to patterns not included in our stimulus set (cf. Kobatake and Tanaka 1994), but they still would convey a large amount of information about a wide range of suboptimal stimuli. We therefore consider it more appropriate to consider most V4 neurons to be nonlinear filters that are broadly tuned along several dimensions relevant to stimulus form and color.

**IMAGE FEATURES.** Particular classes of features may be critically important for image segmentation and the perception of three-dimensional structure. Curved edge boundaries are a common image feature and may be important for segmenting natural images (Biederman 1987). Although psychophysical studies have shown that humans are very sensitive to curvature (Foster et al. 1993; Whitaker et al. 1993), we know of no systematic neurophysiological studies of this aspect of image structure. V4 cells that respond to polar non-Cartesian grating patterns could mediate the perception of curvature in much the same way that V1 cells selective for Cartesian gratings are thought to mediate our perception of straight edges. Similarly, line crossings and terminators are an important source of information about image structure and may facilitate image segmentation and identification (Julesz 1981); the subset of V4 neurons that are highly selective for hyperbolic non-Cartesian gratings patterns would, we suspect, respond well to line crossings. The correlation between non-Cartesian responses and certain image features suggests that these cells could form part of an intermediate representation of visual structure. If so, then other important aspects of image structure should be represented in this area as well.

**TEXTURE.** Behavioral evidence suggests that the primate visual system contains mechanisms that are sensitive to texture gradients (Glass and Switkes 1976; Nothdurft 1992). These mechanisms may be critical for many intermediate visual functions, such as segmentation and grouping (Knirerim and Van Essen 1992; Nothdurft 1992) and estimation of shape from shading (Erens et al. 1993; Pentland 1989). Some of the properties that we observed in V4 neurons suggest that they are involved in the representation of texture gradients. V4 cells are broadly tuned for both Cartesian and non-Cartesian patterns, and are relatively insensitive to grating phase, position, and (for spiral gratings) polarity. These properties are consistent with cells that estimate texture energy or power with respect to a particular gradient direction rather than responding to a specific constellation of image features.

**NON-CARTESIAN FILTERING IN COMPUTATIONAL AND THEORETICAL VISION.** Filters similar to our non-Cartesian patterns have several desirable computational properties. Perona (1991) showed that a finite set of filters closely related to the polar patterns used here can represent a continuous range of signals without discretization error. Koenderink and van Doorn (1990) predicted the existence of receptive field profiles closely related to the three stimulus classes used here on the basis of a few simple design constraints such as scale invariance and the absence of spurious resolution in the representation. Li and Atick (1994) proposed a specific representation that might permit efficient image segmentation by preserving information about spatial relationships in an image, and their filters are highly correlated with hyperbolic patterns.

Non-Cartesian filters have also been proposed as key components in a system meant to normalize images with respect to transformations arising from differences in viewpoint. Views of an object from different vantage points are related, up to a point, by several simple affine transformations, including translation, rotation, dilation, and shear. Although these relationships break down eventually because of self-occlusion of solid objects, they may form an important component of the process that maps visual input into memory (Pitts and McCulloch 1947).

One specific proposal for how the visual system may compensate for changes in vantage point is based on Lie group theory (Dodwell 1983, Dodwell and Humphrey 1990, Hoffman 1966; Hoffman and Dodwell 1985). According to this hypothesis the visual system converts all images into a normalized representation in which the effects of all affine transformations have been minimized. The elements of such a system, known as Lie kernels, are closely related to the Cartesian and non-Cartesian patterns used here. Thus our discovery of non-Cartesian selectivity is superficially consistent with this hypothesis. Many of the specifics of this proposal are still in development, however (Caelli 1976; Eagleson 1992, Ferraro and Caelli 1988), and there are no neurobiologically oriented models available. It is therefore difficult to know how well a model based on this theory would account for our data.

**Concluding remarks**

Non-Cartesian stimuli are a valuable addition to the repertoire of patterns available for characterizing cells in the visual system, not because cells are uniquely tuned to these patterns, but rather because the patterns elicit responses that were not predicted by the responses obtained to Cartesian gratings and lines used in earlier studies. In addition, non-Cartesian stimuli are a principled, mathematically tractable set of patterns that have a strong theoretical basis and are related to patterns that have proven useful in studies of other visual areas. Still, these patterns represent just one of many systematic stimulus sets that might be used to probe the response properties of cells in extratemporal visual cortex. Other sets, chosen for different reasons, are likely to reveal further aspects of shape processing in area V4.

The complex response profiles we observed in area V4 are likely to reflect the competing demands on an intermediate visual area that must faithfully represent a wide range of stimulus characteristics while enabling the system to preferentially detect and process behaviorally relevant structure within a visual scene. Thus V4 transmits significant information about many different classes of stimuli and has recognizable tuning curves within each of these classes. At the same time, certain portions of the multidimensional stimulus space produce larger responses in the population of V4 cells, and may therefore gain preferential access to higher-order mechanisms involved in attention and object recognition.

We thank J. Braun for help in the early phases of these experiments, C. Anderson, B. Olshausen, and J. Mazer for advice and discussion of the project, and M. Walsh and C. Adams for technical assistance.

This work was supported by National Eye Institute Grant EY-02091.
REFERENCES


