We compared aspects of shape representation in extrastriate visual areas V2 and V4, which are both implicated in shape processing and belong to different hierarchical levels. We recorded responses of cells in awake, fixating monkeys to matched sets of contour and grating stimuli of low or intermediate complexity. These included simple stimuli (bars and sinusoids) and more complex stimuli (angles, intersections, arcs, and non-Cartesian gratings), all scaled to receptive field size. The responses of cells within each area were substantially modulated by each shape characteristic tested, with substantial overlap between areas by many response measures. Our analyses revealed many clear and reliable differences between areas in terms of the effectiveness of, and response modulation by, various shape characteristics. Grating stimuli were on average more effective than contour stimuli in V2 and V4, but the difference was more pronounced in V4. As a population, V4 showed greater response modulation by some shape characteristics (including simple shape characteristics) and V2 showed greater response modulation by many others (including complex shape characteristics). Recordings from area V1 demonstrated complex shape selectivity in some cells and relatively modest population differences in comparison with V2. Altogether, the representation of 2-dimensional shape characteristics revealed by this analysis varies substantially among the 3 areas. But surprisingly, the differences revealed by our analyses, individually or collectively, do not parallel the stepwise organization of the anatomical hierarchy. Commonalities of visual shape representation across hierarchical levels may reflect the replication of neural circuits used in generating complex shape representations at multiple spatial scales.

Keywords: contours, extrastriate cortex, hierarchical processing, non-Cartesian gratings, receiver operating characteristic (ROC) analysis

Introduction

In the macaque monkey, the most intensively studied non-human primate, visual cortex contains dozens of distinct areas that are interconnected by hundreds of distinct pathways (Felleman and Van Essen 1991; Van Essen 2004). Based on objective anatomical criteria, most notably the laminar patterns of forward and feedback connections, these areas can be arranged as a distributed hierarchy that by some estimates includes 10 cortical processing stages (Maunsell and Van Essen 1983; Felleman and Van Essen 1991). The connections among areas are usually reciprocal, and they subserve extensive 2-way flow of visual information within and across hierarchical levels.

Given the anatomical evidence for hierarchical organization, it is important to ascertain whether the functional properties of neurons in different visual areas are closely related to their hierarchical level. The clearest correlation is that receptive field sizes (after scaling for eccentricity) increase progressively through several levels of the hierarchy, in qualitative agreement with the anatomical convergence of forward connections (Maunsell and Van Essen 1983; Felleman and Van Essen 1991). For receptive field characteristics related to the functional specializations of each area, a number of complex receptive field characteristics have been described in higher levels that appear to be absent or less prevalent at lower levels (Felleman and Van Essen 1991; Ungerleider and Haxby 1994; Van Essen and Gallant 1994; see Discussion). However, most of the comparisons across studies involve divergent sets of stimuli and experimental paradigms. For example, selectivity for non-Cartesian gratings has been reported in both V4 (Gallant and others 1993) and V2 (Hegdè and Van Essen 2000), but whether there are significant areal differences in the patterns of selectivity has been unclear owing to several methodological differences between the 2 studies.

In order to facilitate direct comparisons across visual areas, we studied the responses of individual cells in extrastriate areas V2 and V4 to matched sets of line ("contour") and grating stimuli presented using comparable experimental paradigms. Both V2 and V4 have been implicated in the processing of contour and texture characteristics sampled by our stimuli (for an overview, see Van Essen and Gallant 1994). V2 and V4 are well separated in the anatomical hierarchy—at the second and fifth level, respectively, in the Felleman and Van Essen (1991) scheme. Our experiments tested, within the constraints imposed by our stimuli and experimental parameters, whether response modulation by 2-dimensional (2D) shapes in V2 and V4 reveal distinct profiles of shape representation in the 2 populations. We find that while the 2 areas do differ in many important respects, these differences do not follow a systematic pattern that correlates with the anatomical hierarchy. Furthermore, the 2 areas were statistically indistinguishable in a number of respects. Analyses of V1 data from one monkey revealed complex shape selectivity in a subset of cells, consistent with other recent reports (Mahon and De Valois 2001; Victor and others 2005).

Materials and Methods

***Surgical and Recording Procedures***

In this study, responses of single units from areas V4 and V1 were recorded in awake, fixating macaque monkeys (Macaca mulatta) using the same set of procedures as described previously for area V2 (Hegdè and Van Essen 2000, 2003, 2004) except where noted otherwise. Prior to fixation training, each animal was implanted with a head post, scleral search coil, and an acrylic cranial patch using sterile surgical procedures. After the animal was fully trained in the fixation task, a small craniotomy (5 mm in diameter) was made through the acrylic patch over the recording site, and a recording chamber was
Stimuli

The stimulus set consisted of 48 grating stimuli and 80 contour stimuli (see Fig. 1). For the various analyses performed in this study, the stimuli were grouped into equal-sized subclasses that (with one exception, see below) shared common shape characteristics but varied in orientation, size, and/or spatial frequency. Grating stimuli were subdivided into 4 subclasses (with 12 stimuli each) that shared common shape characteristics but varied in orientation and/or spatial frequency (Fig. 1A): 1) sinusoidal gratings, 2) hyperbolic gratings, 3) concentric-like polar gratings, and 4) radial-like polar gratings. For the concentric-like gratings, the concentric frequency exceeded the radial frequency, and for the radial-like gratings, the radial frequency exceeded the concentric frequency. Of the 4 polar gratings in which the concentric frequency and the radial frequency were equal, the pair with the 2 highest frequencies was assigned to the concentric-like grating subclass, and the remaining pair was assigned to the radial-like grating subclass.

The contour stimuli were grouped into 10 subclasses, each containing 8 stimuli varying in orientation and size (and also in shape in the case of subclass 4; see Fig. 1B): 1) bars, 2) 3-way intersections ("tristars"), 3) crosses, 4) 5 and 6-armed stars plus circles, 5) acute angles, 6) right angles, 7) obtuse angles, 8) quarter arcs, 9) semicircles, and 10) 3-quarter arcs. Each contour shape was presented in 2 sizes; the larger size of the contour stimuli as described above, was no larger than the nominal diameter of the CRF for any V4 cell in our sample. The line width of contour stimuli was set at the cell’s preferred bar width and the smaller ones at half that size.

Our stimulus set was designed to explore the selectivity for a broad (but obviously nonexhaustive) set of low and intermediate-level form cues (see Gallant and others 1993, 1996; Hedgée and Van Essen 2000, 2003, and references therein). The grating stimuli helped probe the selectivity for conventional spatial frequency and orientation (sinosoids, designated as "simple gratings"), as well as more complex textural characteristics (non-Cartesian gratings, designated as "complex gratings") which the visual system may use as basis functions for surface representation. The contour stimuli were chosen to probe the selectivity for conventional orientation (bar stimuli, designated as "simple contours"), along with selectivity for the angles, intersections, orientations, and curvature of visual contours (designated as "complex contours"), which may play an important role in image segmentation and object recognition.

V4 Recordings

Single V4 units were isolated based on both the shape and the amplitude of the waveform using a window discriminator (Bak Electronics, Germantown, MD). Cells were identified as belonging to V4 based on visual topography and the receptive field size (Van Essen and Zeki 1978;Gattass and others 1988) and by location anterior to V1 and V2.

We recorded and analyzed data from all visually responsive cells that met the isolation criteria described above. The cell’s classical receptive field (CRF) was mapped using mouse-driven bar and sinusoidal and non-Cartesian grating stimuli on the computer’s monitor, and the diameter of the CRF was estimated conservatively. The cell’s preferred bar parameters, including preferred length, width, color, and orientation, were also determined subjectively. Prior to the quantitative test for each cell, the stimulus set was reoriented according to the cell’s preferred orientation (see Fig. 3). The grating stimuli had the same diameter as the cell’s CRF. The preferred bar length, which determined the size of the contour stimuli as described above, was no larger than the nominal diameter of the CRF for any V4 cell in our sample. The line width of contour stimuli was set at the cell’s qualitatively determined preferred bar width. The grating stimuli had a spatial frequency of 2, 4, or 6 cycles per receptive field diameter and a Michelson contrast of 1.0. All stimuli were presented in the cell’s preferred color, selected during the manual mapping from a palette of 7 colors with varying luminances (red, 1.18 cd/m²; green, 5.13 cd/m²; blue, 0.51 cd/m²; aqua, 5.70 cd/m²; pink, 1.82 cd/m²; yellow, 7.02 cd/m²; white, 7.76 cd/m²; all measured using Tektronix J17 photometer). Note that these measures for customizing the stimuli for the cell under study meant that the stimuli were not physically identical from one cell to the next within or across the visual areas. We opted for this design over the comparably principled method of using physically identical stimuli for all cells in all areas because we reasoned that the differences between the areas, to the extent that they may exist, were more likely to be manifest when individual cells were driven well by the stimuli.

Each stimulus was presented in each of 3 jitter positions centered 12.5% of the CRF diameter away symmetrically around the receptive field center (mean jitter magnitude 0.6° for V4, 0.17° for V2, and 0.14° for V1). Stimuli were presented sequentially for 300 ms each with a 300 ms interstimulus interval while the animal fixated within a window of 0.5° radius for a liquid reward. Up to 6 stimuli were presented per...
trial in this fashion. Only the data from the trials throughout which the animal maintained fixation within the fixation window were used in this study. The response to each stimulus was recorded over 12 randomly interleaved repetitions. We recorded from a total of 126 V4 cells from 2 animals (65 from animal C and 61 from animal D). The receptive field eccentricities of these cells ranged from 0.6° to 12.8° (mean 5.6°), and the receptive field diameters ranged from 1.0° to 6.8° (mean 4.8°).

The V2 data were collected from a total of 196 cells from 3 animals (122 from animal A, 11 cells from animal B, and 63 cells from animal C). The receptive field eccentricities of the V2 cells ranged from 2.8° to 9.7° (mean 4.6°). Receptive field diameters ranged from 1° to 3.4° (mean 1.4°) (for details, see Hegdé and Van Essen 2000, 2003).

**V1 Recordings**

We recorded the responses of V1 cells from one animal (animal C) to 152 grating and contour stimuli (see Fig. B), which consisted of the 128 stimuli shown in Figure 1, plus 24 additional sinusoidal and bar stimuli so as to sample selectivity for spatial frequency, phase, and the orientation more densely (see Fig. B). The V1 recordings were carried out exactly as described above. Given the relatively small receptive sizes, especially for opercular V1, there was greater likelihood that the stimuli completely stimulated the nonclassical surround. We did not attempt to classify V1 cells as simple or complex cells, given the difficulties of reliable classification in alert monkey recordings.

Of the 82 V1 cells studied in this experiment, 46 were from opercular V1 (eccentricity range, 4.9°–6.0°; mean, 5.6°; CRF diameter range, 0.5°–1.2°; mean 0.8°), and the remaining were from calcarine V1 (eccentricity range, 9.9°–15.2°; mean, 11.5°; CRF diameter range, 1.0°–1.7°; mean 1.5°).

**Data Analyses**

Data analyses were carried out using programs custom written in C, S-Plus (Statsci Inc., Seattle, WA), or Matlab (The Mathworks Inc., Natick, MA). Data from all 3 areas were analyzed using the same set of procedures. V1 data were analyzed using the responses to the 128 stimuli used for all 3 areas (Fig. 1), except where indicated otherwise. For each cell from each area, the response to each stimulus was averaged from the net firing rate from 12 repetitions of the stimulus, with 4 repetitions at each jitter position (9 repetitions for 9 V4 cells and 62 V2 cells). The net firing rate was calculated for each presentation of the given stimulus by subtracting the background rate from the corresponding visually evoked response. The time windows for calculating these firing rates were customized for each cell so as to take into account the duration of its evoked responses and to yield the most reasonable estimate of its background firing rates. For V4 cells, the background firing rate was calculated using a window 50–295 ms in duration (depending on the cell) immediately preceding the stimulus onset. The evoked response was calculated using a 50 to 290 ms time window starting 100–200 ms after the stimulus onset, during which the overall firing rate of the cell (across all repetitions of all stimuli) was used to determine the firing rate beyond the evoked response. For V1 cells, the background firing rate was calculated using a window 50–295 ms in duration (depending on the cell) immediately preceding the stimulus onset. The evoked response was calculated using a 50 to 290 ms time window starting 100–200 ms after the stimulus onset, during which the overall firing rate of the cell (across all repetitions of all stimuli) remained above background levels. The onset and duration of the time windows used did not differ significantly across the 3 areas (1-way analyses of variance [ANOVAs], P > 0.05 in both cases).

We repeated many of the analyses, including all population-level analyses, using an identical set of window parameters (100 ms window for background firing rate, 0–300 ms window for evoked rates) for all cells in all areas. The results (not shown) were qualitatively similar to those obtained using the window parameters above, except that the similarities among the areas were more prominent, and the differences less prominent, using these windows. Also, the responses of cells did not show significant dependence on the cell's preferred color determined during manual mapping (multivariate analysis of variance; area x stimulus x preferred color; unequal replicates design; preferred color, P > 0.05; preferred color-stimulus interaction, P > 0.05; preferred color-area-stimulus interaction, P > 0.05).

Each cell included in this study had at least one stimulus for which the evoked response differed from the background response at P < 0.05 (2-tailed t test with Bonferroni correction for multiple comparisons). Of a total of 126 V4 cells recorded from the 2 animals, 108 cells (63 cells from animal C and 45 cells from animal D) passed this test and were included in this study, as were 180 V2 cells (108 cells from animal A, 11 cells from animal B, and 61 cells from animal C) and 81 V1 cells (all from animal C).

**Tests of Significance**

Except where noted otherwise, tests of significance were carried out using randomization (for an overview, see Manly 1991). For each test, an appropriate test statistic was first calculated using the actual neural response data. The data were then randomized in a manner appropriate for the given test, and the test statistic was recalculated using the randomized data. The randomization process was repeated 10^5 times in case of the ROC analyses and multidimensional scaling [MDS] analyses described below). The proportion of times the randomized test statistic exceeded the actual test statistic constituted the 1-tailed probability that the actual test statistic was indistinguishable from random. Except where noted otherwise, tests involving multiple comparisons were carried out using Tukey's Honestly Significant Difference test (S-Plus function "multicomp"; Crawley 2002) that adequately balances Type I error (in the present context, probability of overestimating the difference between the areas) with Type II error (i.e., probability of underestimating the differences between the areas).

**Indices of Response Modulation**

We measured the modulation, or variation, of a given cell's responses across a given subset of stimuli relative to chance fluctuations using the corresponding modulation index. To calculate the overall modulation index (OMI), we first calculated the F ratio of the cell's responses to all 128 stimuli, given by F = MS_between/MS_within, where MS_between is the stimulus-to-stimulus variance and the MS_within is the average trial-to-trial variance (see Snedecor and Cochran 1989). We next randomized the responses across the stimuli and recalculated the F ratio. The value of OMI was defined as the F ratio calculated from the actual data divided by the average F ratio from the randomization rounds. Other modulation indices (response modulation indices [RMs] and within-subclass modulation indices [WMIs] etc, see Results) were calculated in a similar manner; the only difference among the various modulation indices was the subset of stimuli used for analysis. Thus, a given modulation index provided a measure of the "information" conveyed by a given neuron about a given shape characteristic (Hegdé and Van Essen 2000, 2003, 2004). This informal measure of information is related but not identical to formal information theoretic measures such as mutual information (Hegdé and Van Essen 2000; also see Rieke and others 1997; 327–362). The RMI is preferable to more conventional metrics of response tuning in our context because the underlying shape characteristics did not always vary smoothly in our stimulus set. We have previously demonstrated the utility of this metric in evaluating shape processing in V2 (Hegdé and Van Essen 2000, 2003, 2004).

To measure the sharpness of the selectivity of a given cell's responses to a given subset of stimuli, we used the "sparseness index" (SI), given by SI = 1 –(Σr_i/n)^2 / (Σr_i/n)^2 / [1–(1/n)], where r_i is the cell's response to stimulus i (averaged across repitions) and n is the total number of stimuli in the given subset (Vinje and Gallant 2000; Friedrich and Laurent 2001). For each of the modulation indices described above, a corresponding SI was calculated.

**Comparisons across Areas**

To assess the extent to which the various indices varied between V2 and V4, we carried out conventional parametric and nonparametric tests of significance (t tests and Wilcoxon rank-sum tests, respectively). However, conventional tests of significance are problematic for our analyses, insofar as they take into account Type I error (i.e., concluding that the areas are different when they are not) but not Type II error (i.e., concluding that the areas are not different when they are). Thus, making decisions about whether the 2 areas are different based solely on the magnitude of the Type I error (or, equivalently, of P value) measured by the conventional tests incurs an indeterminate risk of committing a Type II error (see Fig. 2A). Therefore, we used as our primary approach a novel application of receiver operating characteristic (ROC) analyses (for reviews, see Green and Swets 1988; Swets 1988, 1996; Macmillan and Creelman 1991; Swets and others 2000). ROC

1102 Shape Processing in V2 versus V4 • Hegdé and Essen
analyses are preferable to conventional tests of significance for this purpose because these analyses 1) are criterion independent, 2) are not reliant upon parametric assumptions, and 3) explicitly take into account both Type I and Type II errors, in that the area above the ROC curve (ARC) is a measure of both types of error for all criterion values (see Fig. 2B). Note that the manner in which ARC measures Type I and II errors is analogous to the manner in which conventional tests of significance measure Type I error (Green and Swets 1988; Macmillan and Creelman 1991).

ROC analyses were carried out as described by Swets (1988, see Fig. 2B). Depending on the analysis, the response metric was one of the sparseness or RMSs described above. For example, to compare V2 and V4 using OMI as the response metric, 100 criterion (or “cutoff”) values, which spanned the minimum and maximum observed OMI value (from either area) in equal increments were determined. The probability of a given cell having a larger index value than the given criterion $k_i$ was determined separately for V4 and V2 cells (i.e., $P(V4) > k_i$ and $P(V2) > k_i$, respectively) for each criterion value. The ROC curve was determined by plotting the $P(V2)$ values on the y-axis against the $P(V4)$ values on the x-axis. The “area under the ROC curve” ($	ext{ARC}_{OMI}$) was determined using numerical integration (adaptive Simpson quadrature; Matlab function “quad”).

Figure 2. Comparing visual areas using ROC analysis. (A) Schematic comparison of conventional tests of significance versus ROC analysis. The probability distributions of a hypothetical response metric (e.g., selectivity for a given shape characteristic as measured by a given index) are shown for area V2 (top) and V4 (bottom). The 1-tailed alternative hypothesis to be tested is that the response metric is larger on average for V4 than for V2; the null hypothesis is that the 2 areas are indistinguishable in terms of the metric. For a given criterion (cutoff or “response bias”) value $k_i$, the probabilities of correctly judging the response metric to be larger for V4 and smaller for V2 are indicated by the light and dark gray areas, respectively. The probabilities of erroneously rejecting or accepting the null hypothesis, Type I and Type II errors ($\alpha$ and $\beta$, respectively, are denoted by the obliquely and horizontally hatched areas. For parametric tests of significance, the prespecified value of $\alpha$ determines the balance between $\alpha$ and $\beta$ for a given sample size. Note that this criterion dependency of the $\alpha: \beta$ ratio, while not critical for most neurophysiological purposes, has, in our context, the effect of arbitrarily biasing the analyses in favor of either the null or the alternative hypothesis, depending on the $k$ chosen. When the distributions are nonparametric (e.g., as shown), $\beta$ is neither predetermined by specifying $\alpha$ for any sample size nor explicitly accounted for by available nonparametric tests. ROC analysis (panel B) does not have these shortcomings. To carry out the analysis, an ROC curve (thick solid line in B) is constructed by plotting the $\alpha$ values ($\alpha$ axis) against the corresponding 1- $\beta$ values ($y$ axis) over a large number of $k$ values (4 of which are denoted here by $X$) over the range of the observed values of the response metric. Note that the area under the ROC curve (ARC, denoted by the medium gray area) is a criterion-independent and nonparametric measure that takes both $\alpha$ and $\beta$ into account; ARC is a measure of the reliability of distinguishing between areas using the given response measure. The diagonal denotes chance level performance; the area below the diagonal is 0.5. ARC values >0.5 denote that the values of the response metric are reliably higher for area V4 than for V2; ARC values <0.5 denote the opposite. Larger deviations from 0.5 generally represent greater reliability. See Materials and Methods for details.

The results of the ROC analyses were usually, but not always, consistent with those from conventional tests of significance (e.g., $t$-test, Wilcoxon rank-sum test, or Kolmogorov-Smirnov test; not shown). The results of the ROC analyses were not corrected for multiple comparisons because both types of error are subject to the same number of multiple comparisons. For ROC analyses comparing V1 with V2 (or V4), ROC analyses were plotted on the $x$ axis. Because each of these tests has strengths and weaknesses in relation to our objectives, we present results from multiple tests where appropriate.

We also carried out ROC analyses using, instead of the RMSs, spike counts as the response metric (e.g., the responses to the preferred bar in V2 vs. in V4) as in some previous studies (see e.g., Britten and others 1992). However, the ARC values were <0.5 in all such analyses (in most cases at $P < 0.05$; not shown), owing to the fact that the firing rates were generally higher in V2 than in V4 (see below).

Assessing the Absolute Response Levels and Noisiness of the Data sets

The net firing rate averaged across all stimuli and all cells was $8 (\pm 1.3$ standard error of the mean) Hz in V4, 15 ($\pm 1.1$) Hz in V2, and 13 ($\pm 1.7$) Hz in V1. These values were statistically indistinguishable across the 3 areas by a 1-way ANOVA ($P = 0.056$; not shown). Using $t$-tests pairwise between areas (uncorrected for multiple comparisons), the mean firing rates were significantly larger in V2 than in V4 ($P = 0.0015$), but indistinguishable between V1 and V2 ($P = 0.40$), and between V1 and V4 ($P = 0.22$). The maximum net firing rate (averaged across trials) across all stimuli and all cells was 131 Hz in V4, 193 Hz in V2, and 206 Hz in V1. The average maximal responses across all cells were comparable across the 3 areas: 55 Hz in V4, 44 Hz in V2, and 40 Hz in V1. Using $t$-tests pairwise between areas (uncorrected for multiple comparisons), the maximal firing rates were significantly larger in V2 than in V4 ($P = 0.0005$), but indistinguishable between V1 and V2 ($P = 0.34$), and between V1 and V4 ($P = 0.11$).

For visual cortical cells, the mean firing rate, $m$, and the noise (i.e., trial-to-trial variation of the responses), $n$, tend to have a log-linear relationship given by $\log(m) = \alpha + [\log(n)]$, where $\alpha$ and $\beta$ are the offset and slope, respectively (Dean 1981; Vogels and Orban 1991).

The noise levels in the data from the 3 areas, as measured by the correspondence $\alpha$ and $\beta$ values (V4: $k$: $k_1 = 1.12$, $\alpha = 0.78$; V2: $k_1 = 1.21$, $\alpha = 0.58$; V1: $k_1 = 1.09$, $\alpha = 0.92); data not shown), were indistinguishable across the 3 areas (1-way ANOVAs, $P > 0.05$ in both cases), and comparable with those from previous studies (Dean 1981; Vogels and Orban 1991; Snowden and others 1992), so that the observed similarities among the areas were not attributable to unusually high noise levels in the data.

We measured the degree of response variance, or “noise,” introduced by jittering of the stimuli. To do this, we carried out a 1-way ANOVA for each cell across the 3 jitter positions. For area V4, the mean $F$ ratio for the jitter factor was 0.84 (1st quartile 0.20, median 0.56; 3rd quartile 1.10), indicating that the response variance across the 3 jitters is $84\%$...
of the average within-jitter response. The mean value of the jitter 

\( F \) ratio was 0.81 (1st quartile 0.21, median 0.62; 3rd quartile 1.07) for area V2 and 0.91 (1st quartile 0.23, median 0.58; 3rd quartile 1.31) for area V1. Thus, for all 3 areas, the average response variance introduced by stimulus jitter was less than the average random trial-to-trial variance of the responses.

**Analyses of Response Correlations**

We analyzed the patterns of response similarities within each of the 3 areas using MDS and principal components analyses (PCAs) (for overviews, see Kruskal and Wish 1978; Dunteman 1989; Kagichian 1991). To carry out these analyses, we constructed a \( 128 \times 128 \) correlation matrix, each element of which represents the correlation coefficient of the responses of cells in a given area (averaged across trials) to a given pair of the 128 stimuli (see Fig. 8 of Hegdê and Van Essen 2003, 2004). We then used MDS or PCA to analyze the global pattern of response correlation in these matrices.

**Analysis of MDS Clusters**

MDS plots the data so that stimuli that elicit similar responses from a given population are clustered together and those that elicit disparate responses are dispersed. MDS clusters were identified independently by visual inspection (see Kruskal and Wish 1978) or using linear discriminant analysis (S-Plus function “lda”); Venables and Ripley 1999, p 344–349). The clusters determined by these methods were identical in terms of cluster membership to those revealed by hierarchical cluster analysis (S-Plus function “agnes”, Venables and Ripley 1999, p 356–359; data not shown). To measure the distortion involved in reducing the high-dimensional matrix into a 2D MDS plot, we calculated the normalized stress \( S_k \) for each dimension \( k \), defined as

\[ S_k = \sum_{ij} (d_{ij} - \hat{d}_{ij})^2 / \sum_{ij} (d_{ij} - \bar{d})^2 \] ,

where \( d_{ij} \) is the distance between any 2 stimuli \( i \) and \( j \) in the correlation matrix, \( \hat{d}_{ij} \) is the distance between the same 2 stimuli in the MDS plot, and \( \bar{d} \) is the mean of all \( d_{ij} \) (Kruskal and Wish 1978). \( S_k \) values were calculated for \( k \) values ranging from \( k = 2 \) (i.e., the conventional 2D MDS plot) to \( k = 128 \) (i.e., a hypothetical MDS plot with the largest dimensionality possible for the current data set). To determine whether the clustering of stimuli, if any, in a given MDS plot was significantly nonrandom, we used the \( D \) ratio test, which was directly analogous to the \( F \) ratio (see Hegdê and Van Essen 2003).

To calculate the correlation between a given pair of MDS plots, we used cophenetic correlation coefficient \( r_c \) (see Hegdê and Van Essen 2003). Like the conventional correlation coefficient, the values of \( r_c \) vary from 1.0 (perfect correlation) through 0.0 (no correlation) to -1.0 (perfect anticorrelation). Also, \( r_c \) is a scale-invariant metric just like the conventional correlation coefficient \( r \) so that 2 matrices which are scalar multiples of each other will have the same \( r_c \) of 1.0.

In a different MDS analysis, we plotted individual cells (as opposed to individual stimuli) according to their response similarities. The input to MDS consisted of a \( 369 \times 369 \) correlation matrix in which each cell represented the correlation coefficient of the responses of a given pair of the 369 cells from the 3 areas to the 128 stimuli.

**Results**

We recorded the responses of cells in area V4 to grating and contour stimuli (see Materials and Methods and Fig. 1) matched to those used previously for area V2 (Hegdê and Van Essen 2000, 2003, 2004). Here we systematically compare the representation of various shape characteristics in the 2 areas, bearing in mind certain limitations inherent in these comparisons, such as the fact that stimuli were scaled by receptive field dimensions and thus were systematically larger in V4 (see Discussion). All the analyses presented in this report, including those from Fig. 2, are original to this report except where indicated otherwise.

**Overlapping Patterns of Shape Selectivity in V4 and V2**

Cells in V2 and V4 varied greatly in the degree of stimulus selectivity. In both areas, some cells were very broadly tuned, whereas others were highly selective for a few stimuli. Qualitative similarities in shape selectivity were evident for narrowly tuned V2 and V4 cells, as illustrated by the examples in Figure 3, as well as for broadly tuned cells.

Figure 3A,B shows a pair of cells (V4 cell in Fig. 3A and V2 cell in Fig. 3B) that preferred large-angle stimuli, were narrowly tuned for contour stimuli, and were poorly responsive to grating stimuli. The V4 cell preferred large and small right angles and acute angles, all at 90° orientation (second row), and gave much smaller responses to angle stimuli at other orientations (rows 1, 3, and 4). The cell responded moderately well to a few other stimuli, mainly large intersections. The V2 cell also preferred large angles, but at orientations different from the V4 cell. This cell was more responsive to many of the larger arcs and less responsive to many of the large intersections. Thus, the 2 cells shared many, but not all, shape selectivities.

This was also true for the pair of cells shown in Figure 3C,D. The 2 cells were similar, in that each responded best to a hyperbolic grating at the lowest spatial frequency, and was largely unresponsive to contour stimuli. The 2 cells differed in that the V4 cell (Fig. 3C) was more broadly tuned for grating stimuli, and the V2 cell (Fig. 3D) responded moderately well to a few smaller contour stimuli. Together, the exemplar cells illustrate the overlap between the observed shape selectivity in the 2 areas and that this overlap was not simply attributable to a lack of shape selectivity in either area.

**The Population Average Response**

Figure 4 shows a scatter plot of the average response of the V4 population versus that of the V2 population to each stimulus in our stimulus set. To calculate the population response of a given area, the responses of each cell from the given area to all 128 stimuli were normalized so that the cell’s responses to its most and the least effective stimuli were 1.0 and 0, respectively, using the formula

\[ R = (R_s - R_{min})/(R_{max} - R_{min}), \]

where \( R_s \) and \( R_t \) are, respectively, the normalized and nonnormalized responses to stimulus \( s \) and \( R_{min} \) and \( R_{max} \) are the responses to the least and the most effective stimuli overall. This normalization ensured that all cells contributed equally to the population average. The normalized responses were then averaged across all cells from the given area.

The V4 population response (y axis) was largely correlated with that of area V2 (correlation coefficient \( r = 0.76 \); degrees of freedom, 127; \( P < 0.05 \)), and had a coefficient of determination, \( r^2 \), of 0.58, indicating that 58% of the V4 population average response could be accounted for as a linear function of that of V2 (solid lines in Fig. 4). However, the residuals of this regression (i.e., V4 population average response not accountable as a linear function of that of V2) were not randomly distributed (Kolmogorov-Smirnov test for Goodness of Fit [KS GOF], \( P > 0.05 \); data not shown), indicating that the underlying relationship between the 2 sets of responses was not completely linear. The nonlinear function that best fit the 2 sets of data had the form

\[ y = 0.06 + 0.8(x^{0.6}), \]

where \( x \) and \( y \) were the population average responses of areas V2 and V4, respectively.

Grating stimuli were on average substantially more effective than contour stimuli in V4. In V2, the difference was quantitatively much smaller (albeit statistically highly significant). As a result, grating stimuli lie mainly above the regression line in Figure 4 (1-tailed t-test; null hypothesis, mean signed distance from the regression line = 0; \( P < 0.01 \)), whereas most contour stimuli lie below the regression line (\( P < 0.01 \)).
We obtained similar results by comparing the aggregate responses of the 2 populations using a 2-way ANOVA (areas × stimuli; unbalanced replicates design) that took the responses of individual cells into account. By this measure, the 2 areas were indistinguishable (P > 0.05 for area and interaction factors), although the responses were significantly modulated across stimuli (P < 0.05 for the stimulus factor). Of course, this does not imply that individual cells from the 2 areas tended to respond similarly on average. Indeed, the responses of random pairs of V4 and V2 cells were not correlated (mean pairwise correlation, 0.057; range, -0.69 to 0.82; standard deviation, 0.21; based on 10⁶ random pairwise correlations). For comparison, the coefficient of correlation between the responses of exemplar cells shown in Figure 3 A,B was 0.32 and that for cells shown in Figure 3 C,D was 0.63.

The next 2 sections involve systematic comparisons of the responses with various subsets of stimuli, using either 1) the peak response (i.e., response to the cell’s most effective stimulus) or 2) the response modulation by a given subset of stimuli. These 2 response measures address different aspects of shape representation; the results from one set of analyses are not necessarily predictive of those from the other.

**Peak Response Analysis: Preference for Various Stimulus Types in V4 and V2**

To compare the effectiveness of various types of stimuli for cells in V4 versus V2, we classified cells from each area according to their preferred grating subclasses (see Fig. 1A) and separately for their preferred contour subclasses (see Fig. 1B). The distribution of V4 cells that preferred the various grating subclasses (Fig. 5A, top) was statistically indistinguishable from that of V2 cells (Fig. 5A, bottom) in terms of either the overall distribution of cells or the distribution of cells with significant preferences for a given grating subclass (2-sample KS GOF tests, P > 0.05 in both cases) indicating that the preference for complex gratings did not differ between V2 and V4. The percentage of cells responding better to non-Cartesian gratings than to Cartesian gratings was 65% for V4 and 61% for V2; this

Figure 3. Response profiles of exemplar V4 and V2 cells. Each panel shows the responses of an individual cell from V4 (panels A and C) or V2 (panels B and D). In each panel, the color of the given stimulus represents the net responses of the cell to the stimulus (averaged across trials) plotted according to the color scale at bottom. The negative firing rates represent suppression of the responses below background levels. The trial-to-trial variation of the responses and the response variances associated with systematic variations in the spatial placement of the stimulus within the receptive field were generally low (not shown; see Materials and Methods for additional information). The receptive field diameters for the cells shown were 4.7° (panel A), 1.6° (B), 4.1° (C), and 1.7° (D); preferred bar lengths were 4.0° (A), 1.5° (B), 3.7° (C), and 1.4° (D); preferred bar orientations were 15° (A), 60° (B), 135° (C), and 90° (D). In this figure, stimulus orientations have been normalized so that the preferred bar orientation, as determined during the manual mapping of the receptive field, is shown as vertical (‘|’) in each panel. The preferred bar orientations determined during the manual mapping occasionally differed from that from the actual recording, usually when the cell was not strongly responsive to bars, as in the case of cell shown in panel (A).
Modulation of the each cell’s responses across all 128 stimuli

To determine the extent to which individual cells convey information about the stimulus set as a whole, we measured the variation, or modulation, of a given cell’s responses across all stimuli, and across specific subsets of stimuli, in order to obtain objective measures of information conveyed by each cell about our stimulus set.

Response Modulation across All Stimuli

To determine the extent to which individual cells convey information about the stimulus set as a whole, we measured the modulation of the each cell’s responses across all 128 stimuli using the OMI. The OMI is in essence the signal-to-noise ratio of the cell’s responses across all stimuli, corrected for deviations from normality (see Materials and Methods for details). The distribution of the OMI values for V4 cells is shown in Figure 6A. The average OMI for all V4 cells was 4.9 (gray arrow), indicating that on average, the modulation of responses of V4 cells across all stimuli was 4.9-fold greater than chance level fluctuations. This modulation was statistically significant ($P < 0.05$) for most V4 cells (102/108, 95%; filled bars), and the mean OMI value for these 102 cells was 5.1 (black arrowhead). For area V2, the mean and the median OMI values were 4.03 and 2.72, respectively, and the OMI values were statistically significant ($P < 0.05$) for about nine-tenths of the cells (163/180, 91%; data not shown; see Fig. 5A of Hegde and Van Essen 2003).

Are V2 and V4 Discriminable by Their OMI Values?

To determine whether the 2 areas can be discriminated by their OMI values, we used the ROC analysis, which measured the reliability with which an ideal observer could discriminate between the 2 areas using OMI values (see Fig. 2B and Materials and Methods for details). The area under the ROC curve resulting from this analysis, AR_Comi, was 0.59, which was significantly larger than the chance value of 0.5 ($P < 0.05$). Thus, the OMI values for the population are significantly larger in V4 than in V2. OMI values from V4 were also significantly larger than those from V2 using the nonparametric Wilcoxon rank-sum test and differed significantly using the 2-tailed Kolmogorov–Smirnov test ($P < 0.05$).

We repeated the above analysis using the SI, which is essentially a measure of the peakedness of the cell’s response profile (as distinct from the RMIs, which are proportional to the stimulus-to-stimulus variance of the response profile; see Materials and Methods for details). The distribution of the SI values for V4 cells is shown in Figure 6B. The mean SI value for all 108 cells was 0.54 (gray arrowhead). The SI values were statistically significant for most V4 cells (102/108, 94%; filled bars), and the mean SI value for these 102 cells was 0.55 (black arrowhead).

The SI values of V4 cells were significantly correlated with the OMI values (correlation coefficient $r = 0.74$; $P < 10^{-6}$ by randomization), suggesting that the 2 indices measured related aspects of shape selectivity. Furthermore, the SI values were not significantly different between V2 and V4 by ROC analysis (AR_SI = 0.47), Wilcoxon rank-sum test, or Kolmogorov–Smirnov test ($P > 0.05$ in all cases).

For each stimulus subset for which SI was calculated (see Materials and Methods), the SI values were correlated with the corresponding response modulation index values at $P < 0.0001$ (randomization test), indicating that the strong correlation between the 2 indices was not limited to a subset of stimulus characteristics. In addition, the differences between the 2 areas were more often evident ($P < 0.05$) by RMIs than by the corresponding sparseness indices. Therefore, RMIs were used for the analyses in the next 2 subsections.

Response Modulation across Subsets of Shape Stimuli

Gratings and Contours

The greater response modulation in V4 for the entire stimulus set revealed by the OMI analysis raises the question of whether
the response modulation is greater in V4 for particular stimulus classes. As shown in Figure 7, we compared the modulation of responses in V4 versus V2 for gratings and contour stimuli separately using RMIs $RMI_{\text{grat}}$ and $RMI_{\text{cont}}$, respectively, also based on the $F$ ratio (see Materials and Methods for details).

The response modulation across gratings, as measured by $RMI_{\text{grat}}$, was marginally higher in V4 (Fig. 7A, top) than in V2 (bottom). ROC analysis indicated that the 2 areas are not discriminable by the extent to which they convey information about grating stimuli ($ARC_{\text{RMIgrat}}$, 0.54; $P = 0.12$). On the other hand, 2 areas differed significantly in terms of the information they conveyed about the contour stimuli; response modulation across the contour stimuli was reliably higher in V2 relative to that in V4 ($ARC_{\text{RMIcont}}$, 0.39; $P = 0.001$). Similar results were obtained using the Wilcoxon rank-sum test ($P = 0.26$ for $RMI_{\text{grat}}$ and $P = 0.0007$ for $RMI_{\text{cont}}$) and the Komogorov-Smirnov test ($P = 0.10$ for $RMI_{\text{grat}}$, and $P = 0.02$ for $RMI_{\text{cont}}$).

### Response Modulation within Stimulus Subclasses

We extended the modulation index analyses to a finer-grained level by comparing the response modulation in the 2 areas across each of the 4 grating stimulus subclasses and 10 contour stimulus subclasses (defined in Fig. 1) individually. For each of the 4 grating subclasses, we calculated 3 (or fewer for some subclasses, see below) different WMI s, each based on the $F$ ratio (see Materials and Methods). The results are shown in Table 1. $WMI_{\text{all}}$ for each subclass measured the overall response modulation across all the stimuli in the subclass. For the sinusoidal and hyperbolic gratings, response modulation by orientation and spatial frequency was measured by $WMI_{\text{ori}}$ and $WMI_{\text{sf}}$, respectively (see footnote [a] in Table 1). In each case, we calculated an ARC value to measure the extent to which V2 and V4 were discriminable by the response modulation measured by the given WMI.

For the sinusoids, the average $WMI_{\text{ori}}$ values for V4 and V2 cells were 2.58 and 2.46, respectively (far left data column, lines 1-3). Similarly, the average $WMI_{\text{sf}}$ values for V4 and V2 cells were 2.58 and 2.46, respectively (far left data column, lines 4-5). The response modulation across the 2 classes of sinusoids was reliably higher in V2 relative to that in V4 ($ARC_{\text{WMIori}}$, 0.39; $P = 0.001$) and the Komogorov-Smirnov test ($P = 0.10$ for $WMI_{\text{ori}}$, and $P = 0.02$ for $WMI_{\text{sf}}$).

**Table 1.** Response Modulation within Stimulus Subclasses

<table>
<thead>
<tr>
<th>Subclass</th>
<th>V4</th>
<th>V2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sinusoidal</td>
<td>2.58</td>
<td>2.46</td>
</tr>
<tr>
<td>Hyperbolic</td>
<td>2.58</td>
<td>2.46</td>
</tr>
<tr>
<td>Concentric-like</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
1 and 2), signifying a response modulation averaging about 2.5 times higher than chance in both areas. The 2 areas were not discriminable in terms of the WMIall values (ARCWMIall: 0.52 [line 3]; \( P > 0.05 \)). V2 cells conveyed more information than V4 cells about sinusoid orientation (ARCWMIori: 0.41; \( P < 0.05 \), open box), whereas V4 cells conveyed more information than V2 cells about sinusoidal spatial frequency (ARCWMIsf: 0.57; \( P < 0.05 \), hatched box). Comparable percentages of cells in the 2 areas conveyed significant information (\( P < 0.05 \)) about sinuosids as measured by each index (rightmost 3 columns), indicating that the differences between areas was not attributable to general unresponsiveness to sinuosids. A similar pattern of greater modulation across orientations in V2 and greater spatial frequency modulation in V4 occurred for response modulation by hyperbolic gratings (lines 5–8). The 2 areas were indistinguishable in terms of the response modulation by polar gratings (lines 9–16).

We carried out similar analyses for the 10 contour stimulus subclasses (Table 2). For each subclass, WMIall measured the overall response modulation across the entire subclass, and WMIori and WMIsize measured the response modulation by, respectively, orientation and the size of the stimuli within each subclass (see footnote [a] in Table 2). Two notable trends emerge from this analysis. First, the 2 areas showed reliable differences in the information they conveyed about each subclass (boxes), except for the stars/circles subclass (lines 13–16). Second, the response modulation was greater in V2 than in V4 for each measure by which the 2 areas were discriminable (open boxes), except for the size of obtuse angles (hatched box). This implies that V2 cells generally conveyed greater information than did V4 cells about contour shape characteristics, including complex contours as well as simple contours (i.e., bars). Nonetheless, substantial proportions of V4 cells conveyed information about each contour shape characteristic.

Figure 6. Measures of response selectivity of individual V4 cells across all 128 stimuli. For each cell, the modulation of responses across the stimulus set as a whole was measured using the response modulation index (OMI) and the SI as described in Materials and Methods. This figure shows the distribution of the OMI values (panel A) and the SI values (panel B) for the 108 V4 cells; the corresponding distributions for the V2 cells are illustrated in Hegdé and Van Essen (2003). In panel (A), outliers with OMI > 10 are rounded out to 10. In both panels, the filled bars represent those cells for which the corresponding index values were statistically significant (\( P < 0.05 \)) as determined by the randomization. Cells with \( P > 0.05 \) are denoted by open bars. The gray arrow and the black arrowhead denote the average index value for all cells, and for cells with \( P < 0.05 \), respectively, each calculated before the outliers were rounded out.

Figure 7. Response modulation by gratings and contours in areas V4 and V2. Modulation of each cell’s responses across all grating stimuli and all contour stimuli was measured using RMIgrat and RMIcont, respectively, as described in Materials and Methods. This figure shows the distribution of the index values (A RMIgrat; B RMIcont) for both V4 (top) and V2 (bottom). The gray arrow and the black arrowhead denote the average index value for all cells, and for cells with \( P < 0.05 \), respectively, each calculated before the outliers were rounded out.
The mean values shown in Tables 1 and 2 are based on conventional arithmetic means. Because some of the WMI values in either table showed positively skewed distributions, we also calculated the geometric mean values for all the WMIs (see the corresponding tables in the Supplemental Material).

### Selectivity for Complex Shape Characteristics in Area V1

To assess whether selectivity for complex shape characteristics is absent in V1 and emerges de novo in V2, we analyzed the responses of 81 V1 cells in one animal. The stimulus set included all stimuli in Figure 1 plus 2 dozen additional sinusoidal and bar stimuli in order to sample phase, spatial frequency, and orientation more densely (see Fig. 8 and Materials and Methods).

Given the small size of V1 receptive fields, we assessed the extent to which the responses of V1 cells were sensitive to the positioning of the stimuli over the CRF. To do this, we compared the responses of each cell across the 3 jittered positions of the stimulus (see Materials and Methods) using a 2-way ANOVA (stimulus type × stimulus position). The effect for the stimulus position was statistically significant ($P < 0.05$) for only 4 (5%) of the 81 cells and for only 3 (4%) cells using the responses to the common set of 128 stimuli. The stimulus type–stimulus position interaction factor was significant for 3 cells (4%) using all 152 stimuli and 5 cells (6%) using the common set of 128 stimuli. Thus, for most V1 cells, the responses were not attributable to a fortuitous positioning of the preferred stimulus over a receptive field ‘hot spot.’ By comparison, the stimulus position factor was significant for 5 cells (3%) in V2 and 6 cells (6%) in V4, and the interaction factor was significant for

---

**Table 1**

<table>
<thead>
<tr>
<th>Subclass</th>
<th>WMI&lt;sub&gt;s&lt;/sub&gt;</th>
<th>WMI&lt;sub&gt;all&lt;/sub&gt;</th>
<th>% Cells at $P &lt; 0.05^b$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sinusoids V4</td>
<td>2.58</td>
<td>2.59</td>
<td>3.73</td>
</tr>
<tr>
<td>V2</td>
<td>2.46</td>
<td>3.96</td>
<td>2.60</td>
</tr>
<tr>
<td>ARC</td>
<td>0.52</td>
<td>0.41</td>
<td>0.57</td>
</tr>
<tr>
<td>Bars V4</td>
<td>2.84</td>
<td>3.04</td>
<td>2.72</td>
</tr>
<tr>
<td>V2</td>
<td>4.30</td>
<td>6.78</td>
<td>1.97</td>
</tr>
<tr>
<td>ARC</td>
<td>0.43</td>
<td>0.39</td>
<td>0.54</td>
</tr>
</tbody>
</table>

---

**Table 2**

<table>
<thead>
<tr>
<th>Subclass</th>
<th>WMI&lt;sub&gt;s&lt;/sub&gt;</th>
<th>WMI&lt;sub&gt;all&lt;/sub&gt;</th>
<th>% Cells at $P &lt; 0.05^b$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bars V4</td>
<td>2.84</td>
<td>3.04</td>
<td>2.72</td>
</tr>
<tr>
<td>V2</td>
<td>4.30</td>
<td>6.78</td>
<td>1.97</td>
</tr>
<tr>
<td>ARC</td>
<td>0.43</td>
<td>0.39</td>
<td>0.54</td>
</tr>
</tbody>
</table>

---

^aMean modulation indices. WMI<sub>s</sub> measured the response modulation across all stimuli within a given subclass, whereas WMI<sub>all</sub> and WMI<sub>ori</sub> measured the modulation of the responses to the most effective stimulus within the given subclass by its spatial frequency or orientation, respectively. 
^b denotes the cases in which the corresponding index value was reliably larger for V2 than for V4 as measured by ROC analysis (ARC < 0.5; $P < 0.05$). 
^c denotes the opposite, that is, V4 > V2; ARC > 0.5; $P < 0.05$. See Materials and Methods and Figure 2B for details of the ROC analyses. 
^denotes instances in which the 2 areas were different by a 2-tailed Kolmogorov–Smirnov (K-S) test (V4 ≠ V2; $P < 0.05$). 
^d denotes instances in which the 2 areas were different by a 2-tailed Kolmogorov–Smirnov (K-S) test (V4 ≠ V2; $P < 0.05$). 
^e denotes instances in which the 2 areas were different by a 2-tailed Kolmogorov–Smirnov (K-S) test (V4 ≠ V2; $P < 0.05$).

---

Cerebral Cortex May 2007, V 17 N 5 1109
2 cells each in V2 and in V4 (2% and 1% cells, respectively). This indicates that the effect of stimulus jitter was comparable across the 3 areas, which may be attributable to the fact that the average stimulus jitter in all 3 areas was smaller than the fixation window (see Materials and Methods).

Figure 8 illustrates the responses of 2 V1 cells that showed clear preferences for complex contours over simple bars or gratings. The cells were recorded from the calcarine cortex and had relatively large receptive fields, so that the stimuli were likely located largely, if not exclusively, within the CRF throughout the presentation (see legend for details). The cell shown in Figure 8A was sharply selective for the smaller tristar at 30° (second row), which elicited 78 spikes/s. The second most effective stimulus, the low-spatial frequency sinusoid at 45° (second row) elicited less than half of this response (37 spikes/s). Responses to the bar stimuli were even lower (response range, 2–15 spikes/s), indicating that the cell’s response to the tristar stimulus was not readily predictable from the responses to the simple bar stimuli. Responses to gratings stimuli were generally low (response range, 5–37 spikes/s). The response of this cell to its most effective stimulus and to its preferred bar did not vary significantly across the 3 jitter positions (Fig. 8B,C; see legend for details). The mean eye position of the animal did not vary significantly across the 3 jitter positions (Fig. 8D).

For the exemplar cell shown in Figure 8E, the most selective stimuli were the large 3-quarter arcs at 0° and 180° (first and third rows). The responses to these curved stimuli were not readily predictable from its orientation selectivity for bar stimuli. Neither the responses to preferred stimuli (Fig. 8E,G) nor the eye positions (Fig. 8H) varied significantly across jitter positions for this cell.

Quantitative Comparisons of Shape Selectivity in V1 versus V2

The above results confirm that selectivity for complex shape characteristics can occur in area V1. To explore this issue more systematically, we repeated for the V1 data set each of the analyses we carried out for V4 and V2, using the responses to the 128 stimuli common to all 3 areas (see Materials and Methods).

Figure 9A shows the distribution of V1 cells according to the subclass of their preferred grating. The proportion of V1 cells preferring a sinusoid (35/81, 43%) was indistinguishable from the proportions of cells preferring a non-Cartesian grating (binomial proportions test, P = 0.05), indicating that selectivity for complex gratings was just as prevalent in V1 as for the selectivity for simple gratings.

The distribution of cells that preferred the various contour stimulus subclasses is shown in Figure 9B. Notably, the proportion of V1 cells with a statistically significant preference for a given contour subclass (denoted by filled bars; see Fig. 5 for significance criteria) was substantially larger for many complex contours than for bar stimuli. In particular, the preference for a 3-quarter arc was statistically significant for about 14% (11/81) of the cells, whereas for bar stimuli only a single cell showed a significant preference. Both subclasses elicited preferred responses from equal numbers of V1 cells (12/81, 15%). The distribution of V1 cells among both grating and contour subclasses was indistinguishable from that of V2 or for V4 (see Fig. 5; 2-sample KS GOF tests, P > 0.05 in all cases).

To determine whether the response modulation across simple bar stimuli is more pronounced in V1 than in V2, we compared the WMI_{all\,bar} values for the 2 areas (Fig. 9C). For cells which conveyed significant information about bars (i.e., with WMI_{all\,bar} at P < 0.05; filled bars in Fig. 9C), the WMI_{all\,bar} was reliably higher in V2 than in V1 (ARCWMI_{all\,bar}, 0.40; P = 0.02), although the 2 areas were indistinguishable when the ROC analysis was carried out across all cells regardless of the statistical significance of their WMI_{all\,bar} values (ARCWMI_{all\,bar}, 0.48; P = 0.28). This indicates that the response modulation by simple contours did not decrease from V1 to V2.

Figure 9D compares the response modulation by tristar stimuli in V1 versus V2. Tristars were the least effective contour stimulus type for V1 cells (see Fig. 9B) but were more effective in V2 (see Fig. 5B, bottom). Thus, an increase in complex contour sensitivity from V1 to V2 would arguably be most likely to occur for this stimulus type, if any. However, the information conveyed about the tristar stimuli, as measured by WMI_{all\,tristar} values, was statistically indistinguishable between the 2 areas, either when compared across all cells in either area (ARCWMI_{all\,tristar}, 0.47; P = 0.22) or for cells with significant response modulations across tristar stimuli (i.e., with WMI_{all\,tristar} at P < 0.05; filled bars in Fig. 9D, ARCWMI_{all\,tristar}, 0.46; P = 0.25). Similar analyses of other stimulus subclasses revealed no evidence of any systematic increase in selectivity for complex contour stimuli from V1 to V2 (data not shown).

Together, these results indicate that shape representation at the level of individual cells also varies in a graded fashion between V1 and V2 as it does between V2 and V4.

Higher Order Patterns in the Population Response

We previously reported that V2 cells as a population tend to respond similarly to specific subsets, or clusters, of stimuli, so that stimuli within a given cluster elicit similar responses from the population, but different clusters elicit disparate patterns of population response (Hegdé and Van Essen 2003). These clusters of response correlation are of interest because extracting similarities, or “common denominators,” among visual objects may allow the visual system to reduce the complexity of visual representation with minimal loss of information, analogous to how stenography represents words with fewer symbols yet with low ambiguity (Hegdé and Van Essen 2003; also see Seung and Lee 2000). In other words, to the extent a given population responds similarly to a given set of stimuli, it is less sensitive to differences among them.

We compared the population responses across all 3 areas. For area V4, this entailed construction of a 128 × 128 correlation matrix of the V4 population response, each element of which represented the correlation coefficient of the responses of all V4 cells to a given pair of the 128 stimuli (see Materials and Methods; also see Fig. 8 of Hegdé and Van Essen 2003, 2004). Separate correlation matrices were generated for areas V2 and V1. Each correlation matrix was analyzed for patterns of response correlation using metric MDS and PCA as described previously (Hegdé and Van Essen 2003, 2004; also see Materials and Methods). The results from the 2 analyses were similar, as expected (see Kachigan 1991). Key results from the MDS analysis are presented below.

Patterns of Response Correlation: MDS Analysis

MDS plots the data points, that is, the 128 stimuli in the current context, so that stimuli which consistently elicited correlated responses from a given cell population are clustered together, and stimuli which tended to elicit disparate responses are
dispersed from each other. No clustering is expected if the response correlations vary randomly among stimuli. The results for area V4 are shown in Figure 10A (top). The grating and contours were segregated into 2 distinct clusters. This clustering was highly significant (D ratio test, 0/1000 rounds passing the criterion; see Materials and Methods). No significant subclusters were identifiable within either cluster. Importantly, the 2 statistically significant clusters of contour stimuli that were identifiable in the MDS plot of V2 (Fig. 10A, middle panel; highlighted in green and blue in all panels) were indistinguishable in case of V4. Thus, the stimulus clustering seen in V2 was differentially amplified in V4, with the grating-contour distinction becoming more prominent and the distinction between the 2 contour clusters largely disappearing. This is correlated with the greater separation in average responses to gratings versus contours in V4 compared with V2 (cf., Fig. 4). This suggests that some useful shape discrimination information represented in V2 is not retained to the same degree in V4. Similar results (not shown) were obtained when the MDS analyses were repeated in each area using only the responses to contour stimuli, indicating that the absence of this cluster in V4 was not because this clustering was eclipsed by the stronger grating-contour distinction in V4.

The grating stimuli form a significant cluster of their own in both V4 and V2. However, the grating clusters in the 2 areas were substantially different (coefficient of cophenetic correlation $r_C$, 0.19; $P > 0.05$; see Materials and Methods for more information), indicating that the representation of the grating stimuli in V4 is not a scaled version of that in V2 or vice versa. Similarly, the contour representation in V4 was not a scaled version of that in V2 or vice versa ($r_C$, 0.07; $P > 0.05$), although this fact is also evident from the MDS plots.

The MDS for area V1 is shown in Figure 10C. The mean values for the grating and contour stimuli differ

Figure 8. Exemplar V1 cells. (A) The response profile of an individual V1 cell plotted using the same conventions as in Figure 3. Note that the stimulus used for V1 contained 24 additional sinusoidal and bar stimuli relative to those used for V4 and V2. (B) The placement of the cell’s most effective stimulus in the 3 jitter positions (arbitrarily colored red, green, and blue) within the cell’s CRF (dashed circle). The magenta X symbol denotes the center of the CRF, the bracket denotes 12.5% of the CRF diameter, by which the stimulus placement was jittered. The response elicited by the stimulus in each position is shown by the correspondingly color-coded peristimulus time histogram (PSTH). The long horizontal bar underneath each PSTH denotes the 300 ms stimulus presentation, and the shorter bar below it denotes the time window using which the evoked firing rates were calculated. (C) The placement of, and responses elicited by, the cell’s preferred bar at the 3 jitter positions. (D) The mean eye position (±standard deviation) during stimulus presentation in a given jitter position. The outer square denotes the fixation window, and the intersection of the dotted lines denotes the location of the fixation spot. Panels (E–H) show the corresponding data for another exemplar V1 cell. Both cells shown were recorded from calcaneal V1 (Receptive field eccentricities: A–D, 11.7°; E–H, 12.1°). Receptive field diameters: A–D, 1.5°; E–H, 1.7°. Preferred bar lengths: A–D, 1.2°; E–H, 1.3°. For both cells, the responses were indistinguishable across the 3 jitter positions (1-way ANOVA, jitter factor; A–D, $P = 0.58$; E–H, $P = 0.55$), as were the eye positions across the 3 jitter positions (1-way ANOVA, jitter factor; D, $P = 0.27$; H, $P = 0.81$). See Materials and Methods for additional information.
significantly, but not nearly as much as for V2 or V4, and there is substantial intermixing between stimuli in the 2 groups.

Some distortion (or stress) is unavoidable when a complex, high-dimensional data set is represented in a 2D MDS plot (Kruskal and Wish 1978; Venables and Ripley 1999). To determine the manner in which the magnitude of the distortion increases as the dimensionality of the representation decreases, we measured the normalized stress ($S_k$) as a function of the number of dimensions ($k$) in the MDS representation for each area (see Materials and Methods). For area V4 (Fig. 10B, top), the stress remains relatively constant as number of dimensions is reduced from a possible 128 dimensions to about 40, but increases rapidly thereafter, indicating that the smallest number of dimensions needed for a relatively stress-free representation of this data set is about 40 (see Kruskal and Wish 1978, p 53–60). It also indicates that a 2D representation of this data set (Fig. 10A, top) involves considerable distortion. However, the magnitude of distortion is comparable across the 3 areas (Fig. 10B; $S_k$ values for the plots shown: V4, 0.26; V2, 0.24; V1, 0.21), indicating that the qualitative differences among the 3 MDS plots are unlikely to be primarily attributable to distortion.

Qualitatively similar results were obtained when the MDS analyses were repeated for each area using random subsets of cells (not shown). This suggests that shape representation among subpopulations was similar to that in the corresponding overall population in each area.

Figure 11 shows an MDS plot of individual cells from the 3 areas plotted according to the similarity of the response profiles, so that cells that responded similarly to the 128 stimuli were clustered together, and cells with disparate response profiles were dispersed away from each other (see Materials and Methods for details). Cells from the 3 areas were extensively intermingled with each other. While there are suggestive biases in the distribution from visual inspection (e.g., more V2 and V1 cells near the top and bottom right; more V4 cells in the lower left and far right), no significant clustering was revealed by a $D$ ratio test (526/1000 rounds passing the criterion; see...
Discussion

This study provides a detailed comparison of shape representation in areas V2 and V4 using a diverse set of stimuli that were presented within the CRF. The comparisons reveal numerous similarities as well as many differences in detail between areas. Additional results from area V1 also reveal a combination of major similarities coupled with some differences. Importantly, the data do not reveal a markedly greater preference for complex stimuli in V4 compared with V2 or in V2 compared with V1. Our analyses reveal pronounced areal differences in neither the overall sharpness of stimulus selectivity nor the degree of response modulation. The differences are largely of degree, rather than of kind.

The MDS analyses of the population responses provide some of the clearest evidence for differences of shape representation among the 3 areas. However, these differences do not follow a systematic (e.g., simple to complex) progression from V1 to V2 to V4. Nonetheless, these results raise the possibility that the differences among the 3 areas may be most evident in terms of population-based coding of shape characteristics.

The collective import of these results is not that there are no differences among the areas, but rather that the many differences that do exist fail, individually or together, to clearly parallel the anatomical hierarchy. Thus, even if one were to discount the many similarities among the areas and consider only the differences, the fact remains that our results reveal no

Figure 10. Comparison of the population responses in areas V4, V2, and V1. (A) For each area, the stimulus set was plotted using metric MDS, so that the proximity between any given pair of stimuli within a given plot reflects the similarity of the population response they elicited. To facilitate comparison across areas, the 3 plots are all drawn to the same scale and the same relative orientation. The arrow in each panel denotes the preferred bar. The stimuli are coded in red, green, and blue according to the 3 clusters into which they segregate in the MDS plot for area V2. (B) Normalized stress $S_k$ as a function of different hypothetical dimensionalities $k$ of MDS representation, calculated prior to the aforementioned rescaling of the plots. See Results for details.
difference/s that follow/s a clear stepwise pattern from V1 to V2 to V4. On the other hand, neither do our results mean that there are no functional differences among the 3 areas that follow a hierarchical pattern, but only that our study reveals none (see below).

It is intriguing that our results indicate that the selectivity for simple versus complex shape characteristics is comparable within V1 and that selectivity for complex shape characteristics is comparable in V1 versus V2. These results may reflect genuine complexities of V1 CRFs (Rust and others 2005; Pack and others 2006). Alternatively, given that our stimuli may have stimulated the non-CRF, at least part of the selectivity for complex shape characteristics may reflect nonlinear influences of the nonclassical surround. Also because our results were obtained from a single animal, the characteristics reported here may be unrepresentative of V1 neurons in a population of animals.

**Limitations of Interareal Comparisons**

Important caveats apply to the interpretation of both the similarities and differences among visual areas. First, it is possible that low-level differences in stimuli, such those related to the differences in the average size, retinal resolution, luminance, and/or contrast of the stimuli account for some of the interareal differences in the MDS clusters and of other relevant response measures. Second, while we aimed to minimize the differences in experimental parameters, some differences were unavoidable, ultimately due to the inherent differences of receptive field sizes and preferences across areas. We scaled the stimuli to the match receptive field sizes in order to drive the cells well, but this unavoidably introduces size-related differences in the stimuli. Other possible approaches to addressing this issue, for example, holding the (average) stimulus size constant across areas, raise their own sets of problems. Another limitation is that even though our stimuli elicited significant responses and response modulations from each area, our stimulus set may not have been optimal for revealing important differences among areas (see below). In practical terms, no single experimental approach can provide a comprehensive comparison across areas.

Two general issues arise in considering the significance of these findings for understanding hierarchical processing in visual cortex. One entails relating the similarities in shape representation to the fact that receptive field sizes differ markedly across areas. Another is to place our findings into the context of other studies of receptive field properties in extrastriate cortex, some of which suggest prominent areal differences.

**Shape Selectivity at Multiple Scales**

At any particular eccentricity, average receptive field diameters approximately double between V1 and V2 and again between V2 and V4 (Gattass and others 1981, 1988). Given this, it follows that position-independent selectivity for a complex shape in a higher area (and hence at a coarser scale) cannot be preserved just by passive relaying of comparable selectivity that has been established at a lower cortical level. For example, a V4 neuron that prefers concentric gratings matched to its receptive field size cannot attain this selectivity simply by receiving convergent inputs from a population of concentric-preferring cells in V2. Instead, complex neural circuits are needed in order to generate shape selectivity at each spatial scale. The nature of this circuitry might in principle be similar at every stage where a given type of selectivity occurs (aside from the scale differences). Alternatively, there might be important differences across levels (e.g., concentric-preferring V4 neurons might have subunits generated by selective inputs from curvature-selective cells in V2, whereas concentric-preferring V2 neurons might have subunits generated by selective inputs from conventional orientation-selective neurons in V1). Physiological, psychophysical, and computational studies that characterize receptive field substructure, including spatial inhomogeneities in orientation selectivity and other features, should help clarify these issues (Wilson and Wilkinson 1998; Anzai and Van Essen 2002; Pollen and others 2002; Loffler and others 2003). Regardless of whether the neural circuitry is broadly similar across levels or is highly level specific, the task of generating a diversity of complex shape selectivities at multiple spatial scales presumably requires extensive neural resources at each hierarchical level.

The fact that the spatial scale of shape analysis varies among the 3 areas has 2 important practical implications for interareal comparisons. First, it means that similarities of response measures do not imply that shape analysis is redundant across areas (e.g., the circuits for generating certain properties might be similar in 2 areas, but the representations will differ to the extent that the inputs differ). Second, it highlights the difficulties of the experimentally comparing shape representation across the 3 areas because the optimal stimuli will necessarily differ among the areas.

**Evidence from Other Studies**

To what degree do cells in progressively higher areas display properties that are largely or entirely absent in lower areas? Although we found relatively few differences for the types of form selectivity explored by our stimulus set, studies focusing on other receptive field properties both within and outside
the CRF suggest important differences across hierarchical levels, especially between V1 and V2. Examples of characteristics encountered in V2 but absent or less prevalent in V1 include selectivity for 1) relative rather than absolute retinal disparity (Thomas and others 2002); 2) stereoscopic edges (von der Heydt and others 2000); 3) 3-dimensional surface configurations (Bakin and others 2000); 4) border ownership (also in V4) (Zhou and others 2000); and perhaps 5) figural patterns defined by orientation contrast or illusory contours (Marcus and Van Essen 2002). Other studies have reported that responsiveness to illusory contours is evident in V2 but not in V1 (Peterhans and von der Heydt 1989; von der Heydt and Peterhans 1989), although subsequent studies have reported some illusory contour responsiveness even in V1 (Sheth and others 1996; Lee and Nguyen 2001; Ramsden and others 2001). Mahon and De Valois (2001) reported substantial selectivity for non-Cartesian gratings in area V1 but a modestly greater degree of non-Cartesian selectivity in V2, an outcome consistent with our results. Victor and others (2005) reported that many cells in V1 have responses to polar 2D Hermite functions that are not readily predictable from responses to Cartesian 2D Hermite functions.

Examples of properties encountered in V4 but overtly shown to be absent or sparse in V2 are rare. Selectivity for object-centered contour shapes (Pasupathy and Connor 1999, 2002), object saliency (Mazer and Gallant 2003), direction of elemental luminance gradients in texture patterns (Hanazawa and Komatsu 2001), and motion-dependent distortion of retinotopy (Sundberg and others 2006) have been reported in V4, but tests for these properties have not been reported for V2. Using a diverse set of simple and complex shapes customized for each cell, Kobatake and Tanaka (1994) reported a higher incidence of cells selective for various "critical features" in V4 compared with V2, but the differences are difficult to characterize systematically. Top-down influences, most notably attentional effects, are more pronounced in higher visual areas than in lower visual areas (Luck and others 1997; Marcus and Van Essen 2002; Maunsell and Cook 2002; also see Hochstein and Ahissar, 2002), and differences between areas might be more pronounced during natural viewing conditions (David and others 2004). In general, quantitative comparisons across areas using appropriately designed stimulus sets, such as those used in this study, are needed to determine which response properties emerge de novo or are substantially enhanced in any given visual area. It is also possible that differences between areas not evident at the level of conventional single unit response measures may be revealed by population analyses involving more abstract representations (such as that explored by the MDS analyses) or involving temporal correlations between neurons.

Concluding Remarks

The notion of an anatomically based hierarchy of cortical organization provides only an indirect basis for hypothesizing functional relationships and stages of information processing. In the realm of temporal processing, latency differences to stimulus onset are not strongly correlated with anatomical hierarchical level (Schmolsensky and others1998; Vanni and others 2004), but this is likely to reflect factors such as the relative contributions of magnocellular versus parvocellular inputs to different processing streams as well as the fact that numerous anatomical pathways jump across multiple hierarchical levels. In terms of shape representation, the present study indicates that some properties undergo relatively little change in the nature of the representation other than spatial scale, even though they are likely to be central to the functions of V2 and V4. Attaining a better understanding of which properties change with hierarchical level and how such changes are related to the underlying forward, feedback, and intrinsic cortical circuitry remains an important arena for investigation.

Supplementary Material

Supplementary material can be found at: http://www.cercor.oxfordjournals.org/.

Notes

This work was supported by the National Institute of Health grant EY02091 to DCVE. JH wishes to thank Drs Thomas Albright and Gene Stoner for advice and support during the preparation of this manuscript. We thank Drs Odelia Schwartz, Paul Shrater, and Terrence Sejnowski for helpful discussions, and Drs Leanne Chukoskie, Greg Horwitz, Xin Huang, Bart Krekelberg, and Anja Schlack for useful comments on various drafts of the manuscript, and Susan Danker for assistance with manuscript preparation. Conflict of Interest: None declared.

Address correspondence to David Van Essen, Department of Anatomy and Neurobiology, Washington University School of Medicine, Box 8108, Saint Louis, MO 63110, USA. Email: vanessen@brainvis.wustl.edu.

References


